INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.521

Detection Of Cyber Bullying on Social Media Using Machine Learning

R. Arthi, K. Kamalakkannan S

Assistant Professor, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu, India

UG Student, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu, India

ABSTRACT: In the modern era, the usage of the internet has increased tremendously which in turn has led to the evolution of large amounts of data. Cyber world has its own pros and cons. One of the alarming situations in web 4.0 is cyber bullying, a type of cyber-crime. When bullying occurs online with the aid of technology it is known as cyber bullying. This research paper has surveyed the work done by 30 different researchers on cyber bullying, and elaborated on different methodologies adopted by them for the detection of bullying. Three types of features namely textual, behavioural and demographic features are extracted from the dataset as compared to earlier study over the same dataset where only textual features were considered. Textual features include certain bullying words that if exists within the text may lead to a true outcome for cyber bullying. Personality trait features are extracted for the user if it is involved once in bullying may bully in future too. While demographic features extracted from the dataset include age, gender and location.

I. INTRODUCTION

People from all around the world are using social media in their day-to-day life. Many peoples are facing some issues like cyber bullying. Cyberbullying is repeated behaviour, aimed at scaring, angering, or shaming those who are targeted. It can take place on social media, messaging platforms, gaming platforms, and mobile phones. It is the act of intimidating, threatening, or coercing people online through the use of social media, email, text messages, blog posts, or other digital or electronic methods. It usually involves the use of derogatory, aggressive, or threatening language. Our proposed work, enhances deep learning techniques to overcome the cyberbullying occurs in the social media.

II. LITERATURE REVIEW

A. T. Aind, A. Ramnaney and D. Sethia, "A Reinforcement Learning based Cyberbullying Detection Framework",2020
Cyberbullying can cause severe mental and emotional impacts on a person, especially on minors; hence, there is a requirement to have intelligent automated systems to detect questionable content present on social media platforms and remove it. In this paper, we introduce our novel algorithm Q-Bully which can automatically detect cyberbullying on various social media and online gaming platforms using Reinforcement Learning along with Natural Language Processing techniques. Previously the techniques used to detect cyberbullying have a good accuracy related to the text they have been trained on and do not incorporate new word patterns without complete retraining of model. In this paper, we incorporated the use of Reinforcement Learning and have conducted an experimental study in which we feed the messages and posts of bullies as well as victims to a Reinforcement Learning Agent for classification.[1]

O. C. Hang and H. M. Dahlan, "Cyberbullying Lexicon for Social Media",2019
cyberbullying lexicon can be used to prevent cyberbullying, since user can know which words represent the cyberbullying word in social media. However, there is less research been done for developing the cyberbullying lexicon. Many researches are focusing on cyberbullying detection rather than cyberbullying instances. In order to fill the gaps, this paper aims to propose a cyberbullying lexicon for social media. This paper is focusing on exclusion cyberbullying and proposing an exclusion cyberbullying lexicon by using an ontological approach. This approach consists of several phases, namely: understanding the concepts of exclusion cyberbullying, word list selection, keyword identification, classes and subclasses identification, and lastly cyberbullying ontology and lexicon development. The classes, subclasses and the documentation have been included in the development of exclusion cyberbullying ontology.

Classes are based on the semantic frame, subclasses are based on the lexical entry, while documentation of classes and subclasses are based on the definition of the semantic frame and lexical entry from Frame Net. [2]

J. Zhang, T. Otomo, L. Li and S. Nakajima, "Cyberbullying Detection on Twitter using Multiple Textual Features," 2019
Due to the spread of PCs and smartphones and the rise of user-generated content in social networking service, cyberbullying is also increasing and has become a serious risk that social media users may encounter. In this paper, we focus on the Japanese text on Twitter and construct an optimal model for automatic detection of cyberbullying by extracting multiple textual features and investigating their effects with multiple machine learning models. [3]

