

CBD finder the bully users according to the user behavior Using ML

A. Ramaswami Reddy

Professor, Computer Science Engineering, Malla Reddy Engineering College, Maisammaguda, Hyderabad.

Abstract

In terms of SM (Social Media) management, tracking bully users across Social Media platforms is a complex and important job. In theory, we explorations allow a bird's-eye view of SM user behaviors. However, nearly all recent SM-based studies focus on a different types of data like posts, profile media, these are yielding incomplete data. Therefore, this study investigates the strategy of Cyber Bullying Detection (CBD) find the bully users according to the user behavior based on the chat history, we deeply mined friend relationships. In thispaper, we have calculated the user behavior according to the chat statements of the user which he/she do with others. For Detection of bully users, we approached using Machine Learning models and conducted few experimental results. We conducted classification analysis on the Harassment-Corpus Dataset. In these results, we achieved 97% accuracy with Neural Network algorithm for detection of bully words.

Keywords: - CBD, bullywords, Python, ML

1. INTRODUCTION

Most chat applications in the internet like WhatsApp, Messenger or other social networking apps offering chat communication tools for text messages, media data sharing, web data sharing etc. In current trend many social networking sites created and providing services of communications, multi-media services, e-commerce etc immensely. For example twitter social media provide major services of micro-blogging massively, it has more than 700 million users and 400 million micro-blogs produces per day. According to research survey many more than 30% of

dummy or duplicate or fake accounts are present in all social media services like twitter, facebook, sinaetc [1]. Using fake profiles, the bully users publish the bully data in the social networks. But in the current social sites not focus on services like tracking the user bully behaviour. In current system, social network sites need to focus the user microblogs and need to capture the user behavior whether his/she bully user or not. Few surveys' providing concepts to tracking the attackers like using profile matching techniques and network based techniques etc. But in real-time to apply those concepts in social network is less

practical. Crawling the user information from the user micro blogs is also less practical. Bully Users can easily manipulate the public profile information. In social networks user may share their messages by using the chat applications. For every social networking sites has their own chat applications, for this facebook is main example. And another way is sharing the multimedia data like images or videos. For this Facebook and Instagram best examples. For communication between users chat applications will most useful for share their information, thoughts, views etc. But in the same way it may also cause the security loophole of user's security which is cyber bullying. Such text based content may security threat to the users because of the people can share cyber bullying words to the users with their fake accounts. [2], [3]. Based on these disadvantages detection malicious users is active topic in the study of social media.

2. RELATED WORK

In current trend many social networking sites created and providing services of communications, multi-media services, e-commerce etc immensely. For example twitter social media provide major services of micro-blogging massively, it has more than 700 million users and 400 million

micro-blogs produces per day. According to research survey many more than 30% of dummy or duplicate or fake accounts are present in all social media services like twitter, facebook, sinaetc [1]. But in the current social sites not focus on services like tracking the user behavior based on the published data. In current system, social network sites need to focus the user microblogs and need to capture the user behavior by predicting the bully data or normal data.

Few surveys' providing concepts to tracking the attackers like using profile matching techniques and network based techniques etc. But in real-time to apply those concepts in social network is less practical. Crawling the user information from the user micro blogs is also less practical. Bully Users can easily manipulate the public profile information.

In current trend many social networking sites created and providing services of communications, multi-media services, e-commerce etc immensely. Lot of bully users accounts are creating very rapidly. We need to focus for the tracking the bully users. In our proposed system we are implementing the web based application which will find the bully users according to the user behavior. We calculate the user

behavior according to the chat statements of the user which he/she do with others. By the taking the advantages Machine Learning algorithms we classify the bully users and bully data. Here we are using Naïve Bayes algorithm to perform the classification of the users.

3. IMPLEMENTATION

Architecture

In this architecture I described the flow of the process of the proposed system. This architecture will describe the online social user process of the chat procedure and detection of bully users, in this flow there is a manual procedure and another one is system procedure in the detection process. Let’s describe the architecture components.

using a classification model. Based on this requirement, the proposed architecture is designed based on two flows, namely, classification analysis and user side prediction. Based on the proposed system requirement, this architecture is designed which is represented in figure 1. This section described the workflow of the architecture and main modules. The main purpose of the CBD system is to identify the bully or offensive statements using Machine Learning algorithms. To achieve this requirement, need to implement the classification analysis. The classification analysis is the process of conducting training and testing process for calculating the performance measures between various Machine Learning algorithms. Based on these requirements, the proposed architecture is designed with two flows, namely, classification analysis which is taken care of by the admin and user side prediction. In the architecture (Figure 1), these two flows are represented in two different color formats.

