

Detection of Cyberbullies through Sentiment Analysis Using Svm and Maxent

Ms.T.Sumathi¹, V.Praveen², S.Arun Kumar³

¹Department of Information Technology, Dr Mahalingam College of Engineering and Technology, Pollachi, Coimbatore 642003, India;

²Department of Information Technology, Dr Mahalingam College of Engineering and Technology, Pollachi, Coimbatore 642003, India

³Department of Information Technology, Dr Mahalingam College of Engineering and Technology, Pollachi, Coimbatore 642003,India;

Submitted: 10-07-2022

Revised: 18-07-2022

Accepted: 23-07-2022

ABSTRACT

Cyberbullying has evolved on social media in a variety of ways, both hidden and visible with the use of memes and open as the cyber bullying remarks using fake identities to become violence. Color, origin, language, culture, and, most crucially, religion are now used as grounds for cyberbullying. As a result, because social media is the most common source of cyberbullying opinions, it should be monitored and comments should be identified and blocked. As social media platforms have become more widely used around the world, particularly among young people, cyber bullying and aggressiveness has become a serious and disturbing problem that communities must deal with in recent years. Machine learning (ML) techniques are commonly employed to recognise particular language patterns that bullies use to assault their victims. Also, Sentiment Analysis (SA) of social media information is one of the fastest-growing fields of machine learning research. This study shows how to use a SA model to detect cyberbullying texts on social media sites like Facebook. The supervised machine learning classification tools SVM and MAXENT are used in this model and it gives the better performance.

Keywords: Machine learning, Support Vector Machine, Maxent classifier.

I. INTRODUCTION

Social media is used as a key form of communication by almost everyone, especially young people. It has become more common as the digital environment has grown and technology has advanced, especially among youth. When someone, usually a teenager, bullies or harasses people on the internet and other digital channels, notably social

media platforms, it is known as cyberbullying. As a result, the number of people who use the internet has risen. Cyberbullies who have been physically, emotionally, mentally, or physically harmed.

1.1 RELATED WORKS

NLP (natural language processing) is an emerging field with the help of NLP and machine learning algorithms such as naive bayes, random forest, SVM we are going to identify cyberbullying in facebook. The initiative's main purpose is to create a system that can automatically recognise symptoms of cyberbullying on social media, such as distinct types of cyberbullying and posts by bullies, victims, and bystanders. [3]

A collection of huge datasets was made available for a misbehaviour detection project at the Content Analysis on the Web 2.0 (CAW 2.0) workshop, but this dataset was unlabeled (i.e. it is not known which posts actually contain cyberbullying). Furthermore, the data came from a variety of sources, and all of the datasets, except the Kongregate data, looked to be adult talks (a gaming website). To conduct the research detailed here, we constructed our own labelled dataset using data from a web crawl of Formspring.me. Formspring.me was chosen because it is primarily populated by teenagers and college students, and the data contains a large volume of bullying content. [4]

This study evaluates the performance of three classifiers in a machine-learning sentiment analysis model. Data collection, pre-processing, feature vector extraction, and classification are the four stages of the model. The performance of the classifiers is evaluated using three datasets, with the NB classifier outperforming the others. However, given all of the classifiers perform well,

it can be inferred that the machine-learning approach for sentiment categorization has a lot of potential. Now, depending on the existence or absence of features in a feature vector of the preceding patterns, 1 or 0 is assigned to each attribute for each pattern of the dataset to produce the final dataset to be utilised in the classification. [13]

This research proposes a supervised machine learning technique for recognising and combating cyberbullying. A variety of classifiers are used to learn and recognise bullying behaviours. The recommended approach outperforms SVM on the cyberbullying dataset, with an accuracy of 92.8 percent versus 90.3 percent for SVM. NN also outperforms other similar classifiers on the same dataset. With the growing number of social media users comes a new form of bullying. To increase the performance, more cyberbullying data is required. [2]

II. METHODS:

2.1 SUPPORT VECTOR MACHINE

SVM is a technique for categorising and forecasting data that uses supervised machine learning. Despite the difficulties of regression, classification is the most appropriate method. In an N-dimensional space, the SVM algorithm looks for a hyperplane that clearly categorises data points. To categorise the data, a variety of hyper planes might be used. The hyper plane that reflects the greatest separation, or margin, between the two classes is a reasonable choice. As a result, the hyper plane is chosen so that the distance between it and the closest data point on each side is maximised. SVMs can significantly minimise the demand for labelled training instances in both inductive and transductive contexts, making them effective in text and hypertext categorization. Several shallow semantic parsing techniques use support vector machines.

2.2 MAXENT CLASSIFIER

By parameterizing the model to attain maximum categorical entropy, the maximum entropy (maxent) classifier has been a popular text classifier. Maximum Entropy Start with the least informative weights (priors) and work your way up until you find weights that maximise the likelihood of the data. It's the EM algorithm in essence. In a simple Naive Bayes classifier, the prior weights are proportional to the number of times a phrase appears in the document. However, word relationships are ignored. The earlier loads would be corresponding to the times the word appears in the archive, according to a simple Naive Bayes

classifier. Be that as it may, this disregard relationships between words. The Max Entropy classifier belongs to the class of exponential models and is a classification algorithm. In contrast to the Naive Bayes classifier.