Rahul Ramesh Dalvi, Sudhanshu Baliram Chavan, Aparna Halbe, "Detecting A Twitter Cyberbullying Using Machine Learning", 2020
Social media is a platform where many young people are getting bullied. As social networking sites are increasing, cyberbullying is increasing day by day. To identify word similarities in the tweets made by bullies and make use of machine learning and can develop an ML model automatically detect social media bullying actions. However, many social media bullying detection techniques have been implemented, but many of them were textual based. The goal of this paper is to show the implementation of software that will detect bullied tweets, posts, etc. A machine learning model is proposed to detect and prevent bullying on Twitter. Two classifiers i.e. SVM and Naïve Bayes are used for training and testing the social media bullying content. Also, Twitter API is used to fetch tweets and tweets are passed to the model to detect whether the tweets are bullying or not.[4]

Robert Slonje, Peter K. Smith a, Ann Frisén, "The nature of cyberbullying, and strategies for prevention", 2012.
Cyberbullying has been identified as an important problem among youth in the last decade. This paper reviews some recent findings and discusses general concepts within the area. The review covers definitional issues such as repetition and power imbalance, types of cyberbullying, age and gender differences, overlap with traditional bullying and sequence of events, differences between cyberbullying and traditional bullying, motives for and impact of cyber victimization, coping strategies, and prevention/intervention possibilities. These issues will be illustrated by reference to recent and current literature, and also by in-depth interviews with nine Swedish students aged 13–15 years, who had some first-hand experience of one or more cyberbullying episodes. We conclude by discussing the evidence for different coping, intervention and prevention strategies.[5]

III. PROPOSED SYSTEM

The purpose of this research is to design and develop an effective technique to detect online abusive and bullying messages by using deep learning. In this system, a deep learning algorithm is used to detect cyber bullying activities from social network data

Advantages
- By achieving the best accuracy, our work is definitely going to improve cyber bullying detection to help people to use social media safely.
- Cyberbullying detection model based on deep learning that can detect whether a text relates to cyberbullying or not.
- This feature provides better accuracy and better performance than the existing one

IV. IMPLEMENTATION

Dataset:
The informational collection might incorporate information for at least one individual, with the number of columns compared to the number of individuals. Every one of our pieces of information is kept here as CSV documents. A comma isolated values (CSV) document is a text record that utilizes a comma to isolate values in figuring. A CSV record is a plain text document that contains even information (numbers and text). An information record is addressed by each line in the document. Each record has at least one field, which is isolated by commas. The name for this record design comes from the utilization of the comma as a field separator. Our information is coordinated in a success accounting sheet with values like date, open, high, low, last, low, all-out exchange, and turnover. The dataset downloaded from. It contains the tweet data and their bulling method. More than 7000 data are used.

Data Pre-processing

It is a significant stage in AI drives. Data gathering approaches are generally uncontrolled, coming about in out-of-range regards, missing properties, and different issues. Taking apart information that hasn't been completely examined for such imperfections can prompt incorrect ends. Thus, preceding running an assessment, the depiction and nature of data are basic. Data readiness is every now and again the most tedious component of an AI project, particularly on account of computational information. In the event that there is a ton of irrelevant and overabundant information present, or on the other hand on the off chance that there is a great deal of boisterous and deceitful data, data disclosure during the readiness stage will be more troublesome. The techniques of data readiness and detachment can consume a large chunk of the day to finish. Cleaning, case choice, normalization, change, incorporating extraction and determination, etc are instances of information planning. The last preparation set is the consequence of information preprocessing.

Artificial Neural Network [ANN]

The Artificial Neural Network (ANN) is a deep learning method that arose from the concept of the human brain Biological Neural Networks. The development of ANN was the result of an attempt to replicate the workings of the human brain. There are three layers in the network architecture: the input layer, the hidden layer (more than one), and the output layer. Here we consider age, gender, height and weight as dataset (input layer) and BMI values as hidden layer and the exercise, diet and food recommendation as output layer.