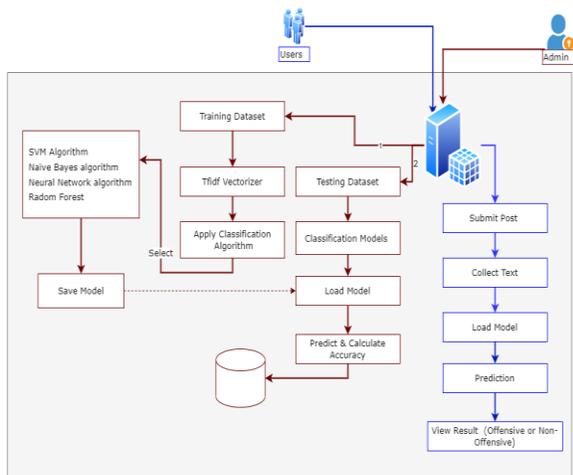


Fig:-1 CBD Architecture

The proposed system ‘Cyber Bullying Detection (CBD) in Social Networking’ is to identify the bully or offensive statements

Admin:

Admin is a main user of our application, admin will process main functionalities of the system. Admin can see the lists of users are available in the system. Admin can handle the classification analysis process of

the Harassment-Corpus dataset [11]. For classification analysis, taken three algorithms, namely, Naïve Bayees, Support Vector Machine and Neural Network algorithms.

User:

In our system user is a end user of our application. We build a social application for users. They can make friends and share data among them. User also chat with his/her friends. When user send any bullying words to the others, the application detects the bully user and send the warning notification to the user. User will get the notification of warning through a message in chat window.

4. EXPERIMENTAL RESULTS

Our Application Home Page

In this result screenshot we can see the home page of our application. This is designed in HTML and CSS technologies. From home page only user and admin can login with their login credentials.



