

Detecting Twitter Cyberbullying Using Machine Learning with Big Data

Akilash SK¹, Arunachalam P¹, Dharanees Kumar S¹, Keshoth U¹, Dr.N. Suguna², Mrs. K. Veena³

¹UG Scholar, Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu, India

²Professor, Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu, India

³Assistant Professor, Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu, India

ABSTRACT

Article Info

Volume 8, Issue 3

Page Number : 282-288

Publication Issue :

May-June-2022

Article History

Accepted: 01 June 2022

Published: 07 June 2022

Online media is a stage where numerous youthful individuals are getting tormented. As person to person communication destinations are expanding, cyberbullying is expanding step by step. To recognize word likenesses in the tweets made by menaces and utilize AI and can build up a ML model naturally recognize online media tormenting activities. In any case, numerous online media tormenting identification methods have been actualized, however numerous of them were printed based. The objective of this paper is to show the execution of programming that will distinguish tormented tweets, posts, and so on An AI model is proposed to distinguish and forestall tormenting on Twitter. Two classifiers for example SVM and RF are utilized for preparing and testing the online media tormenting content. Both SVM (Support Vector Machine) and RF had the option to recognize the genuine positives with 71.25% and 52.70% precision individually. Yet, SVM beats RF of comparable work on the equivalent dataset.

Keywords: Machine Learning, Big Data, Cyberbullying, Detection, NLP, Learning

I. INTRODUCTION

Nowadays technology has become a very important part of our lives and most people can't live without it. The Internet provides a platform to share their ideas.

Many people are spending a large amount of time on social media. Communicating with people is no exception, as technology has changed the way people interact with a broader manner and has given a new

dimension to communication. Many people are illegally using these communities.

Many youngsters are getting bullied these days. Bullies use various services like Twitter, Facebook and Email to bully people. Studies show that about 37% of children in India are involved in cyber bullying and nearly 14% of bullying occurs regularly. Social media also allows bullies to harness the anonymity which satisfies their unkind deeds. Things also get more serious when bullying occurs more repeatedly over time. So, preventing it from happening will help the victim. Thus, a complete solution is required for this problem. Cyber bullying needs to stop. The problem can be tackled by detecting and preventing it by using a machine learning approach, this needs to be done using a different perspective.

To better understand the uses of machine learning, consider some of the instances where machine learning is applied: the self-driving Google car, cyber fraud detection, online recommendation engines—like friend suggestions on Facebook, Netflix showcasing the movies and shows you might like, and “more items to consider” and “get yourself a little something” on Amazon—are all examples of applied machine learning.

II. LITERATURE REVIEW

To better understand the uses of machine learning, consider some of the instances where machine learning is applied: the self-driving Google car, cyber fraud detection, online recommendation engines—like friend suggestions on Facebook, Netflix showcasing the movies and shows you might like, and “more items to consider” and “get yourself a little something” on Amazon—are all examples of applied machine learning. All these examples echo the vital role machine learning has begun to take in today’s data-rich world. Machines can aid in filtering useful pieces of information that help in major advancements, and we are already seeing how this

technology is being implemented in a wide variety of industries.

2.1. Deep learning approach to text analysis for human emotion detection from big data

Emotional recognition has arisen as an essential field of study that can expose a variety of valuable inputs. Emotion can be articulated in several means that can be seen, like speech and facial expressions, written text, and gestures. Emotion recognition in a text document is fundamentally a content-based classification issue, including notions from natural language processing (NLP) and deep learning fields. Hence, in this study, deep learning assisted semantic text analysis (DLSTA) has been proposed for human emotion detection using big data. Emotion detection from textual sources can be done utilizing notions of Natural Language Processing. Word embeddings are extensively utilized for several NLP tasks, like machine translation, sentiment analysis, and question answering. NLP techniques improve the performance of learning-based methods by incorporating the semantic and syntactic features of the text.