V. CONCLUSION

Although social media platform has become an important entity for everyone, cyberbullying has many negative impacts on victim's life which include depression, anxiety, anger, fear, trust issues, low self-esteem. Therefore, detection of cyberbullying in the vast social media network has become immensely important. In summary, we have studied cyberbullying detection for messages and chats to identify cyberbullying text and actors on discord Platform. The work has successfully identified cyberbullying messages using the proposed artificial neural network.
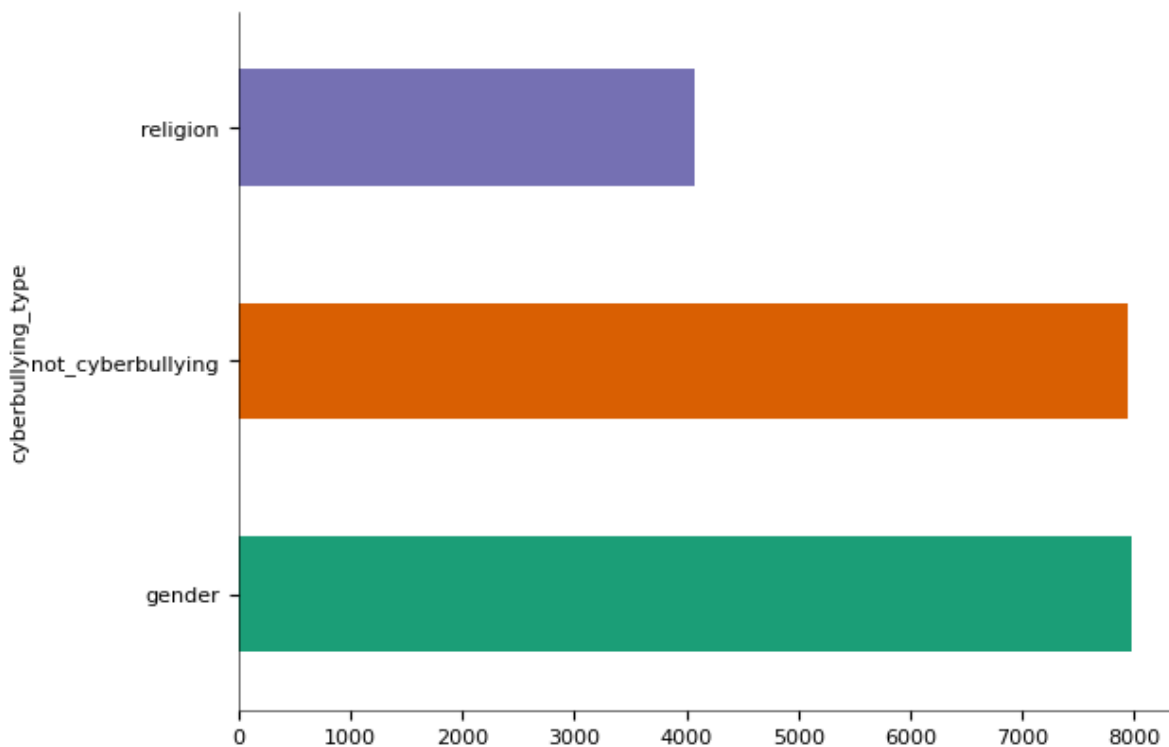
REFERENCES

