WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Detecting Twitter Cyberbullying Using Machine Learning

DR. M. ARUN[1], RAGAVI.S [2]

[1] Research Scholar, Assistant Professor Department of Computer Science
[2] Department of Computer Science Sri Krishna Adithya College of Arts and Science
Kovaipudur , Coimbatore, Tamil Nadu, India
sragavi699@gmail.com

ABSTRACT—

Cyberbullying on social media platforms, particularly on Twitter, has become a pressing societal concern, with detrimental effects on individuals' mental health and well-being. Traditional methods of identifying and mitigating cyberbullying are often insufficient due to the sheer volume and complexity of online interactions. In response, this study proposes a novel approach leveraging machine learning techniques for the automated detection of cyberbullying on Twitter.

Our methodology involves several key steps. Firstly, we gather a large dataset of Twitter conversations annotated with cyberbullying labels. Next, we preprocess the text data, including tokenization, stemming, and removing stopwords, to prepare it for analysis. Subsequently, we extract a diverse set of features, including lexical, syntactic, and semantic features, to capture the nuances of cyberbullying behavior.

Keywords:Naive bayes algorithm,(Confusion metrics),Machine Learning,multinational naive bayes algorithm

Introduction:

Cyberbullying, the act of harassing, intimidating, or threatening others through digital platforms, has emerged as a significant societal issue, particularly on social media platforms like Twitter. With the increasing prevalence of online interactions, cyberbullying poses serious threats to individuals' mental health, well-being, and even their safety. Traditional methods of detecting and addressing cyberbullying are often ineffective due to the sheer volume of online content and the subtlety of abusive behavior.
In response to this challenge, researchers and technologists have turned to machine learning as a potential solution. Machine learning algorithms have shown promise in automating the detection of cyberbullying by analyzing large volumes of text data and identifying patterns indicative of abusive behavior. Leveraging the wealth of information shared on platforms like Twitter, machine learning techniques offer a scalable and efficient approach to identifying and addressing cyberbullying in real-time.

EXISTING SYSTEM:

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).

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. Time consuming and prediction is not perfect

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:
For a particular location, a limited number of tweets will be fetched through Twitter‟s tweet API
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.

SYSTEM REQUIREMENT HARDWARE REQUIREMENT

- .Ram:4GB or 8GB
- Hard Disk 64GB
- Proceeser:i3 or i5

SOFTWARE REQUIREMENT

- Python 3.9.6
- Excel(.CSV)
- Flask
- Back end :Python
- front end :HTML CSS