III. PROPOSED METHODOLOGY:

We use a support vector machine and a maxent classifier in the suggested system. SVM is a technique for categorising and forecasting data that uses supervised machine learning. Despite the difficulties of regression, classification is the most appropriate method. In an N-dimensional space, the SVM algorithm looks for a hyperplane that clearly categorises data points. By parameterizing the model to attain maximum categorical entropy, the maximum entropy (maxent) classifier has been a popular text classifier. The Max Entropy classifier belongs to the exponential models category of classification algorithms. Unlike the Naive Bayes classification algorithm. The conditional independence of the features is ignored by the Max Entropy.

ADVANTAGES OF PROPOSED

The algorithm presents the best result in identifying the cyber bully comments. Large number of data can be implemented in dataset. Better accuracy. Two phase risk assessment is implemented for better result.

3.1. MODULES

3.1.1. LOAD DATASET

The input data set consists of attributes like post ID, user ID, username, gender, age, City, province, country, user, date, posts. As many number of comments are stored in the form of dataset which will be used for the further processing of two phase risk assessment phase.

3.1.2. GROUP IDENTIFICATION PHASE

To achieve such identifications, they commit more resources to the in-group in order to maximise the difference between their in-group and out-groups. The first step is to categorise everything. To better comprehend and recognise objects, we classify them. We categorise people in the same way in order to understand the social context. The second stage, known as social identification, is adopting the identity of the group to which we have been assigned. In this module the loaded dataset is in raw data and it converts into train data. In this module the comments should be categories into four parts Male, Female, less than 18 years, greater than 18 years .

3.1.3. RISK ASSESSMENT PHASE

In this module by using the cyberbullies words it finds the risky and the normal behaviour in dataset. It checks the words with the help of group

identification phase because of categorization. At last, if it finds the words mention in the cyberbullies it marked as risky or otherwise it marked as Normal behaviour.

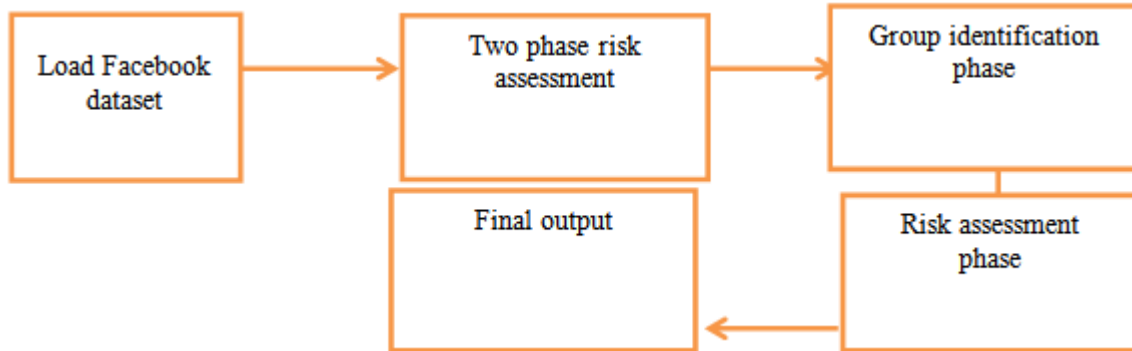


Fig 1: Flow chart

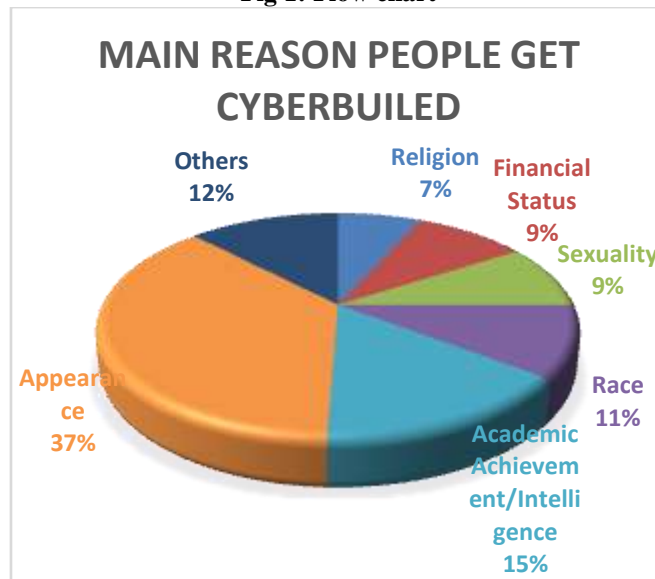
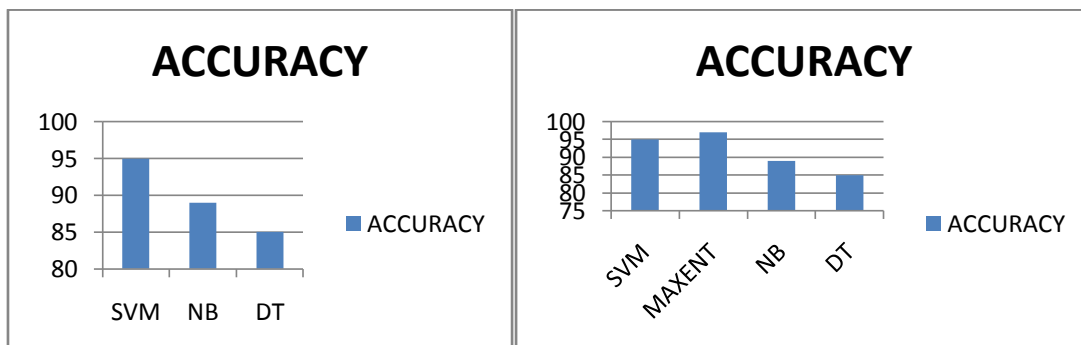