1. A. T. Aind, A. Ramnaney and D. Sethia, "A Reinforcement Learning based Cyberbullying Detection Framework",2020
2. O. C. Hang and H. M. Dahlan, "Cyberbullying Lexicon for Social Media",2019
3. J. Zhang, T. Otomo, L. Li and S. Nakajima, "Cyberbullying Detection on Twitter using Multiple Textual Features," 2019
4. Rahul Ramesh Dalvi, Sudhanshu Baliram Chavan, Aparna Halbe, "Detecting A Twitter Cyberbullying Using Machine Learning", 2020
5. Robert Slonje a, Peter K. Smith a, Ann Frisén, "The nature of cyberbullying, and strategies for prevention", 2012.
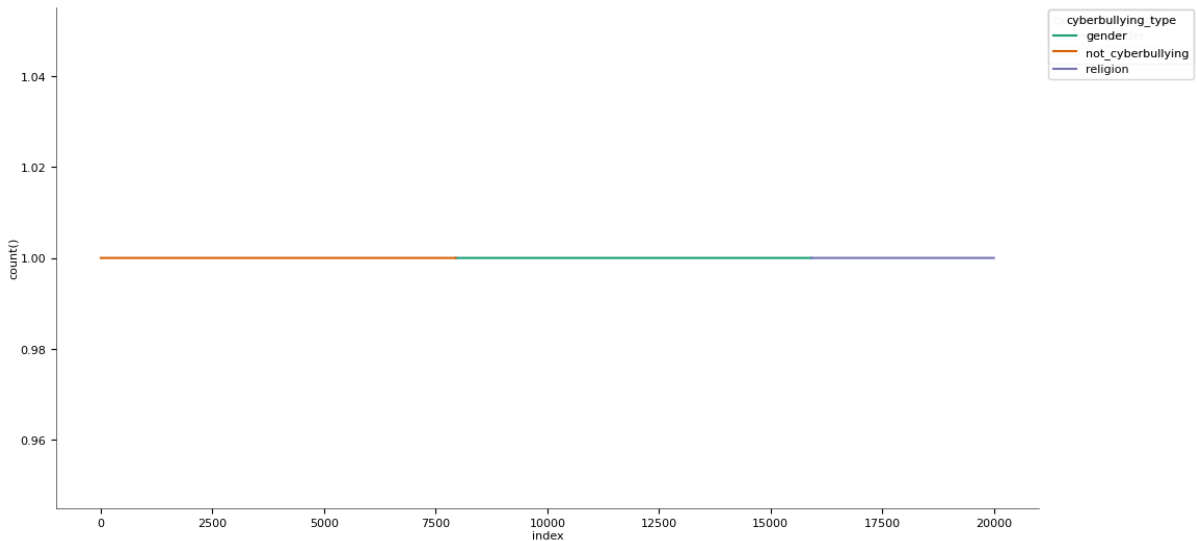