2.2. AI Based Emotion Detection for Textual Big Data: Techniques and Contribution

Online Social Media (OSM) like Facebook and Twitter has emerged as a powerful tool to express via text people’s opinions and feelings about the current surrounding events. Understanding the emotions at the fine-grained level of these expressed thoughts is important for system improvement. Such crucial insights cannot be completely obtained by doing AI-based big data sentiment analysis; hence, text-based emotion detection using AI in social media big data has become an upcoming area of Natural Language Processing research. It can be used in various fields such as understanding expressed emotions, human-computer interaction, data mining, online education, recommendation systems, and psychology

III. OBJECTIVE AND PROBLEM STATEMENT

Emotion can be expressed in many ways that can be seen such as facial expression and gestures, speech and by written text. Emotion Detection in text documents is essentially a content - based classification problem involving concepts from the domains of Natural Language Processing as well as Machine Learning.

In machine learning, the detection of textual emotions is the problem of content-based classification, which is the task of natural language processing. Detecting a person's emotions is a difficult task, but detecting the emotions using text written by a person is even more difficult as a human can express his emotions in any form.

IV. EXISTING SYSTEM

Cyberbullying causes emotional and physiological damage to defenseless victims (Faryadi, 2011) as well as psychosocial problems including inappropriate behaviors, drinking alcohol, smoking, depression and low commitment to academics (Walker et al., 2011). Twitter is listed as one of the top five social media platforms where the maximum percentage of users experience cyberbullying (turbofuture.com, 2019). It enables a user to send a message of 280-characters, with more than 330 million active users at present (Statista, 2018).

Studies on cyberbullying and Twitter often reported extensive cases of the phenomenon, with the potential for serious, deleterious consequences for its victims (Chatzakou et al., 2017a; Balakrishnan et al., 2019; Sterner, 2017). Several measures have been taken by Twitter to mitigate cyberbullying, such as filtering unwanted messages from users without a profile picture, and enabling a time-out feature that bans users using abusive language, among others. Despite these positive attempts, the platform is not completely immune from cyberbullying (Bernazzani, 2017; Twitter, 2019).

4.1 Disadvantages

1. Accuracy will be low. The time complexity is very high because we are working on text data.
2. Covert the text data into numeric form is very big task. We have to text preprocessing like removing the stop words, punctuation marks...etc.

V. PROPOSED SYSTEM

In this project, a solution is proposed to detect twitter cyberbullying. The main difference with previous research is that we not only developed a machine learning model to detect cyberbullying content but also implemented it on particular locations real-time tweets using Twitter API. The entire approach to detect and prevent Twitter cyberbullying is divided into 2 major stages: developing the model and experimental setup.

Stepwise Procedure of SVM and Naïve Bayes utilized in detecting the cyberbullying Steps:

1. For a particular location, a limited number of tweets will be fetched through Twitter's tweet API.
2. The Data Preprocessing, Data Ext reaction will be performed on the fetched Tweets. Preprocessed tweets will be passed to SVM and Naïve Bayes model (see Developing the Model section) to calculate the probabilities of fetched tweets to check whether a fetched tweet is bullying or not.

Dataset preparation and preprocessing

The dataset has 20001 items of which 20001 items have been manually labeled. The labels are divided into following 2 categories:

- 1 (Cyber-bullying) 0 (Non Cyber-bullying)

Pre-processing refers to the transformations applied to our data before providing the data to the algorithm. Data Preprocessing technique is used to convert the raw data into an understandable data set. In other words, whenever the information is gathered from

various sources it is collected in raw format that isn't possible for the analysis.

Pre-processing and Cleaning

News-Punctuation Cleaning

Punctuation is the tool that allows us to organize our thoughts and make it easier to review and share our ideas. The standard English punctuation is as follows: period, comma, apostrophe, quotation, question, exclamation, brackets, braces, parenthesis, dash, hyphen, ellipsis, colon, semicolon.