Fig 2: Main Reason People Get Cyberbullies

IV. EXPERIMENTS AND RESULTS:

We gathered a large number of comments to assess the performance of the machine learning algorithms employed in this study, specifically the Support Vector Machine (SVM) and maxent

(Positive-cyber bullying, negative-no cyber bullying). There were no cyber bullying comments in this group of comments, thus they were manually categorised.



V. CONCLUSION:

SVM classifiers and Maxent classifiers outperformed in almost all performance measures across all language models, according to the findings of the studies. In comparison to past studies, these experiments revealed that our SVM and maxent classifiers had slightly improved performance measures. Finally, we'd like to investigate alternative machine learning techniques, such as Neural Networks and deep learning, using larger sets of comments for direction study in cyberbullying detection. To handle such a large number of comments, it's also necessary to use some tried and true ways for an automated annotation process.

REFERENCES:

- [1]. E. Lee, F. Rustam, P. B. Washington, F. E. Barakaz, W. Aljedaani and I. Ashraf, "Racism Detection by Analyzing Differential Opinions Through Sentiment Analysis of Tweets Using Stacked Ensemble GCR-NN Model," in *IEEE Access*, vol. 10, pp. 9717-9728, 2022, doi: 10.1109/ACCESS.2022.3144266.
- [2]. Internet Monitoring and Web Filtering Solutions", "PEARL SOFTWARE, 2015. Online. Available: <http://www.pearlsoftware.com/solutions/cyber-bullyinginschools.html>. [Accessed Feb 20, 2020]
- [3]. H. Kumar Sharma, K. Kshitiz and Shailendra, "NLP and Machine Learning Techniques for Detecting Insulting Comments on Social Networking Platforms," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2018, pp. 265-272, doi: 10.1109/ICACCE.2018.8441728.
- [4]. Van Hee, Cynthia et al. "Automatic detection of cyberbullying in social media text." *PloS one* vol. 13,10 e0203794. 8 Oct. 2018, doi:10.1371/journal.pone.0203794
- [5]. Dipika Jiandani, Riddhi Karkera, Megha Manglani, Mohit Ahuja, Mrs. Abha Tewari, "Comparative Analysis of Different Machine Learning Algorithms to Detect Cyber-bullying on Facebook", *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, Volume 6 Issue IV, April 2018, pp. 2322-2328.
- [6]. S. Salawau, Y. He, and J. Lumsden, "Approaches to Automated Detection of Cyberbullying: A survey," Vol. 3045, no c, pp 1-20, 2017.
- [7]. Walisa Romsaiyud, Kodchakornna Nakornphanom, Pimpaka Prasertsrip, Piyapon Nurarak, and Piromkonglerd, "Automated Cyberbullying Detection Using Clustering Appearance Patterns", in *Knowledge and Smart Technology (KST), 2017 9th International Conference on*, pages 242-247, IEEE, 2017.
- [8]. B. Sri Nandhinia, and J.I. Sheeba, "Online Social Network Bullying Detection Using Intelligence Techniques", *International Conference on Advanced Computing Technologies and Applications (ICACTA2015)*, *Procedia Computer Science* 45 (2015) 485 – 492
- [9]. Cristina Bosco and Viviana Patti and Andrea Bolioli, "Developing Corpora for Sentiment Analysis: The Case of Irony and Senti-TUT, *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pp. 4158-4162
- [10]. Tavani, Herman. T., "Introduction to Cybernetics: Concepts, Perspectives, and Methodological Frameworks", In H. T. Tavani, *ethics and Technology: Controversaries, questions, and Strategies for ethical Computing*, river University – Fourth Edition, Wiley, pp 1-2, 2013.
- [11]. V. Nahar, X. Li and C. Pang, "An effective Approach for Cyberbullying Detection," in *Communication in Information Science and Management Engineering*, May 2013.
- [12]. Chen, Y., Zhou, Y., Zhu, s. and Xu, H., "Detecting Offensive Language in Social Media to Protect Adolescent Online Safety", In *privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on Social Computing (SocialCom)*, pp 71-80, 2012.
- [13]. K. Reynolds, "Using Machine Learning to Detect Cyberbullying", 2012.