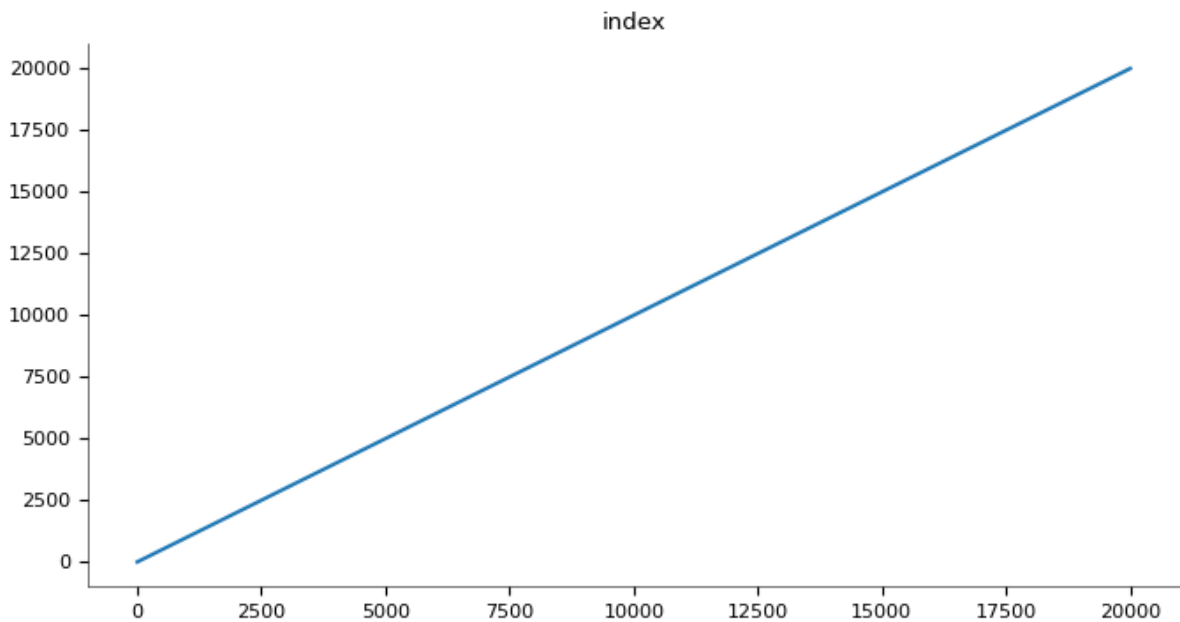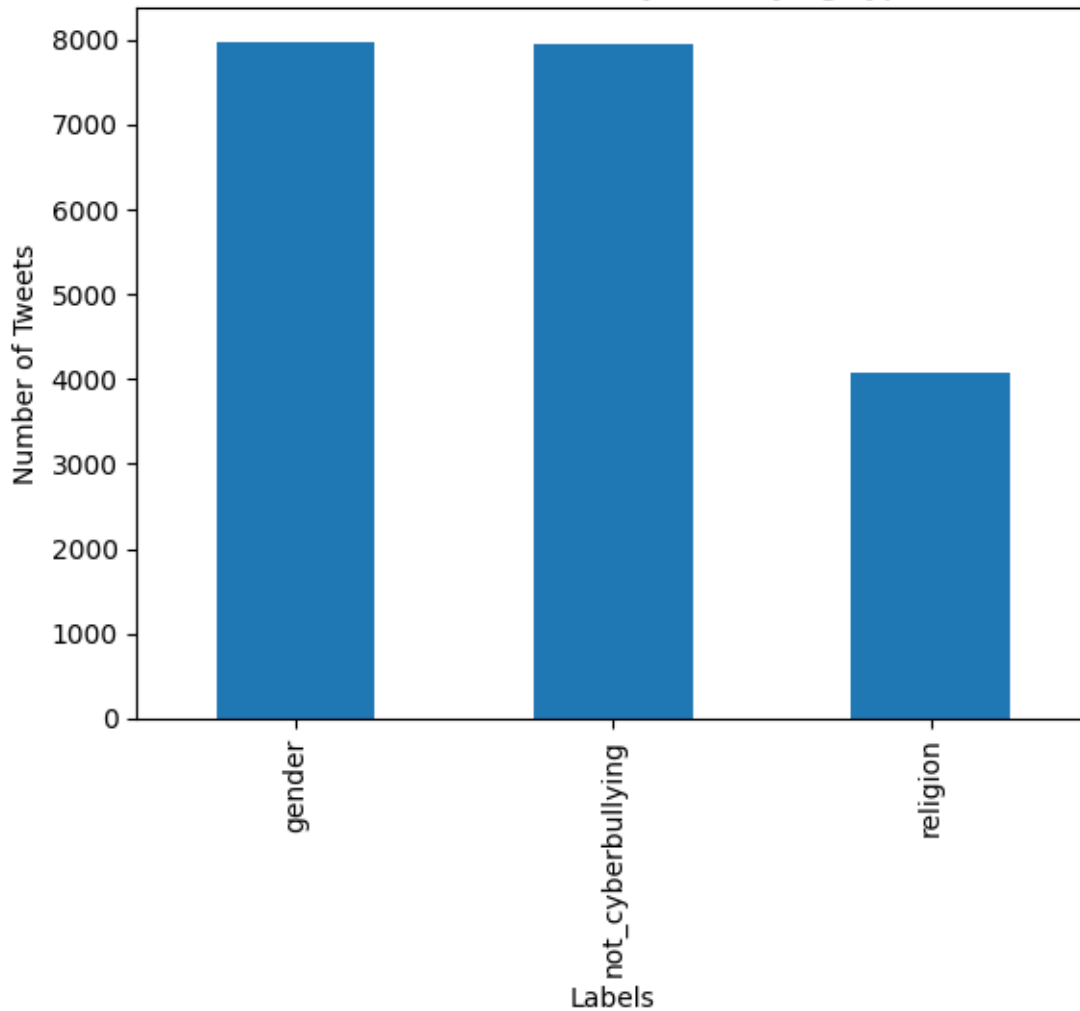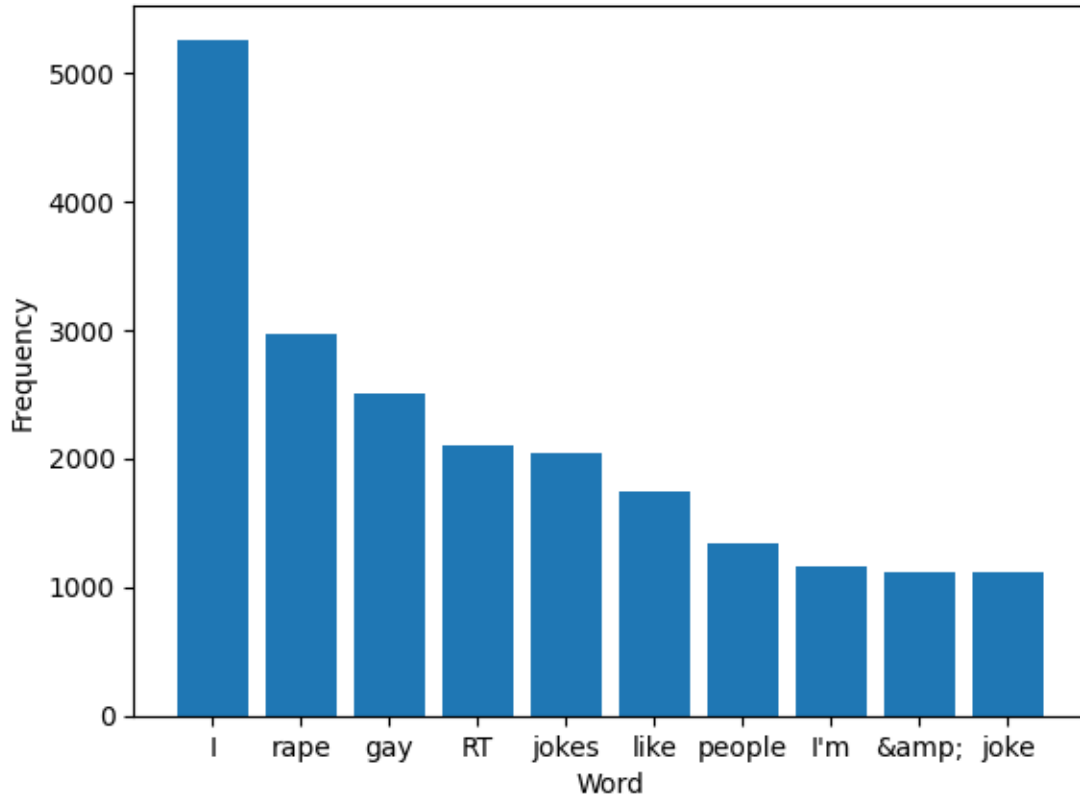**Distributions**
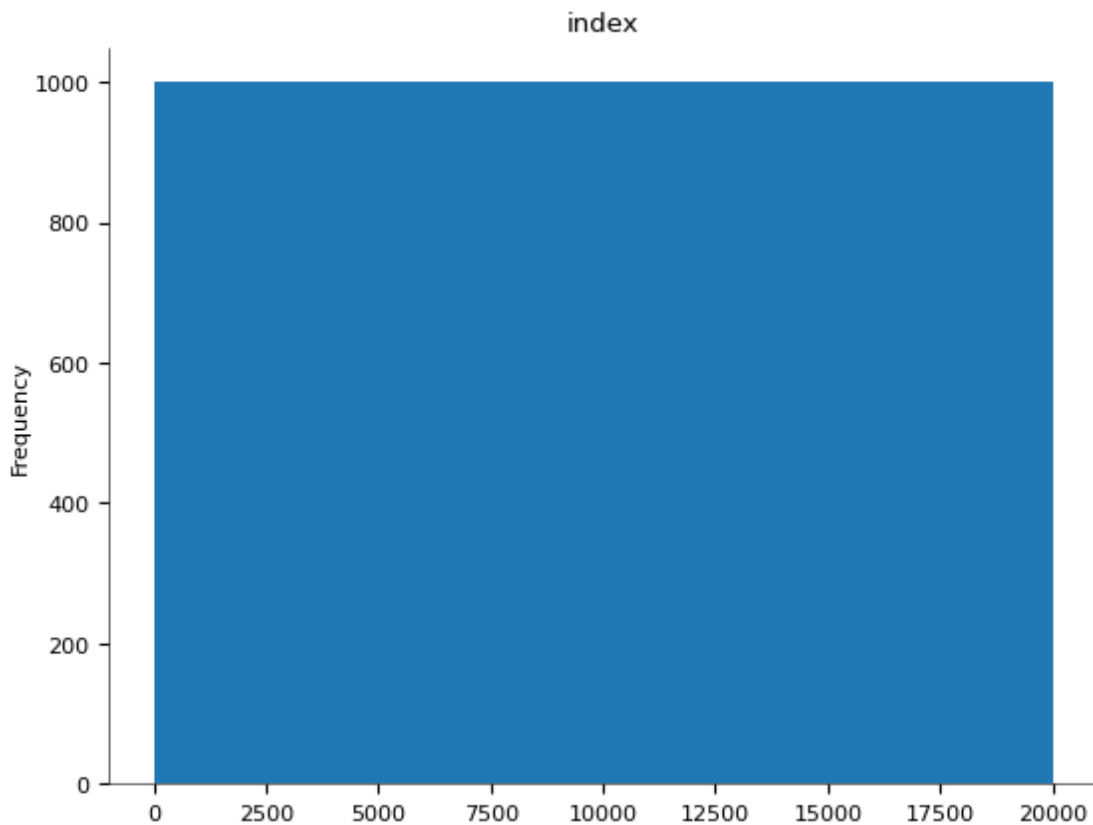
Categorical distributions