Fig:-2 Application Home Screen

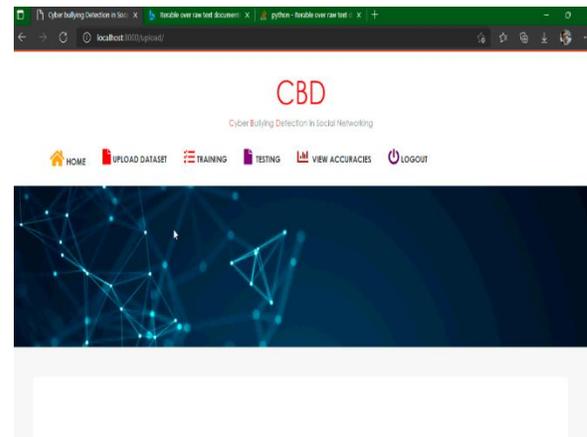


Fig:-3 Admin Dashboard

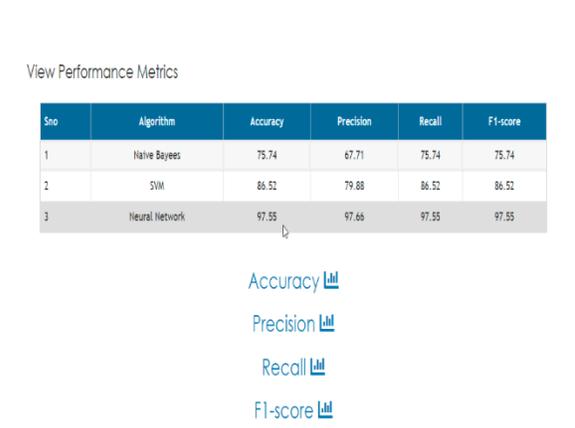


Fig:-4 Performance Metrics

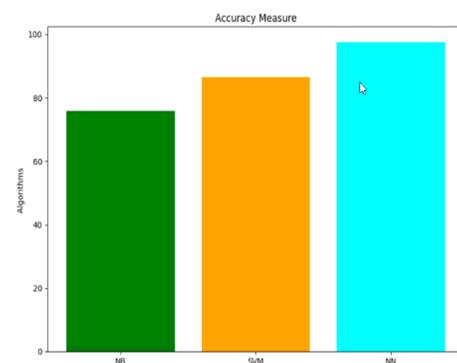


Fig:-5 Accuracy Measure Graph

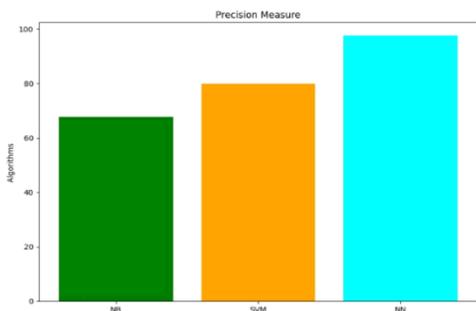


Fig:-6 Precision Measure Graph

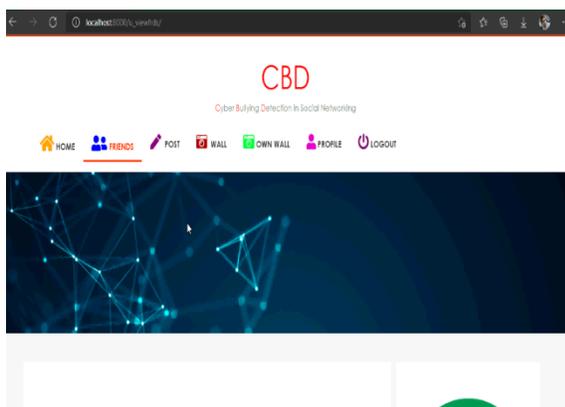


Fig:-7 User Dashboard

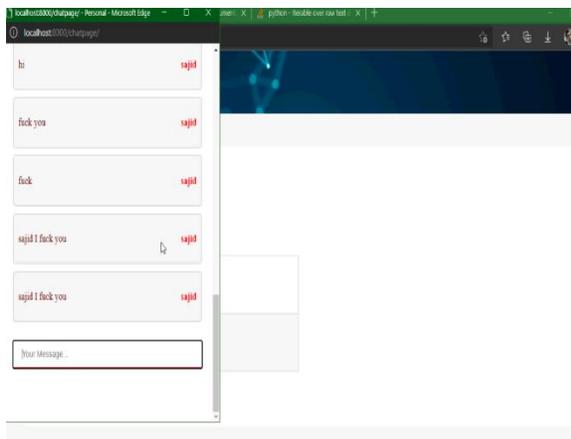


Fig:-8 User Chat Screen

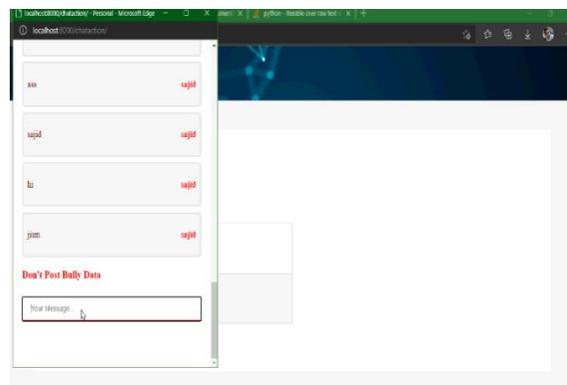


Fig:-9Bullying warning on User Chat Screen

5. CONCLUSION

In current trend many social networking sites created and providing services of communications, multi-media services, e-commerce etc immensely. We need to focus for the tracking the bully users. In this paper, we proposed an architecture with Machine Learning analysis for detecting the bully data published in social networks. The architecture is divided in two ways. One flow is classification analysis with Machine Learning algorithms, namely, Naïve Bayes, Support Vector Machine and Neural Network algorithms. In the classification analysis, performed comparison of performance with four matrices, called, Accuracy, Precision, Recall, and F1-scores. In these results, we achieved 97% accuracy for neural network algorithm. In future work, suggesting continue this work for media data, such as images, videos, etc.

6. REFERENCES

- [1] T. Iofciu, P. Fankhauser, F. Abel, and K. Bischoff, "Identifying users across social tagging systems," in Proc. 5th Int. AAAI Conf. Weblogs Social Media, 2011, pp. 522-525.
- [2] Identifying Users Across Social Tagging Systems TerezaIofciu, PéterFankhauser, +1 author Kerstin Bischoff, prieto et al. 2009
- [3] D. Perito, C. Castelluccia, M. A. Kaafar, and P. Manils, "How unique and traceable are usernames?" in Proc. 11th Int. Conf. Privacy Enhancing Technol., 2011, pp. 1-17.
- [4] J. Liu, F. Zhang, X. Song, Y. I. Song, C. Y. Lin, and H. W. Hon, "What's in a name?: An unsupervised approach to link users across communities," in Proc. 6th ACM Int. Conf. Web Search Data Mining, 2013, pp. 495-504.
- [5] R. Zafarani and H. Liu, "Connecting corresponding identities across communities," in Proc. 3rd Int. ICWSM Conf., 2009, pp. 354-357.
- [6] R. Zafarani and H. Liu, "Connecting users across social media sites: a behavioral-modeling approach," in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 41-49.
- [7] A. Acquisti, R. Gross, and F. Stutzman, "Privacy in the age of augmented reality," in Proc. Nat. Acad. Sci., 2011, pp. 36-53, Available: <https://www.usenix.org/legacy/events/sec11/tech/slides/acquisti.pdf>
- [8] T. Iofciu, P. Fankhauser, F. Abel, and K. Bischoff, "Identifying users across social tagging systems," in Proc. 5th Int. AAAI Conf. Weblogs Social Media, 2011, pp. 522-525.
- [9] M. Motoyama and G. Varghese, "I seek you: searching and matching individuals in social networks," in Proc. 11th Int. Workshop Web Inf. Data Manage., 2009, pp. 67-75.
- [10] O. Goga, D. Perito, H. Lei, R. Teixeira, and R. Sommer, "Large-scale correlation of accounts across social networks," University of California at Berkeley, Berkeley, California, Tech. Rep. TR-13-002, 2013.
- [11] Mrezvan, GitHub, 2018, [online], Available At: <https://github.com/Mrezvan94/Harassment-Corpus>, last accessed: Dec, 30, 2021.