Stop words

A stop word is a commonly used word (such as "the", "a", "an", "in") that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query. We would not want these words to take up space in our database, or taking up the valuable processing time. For this, we can remove them easily, by storing a list of words that you consider to stop words

Featurization

In this step, the proposed model has transformed the data in a suitable form which is passed to the machine learning algorithms. The TFDIF vectorizer is used to extract the features of the given data. Features of the data are extracted and put them in a list of features. Also, the polarity (i.e. the text is Bullying or Non-Bullying) of each text is extracted and stored in the list of features.

Data Split

- For choosing a model we split our dataset into train and test
- Here data's are split into 3:1 ratio that means
- Training data having 70 percent and testing data having 30 percent
- In this split process performing based on train_test_split model
- After splitting we get xtrainxtest and ytrainytest

Model Creation

- Contextualise machine learning in your organisation.

- Explore the data and choose the type of algorithm.
- Prepare and clean the dataset.
- Split the prepared dataset and perform cross validation.
- Perform machine learning optimisation.
- Deploy the model.

Model Evaluation

Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance, as well as its strengths and weaknesses. Model evaluation is important to assess the efficacy of a model during initial research phases, and it also plays a role in model monitoring.

VI. CONCLUSION

The goal of this project is to the automatic detection of cyberbullying-related posts on social media. Given the information overload on the web, manual monitoring for cyberbullying has become unfeasible. Automatic detection of signals of cyberbullying would enhance moderation and allow to respond quickly when necessary. However, these posts could just as well indicate that cyberbullying is going on. The main aim of this project is that it presents a system to automatically detect signals of cyberbullying on social media, including different types of cyberbullying, covering posts from bullies, victims and bystanders.

VII. CONCLUSION

With the rising number of COVID cases from one side of the planet to the other, a framework to supplant people to check veils on the essences of individuals is extraordinarily required. This application fulfills that need. This application can be utilized openly puts at passage level in rail route stations, shopping centers, emergency clinics, and so forth. It will be of an extraordinary assistance in organizations and immense foundations where there

will be a ton of laborers. This application will be an incredible assistance there since it is not difficult to recognize faces and shows that he/she has the cover. So this is exceptionally helpful to forestall the Coronavirus infection by making mindfulness by putting cover.

VIII. REFERENCES

- [1]. John Hani Mounir, Mohamed Nashaat, Mostafa Ahmed, Zeyad Emad, Eslam Amer, Ammar Mohammed, "Social Media Cyberbullying Detection using Machine Learning", (IJACSA) International Journal of Advanced Computer Science and Applications Vol. 10, pages 703-707, 2019.
- [2]. Kelly Reynolds, April Kontostathis, Lynne Edwards, "Using Machine Learning to Detect Cyberbullying", 2011 10th International Conference on Machine Learning and Applications volume 2, pages 241-244. IEEE, 2011
- [3]. Amanpreet Singh, Maninder Kaur, "Content-based Cybercrime Detection: A Concise Review", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-8, pages 1193-1207, 2019
- [4]. Abdhullah-Al-Mamun, Shahin Akhter, "Social media bullying detection using machine learning on Bangla text", 10th International Conference on Electrical and Computer Engineering, pages 385- 388, IEEE Xplore, 2018.
- [5]. NektariaPotha and ManolisMaragoudakis. "Cyberbullying detection using time series modeling", In 2014 IEEE International Conference on, pages 373-382. IEEE, 2014.
- [6]. Ying Chen, Yilu Zhou, Sencun Zhu, and Heng Xu. "Detecting offensive language in social media to protect adolescent online safety". In Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012

International Conference on Social Computing (SocialCom), pages 71- 80. IEEE, 2012

Cite this Article

Akilash SK, Arunachalam P, Dharanees Kumar S, Keshoth U, Dr.N. Suguna, Mrs. K. Veena, "Detecting Twitter Cyberbullying Using Machine Learning with Big Data", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 3, pp. 294-298, May-June 2022.

Journal URL : <https://ijsrcseit.com/CSEIT1228379>