Class Distribution of Cyberbullying Types

Most Commonly Used Words

Most frequent words used in Tweets
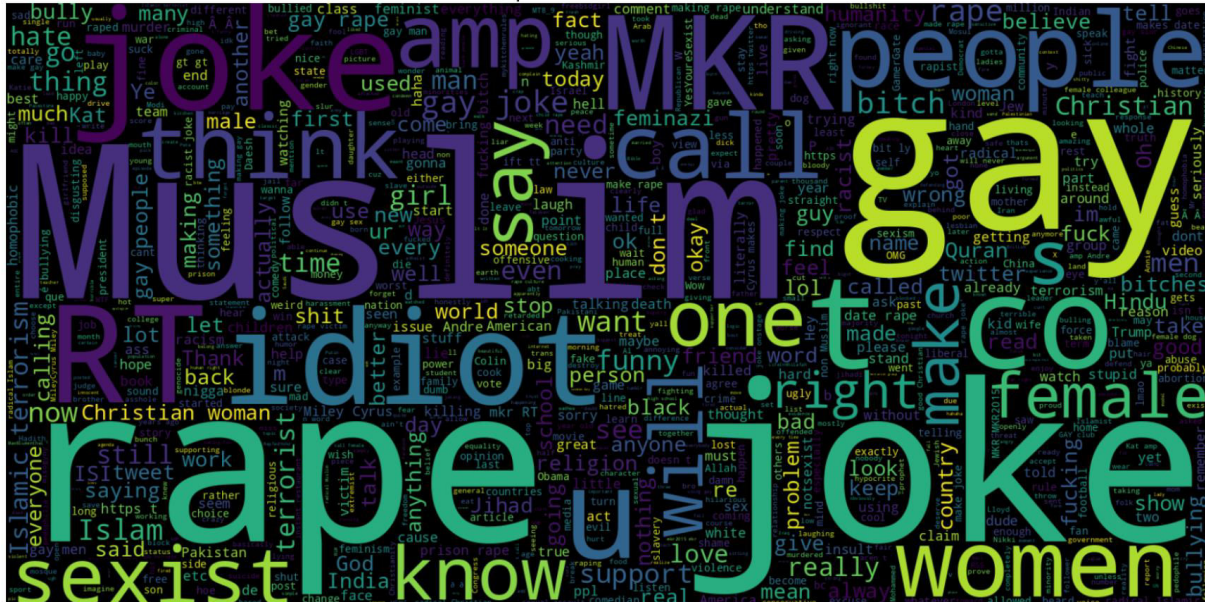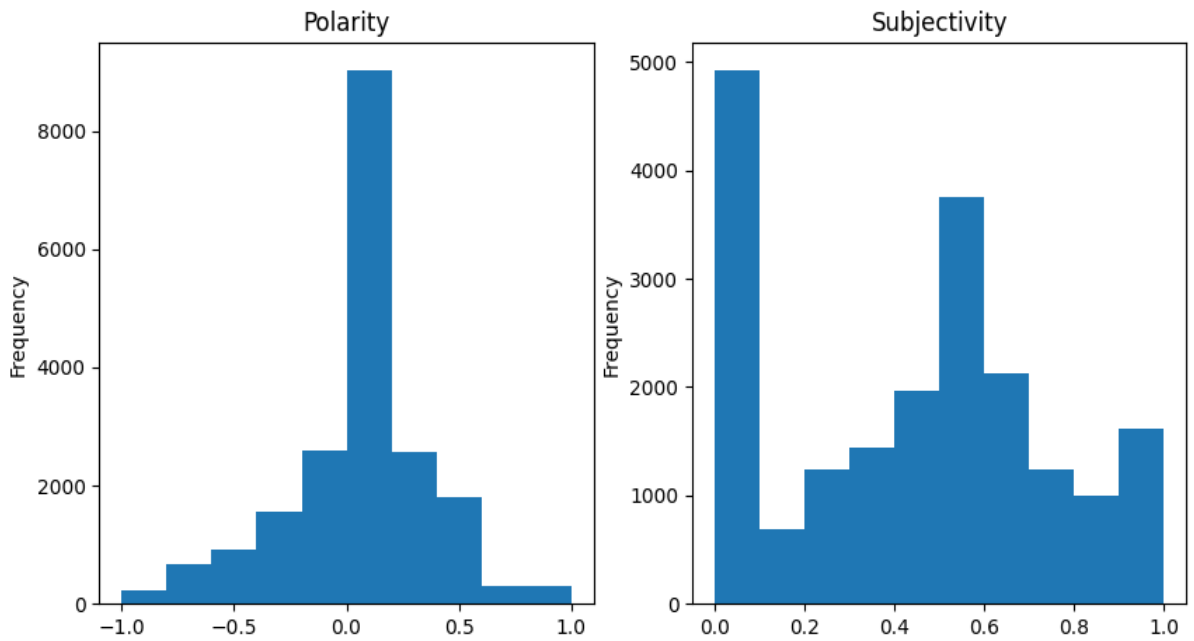
INTERNATIONAL JOURNAL OF

MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |