Department of MSc (IT)

A Mini-project Report
On

"Cyberbullying Prediction using Naïve Bayes Classifier"

Submitted as a part of [**18MSIT3H06L-** Mini-project Lab]
During the **III** Semester [**July 2020- Dec 2020**]

Submitted by

Name of Student: Meghaben Mistry **USN:** 19MSRIT034

Internal Guide

Dr. Ganesh. D
Department of MSc - IT
School of CS & IT, Jain [Deemed to-be University]

Jain Knowledge Campus

#44 – 4, District Fund Road, Jayanagar 9th Block,

Bangalore – 560 069.

Department of Master of Science in Information Technology

Certificate

This is to certify the project entitled **"Cyberbullying Prediction using Naïve Bayes Classifier"** is a bonafide work carried out by **Meghaben M. Mistry [19MSRIT034]** in a partial fulfillment for the award of **Degree of Master of Science in Information Technology** of Jain [Deemed-to-be] University during the year **2020- 2021.** It is certified that all correction / suggestion indicated for internal assessment have been incorporated in the report deposited in the Department library. The project has been approved as it satisfies the academic requirements in respect to the project work prescribed for the Master of Science in Information Technology.

Signature of HOD

Dr. Suchitra. R
Head of Department MSc (IT),
Jain (Deemed-to-be) University.

Signature of Guide

Dr. Ganesh D
Asst. Professor,
School of CS & IT,
Jain (Deemed-to-be) University.

Name of the Examiner **Signature with date**

1. _____ **_____**

2. _____ **_____**

ACKNOWLEDGEMENT

It is not the completion of the project that is most important but more so, the interaction of roles played by various people in the satisfactory completion. I take this opportunity to express my deep gratitude and appreciation to all those who encouraged me to successfully complete the project.

With profound sense of gratitude and regards, I acknowledge with great pleasure the guidance and support extended by, **Dr. Eshwaran Iyer,** Dean, Jain Knowledge Campus, Bangalore, **Dr. Nachappa. MN,** Head, School of CS & IT, JAIN University, **Dr. Dinesh Nilkant** center head, Jain knowledge Campus, **Dr. Suchitra. R,** HOD, Department of MScIT, **Dr. Ganesh. D,** Project Guide, Department of Computer Science & IT, Jain (Deemed-to-be) University, Bangalore, for their interest & encouragement throughout the project.
I would like to express my deep sense of gratitude to my Guide **Mr. Ganesh D,** Asst. Professor, Department of Computer Science & IT, Jain (Deemed-to-be) University, Bangalore for his accomplishment and valuable information, direction and sense of perfection to work. He had been the main source of inspiration for the completion of work and strengthening confidence.

I would also thank my parents for their understanding & encouragement, Department Staffs, teaching and non-teaching, my friends, one and all those who helped me to complete this project successfully.

Ms. Meghaben Mistry
[19MSRIT034]

ABSTRACT

Technology has made our life so easy. Everyone including any age group using technology in many ways like mobile, laptops, computers, etc, . Because of Technology there are to many cybercrimes. In India, the rate of cybercrime is increasing day by day. One of the cybercrime is Cyberbullying or Cyberharassment.

Increase the use of social media has also increased the cases of cyberbullying. In this paper, we are predicting the cyberbullying percentage of Twitter data using the Naive Bayes classifier and text Mining. Naive Bayes Classifier is the statistic analysis algorithm which uses the byes theorem to predict the probability of class member.

Index terms – Cybercrime, Cyberbullying, Cyberharassment, Naïve Bayes Classifier, Text Mining.

CONTENTS

Chapter 1: Preamble

1.1 Introduction

Technology has helped to bridge a global gap during the age of globalization. This bridge has allowed multitudes to learn about the world, and connect with others, in ways that were previously impossible [6]. Technology has both positive and negative impacts on humans. Technology made our life so easy and flexible. But many crimes can happen through technology, which is called Cyber Crime. Cybercrime may threaten a person, company or a nation's security and financial health.

Digital India may have become a soft target for criminals as country recorded a huge increase of 63.5 percent in cyber crime cases in the year 2019, showed the National Crime Record Bureau data. The NCRB's data stated that 4,4546 cases of cyber crimes were registered in 2019 as compared to 28,248 in 2018. The data showed in 60.4 percent of cases, registered fraud was the motive followed by sexual exploitation (5.1%) and causing disrepute (4.2%) [11].

Cyberbullying or Cyberharassment is also one of the Cyber Crimes, which is a form of bullying or harassment using electronic means. Cyberbullying and cyberharassment are also known as online bullying. It has become increasingly common, especially among teenagers, as the digital sphere has expanded and technology has advanced.

Cyberbullying is when someone, typically a teenager, bullies or harasses others on the internet and in other digital spaces, particularly on social media sites. Harmful bullying behavior can include posting rumors, threats, sexual remarks, victims' personal information. Bullying or harassment can be identified by repeated behavior and an intent to harm [11].

Using the Naive Bayes Classifier, the prediction of bullying and non-bullying sentences is generated. Before applying the Naive Bayes classifier Algorithm, First need to preprocess and manage the data using Tokenization, Stop word removal, Lemmatization, etc,.

1.2 Literature Review

The survey paper by Anurag Pandey et.al [6] tried to address this issue by reviewing the steps that can be undertaken to detect cyber bullying on online social networks. This paper aims to review the different methods and algorithms used for detection in cyber bullying and provide a comparative study amongst them so as to decide which method is the most effective approach and provides the best accuracy.

The review paper done by Paul Bocij [7] describes the first study to focus exclusively on the prevalence and impact of cyberstalking. A Web-based questionnaire was used to collect data from a group of respondents who were recruited by snowball sampling via e-mail. A total of 169 respondents completed the questionnaire. The results of the study found that approximately a third of respondents might be considered victims of cyberstalking. Furthermore, when asked to indicate the level of distress felt as a result of their experiences, almost a quarter of respondents chose a value of ten on a ten-point scale.

The review paper by Per Carlbring et.al [8] discusses about the factors which are involved in cyberstalking. He has also mentioned the problems faced by the victims of cyberstalking such as loneliness, anxiety, self-harming behavior, sleeping problems etc. He has also mentioned some therapies for the victims that will help them to deal with the consequences of cyber bullying like REBT (Relational Emotive Therapy), CBT (Cognitive Behavioral Therapy), IPT (Interpersonal Psychological Therapy) etc.

The study done by Qing Li [9] examines the nature and extent of adolescents' cyberbullying experiences, and explores the extent to which various factors, including bullying, culture, and gender, contribute to cyberbullying and cyber victimization in junior high schools. In this study, one in three adolescents was a cyber victim, one in five was a cyberbully, and over half of the students had either experienced or heard about cyberbullying incidents. Close to half of the cyber victims had no idea who the predators were. Culture and engagement in traditional bullying were strong predictors not only for cyberbullying, but also for cyber victimization. Gender also

played a significant role, as males, compared to their female counterparts, were more likely to be cyberbullies.

1.3 Basic theory

Cyber Bullying

Cyberbullying can affect human beings in many ways. It mainly happened with kids, teenagers, women. It can affect human beings in many ways. Cyberbullying victims experience anxiety, fear, depression, negativity, and low self-esteem. It may lead victims to suicide also.

Victims may have lower self-esteem, increased suicidal ideation, and a variety of emotional responses, including being scared, frustrated, angry, and depressed. Cyberbullying may be more harmful than traditional bullying, because there is no escaping it. One of the most damaging effects is that a victim begins to avoid friends and activities, which is often the very intention of the bully. Cyberbullying campaigns are sometimes so damaging that victims have committed suicide.

Text Mining

Expressions of text mining is generally used to indicate the system analyze of the large data in the form of a natural language text and detect usage patterns or linguistic lexical in an attempt to extract useful information.

The process of text mining requires several stages, remembering the text data has characteristics that are more complex than the usual data. Generally a document has the following characteristics:

1) Text on Database has a large size (Large textual database)
2) Have a high dimension, one word one dimension (high dimension)
3) Contain the phrase and among other phrases with one phrase can have different meanings and interdependent of each other (dependency)
4) Many words/sentences contain ambiguous (Ambiguity)
5) Contain data noise, such as an abbreviation. Terminology and spelling mistake.
6) Contains structures that are not raw suppose abbreviations on words.

Process of extracting the information from a set of text documents such as web pages, twitter, and other documents need several interrelated processes. Processing unstructured documents from being more structured by applying some of the techniques of extraction and filtering on words in the document at once by weighting the importance of words with weighting method.

Tokenization:

This involves conversion of the large blob of unstructured text into a set of tokens divided by white spaces and/or punctuation marks, classified into categories such as words, phrases and sentences.

Example: "1 voted for Sammy because he was most inclined and considerate towards my ideals", she said. By applying tokenization we get the following set {[I], [voted],[for],[Sammy],[because],[he],[was],[most],[inclined], [and],[considerate],[towards],[my], [ideals], [she],[said]}

Stop words Removal:

Some very common words such as "a", "and', "are" etc add very little to the meaning of the text and are of little value in helping classify text. These 'Stop Words' are dropped from the text to ensure easier analysis of the text in further steps.

Using our above example the new set we get is
{[I],[voted],[Sammy],[because],[he],[most],[inclined],[considerate],[towards],[my],[ideals],[she],[said]}
The words 'for', 'and' and 'was' are removed.

Lemmatization:

Lemmatization is a refinement of the stemming technique that makes use of a dictionary-based approach for morphological analysis of words to obtain the base form of a word, called a lemma. Porter's algorithm is widely used for lemmatization, as it has been consistently shown to be effective. In reference to the above example, simple stemming would transform it into:

{[Me],[vote],[Sammy],[cause],[he],[most],[incline],[considerate],[toward],[me],[ideal],[she],[say]}

Naïve Bayes Classifier

Naïve Bayes statistics analysis is an algorithm, which performs data processing numeric data using Bayesian probability. Classification of Bayes statistics classification which is able to predict the probability of a class member.

Broadly speaking, the workings of this method can be explained as follows:

1) Take the probability of positive and negative words.
2) Calculate the average probability of both.
3) Specify the classification based on the value of the above probability.

To get the probability of each word, passing the learning each word and the probability. In this learning process, needed a training set, which is a set of sentences is positive and negative have been classified.

Naïve Bayes classification technique is a simple and fast. This technique works well with representation statistics. Unlike the method of rule-based, Naïve Bayesian can learn incrementally. But the shortcomings of the Naïve Bayesian vector is the size of the resulting feature quite large and need a technique to minimize the size of the vector.

Now, with regards to our dataset, we can apply Bayes' theorem in following way:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

where, y is class variable and X is a dependent feature vector (of size *n*) where:

$$X = (x_1, x_2, x_3, ..., x_n)$$

Chapter-2 : Research Findings

2.1 Problem Statement

Cyberbullying is bullying that takes place in cyberspace through various mediums including online chats, text messages and e-mails. It is a big problem on social media websites like Facebook and Twitter. Many individuals, especially adolescents, suffer negative effects such as depression, sleeplessness, lowered self-esteem and even lack of motivation to live when being targeted by bullies on social media. Much is being done to stop regular bullying in schools.

Cyberbullying on the other hand can be difficult to detect and stop due to it happening online, often hidden from the eyes of parents and teachers. The problem we face is to come up with a technological approach that can aid in automatic detection of bullying on social media. The approach we will investigate is a system capable of automatically detecting and reporting instances of bullying on social media platforms.

Victims of cyberbullying can suffer many negative effects such as not being accepted in their peer groups (leading to loneliness

and isolation), low self-esteem, depression, poor academic achievement and decreased emotional well-being . Cyberbullied individuals are also more likely to suffer from headaches, abdominal pain and sleeplessness. In some extreme cases cyberbullying has also been linked to suicide.

India has high occurrences of bullying instances. 79% Indians are aware and worried about cyber bullying in comparison to 54% worldwide. 53% Indians have been bullied compared to a worldwide average of 37%. In addition to this, 50% Indians have been involved in bullying someone online while worldwide only 24% of the population has been involved in similar instances. On an upside 63% Indians are educated about and 76% institutions have a formal policy on cyber bullying in comparison to a worldwide average of 23% and 37% .

Chapter 3: SOFTWARE REQUIREMENT SPECIFICATIONS

Internal Interface Requirements

Identify the product whose software requirements are specified in this document, including the revision or release number. Describe the scope of the product that is covered by this SRS, particularly if this SRS describes only part of the system or a single subsystem.

Describe any standards or typographical conventions that were followed when writing this SRS, such as fonts or highlighting that have special significance. For example, state whether priorities for higher-level requirements are assumed to be inherited by detailed requirements, or whether every requirement statement is to have its own priority.

Describe the different types of reader that the document is intended for, such as developers, project managers, marketing staff, users, testers, and documentation writers. Describe what the rest of this SRS contains and how it is organized.

Suggest a sequence for reading the document, beginning with the overview sections and proceeding through the sections that are most pertinent to each reader type.

Provide a short description of the software being specified and its purpose, including relevant benefits, objectives, and goals. Relate the software to corporate goals or business strategies. If a separate vision and scope document is available, refer to it rather than duplicating its contents here.

The recent explosion in data pertaining to users on social media has created a great interest in performing sentiment analysis on this data using Big Data and Machine Learning principles to understand people's interests. This project intends to perform the same tasks. The difference between this project and other sentiment analysis tools is that, it will perform real time analysis of tweets based on hashtags and not on a stored archive.

Describe the context and origin of the product being specified in this SRS. For example, state whether this product is a; follow-on member of a product family, a replacement for certain existing systems, or a new, self-contained product.

If the; SRS defines a component of a larger system, relate the requirements of the larger system to the functionality of this; software and identify interfaces between the two. A simple diagram that shows the major components of

the overall system; subsystem interconnections, and external interfaces can be helpful.

The Product functions are:

• Collect tweets in a real time fashion i.e., from the twitter live stream based on specified hashtags

• Remove redundant information from these collected tweets.

• Perform Sentiment Analysis on the tweets stored in the database to classify their nature viz. positive, negative and so on.

• Use a machine learning algorithm which will predict the ‘mood’ of the people with respect to that topic.

Summarize the major functions the product must perform or must let the user perform; so only a high-level summary (such as a bullet list) is needed here. Organize the functions to make them understandable to; any reader of the SRS.

A picture of the major groups of related requirements and how they relate, such as a top-level data; flow diagram or object class diagram, is often effective.

Identify the various user classes that you anticipate will use this product. User classes may be differentiated based on frequency of use, subset of product functions used, technical expertise, security or privilege levels, educational level, or experience.

Describe the pertinent characteristics of each user class. Certain requirements may pertain only to certain user classes.

Distinguish the most important user classes for this product from those who are less important to satisfy.
Describe the environment in which the software will operate, including the hardware platform, operating system and versions, and any other software components or applications with which it must peacefully coexist.

External Interface Requirement

We classify External Interface in 4 types, those are:

User Interface:

Describe the logical characteristics of each interface between the software product and the users. This may

include sample screen images, any GUI standards or product family style guides that are to be followed, screen layout constraints, standard buttons and functions (e.g., help) that will appear on every screen, keyboard shortcuts, error message display standards, and so on.

Define the software components for which a user interface is needed. Details of the user interface design should be documented in a separate user interface specification.

Hardware interface:

Describe the logical and physical characteristics of each interface between the software product and the hardware components of the system. This may include the supported device types, the nature of the data and control interactions between the software and the hardware, and communication protocols to be used.

Software Interface:

Describe the connections between this product and other specific software components (name and version), including databases, operating systems, tools, libraries, and integrated commercial components. Identify the data items or messages coming into the system and going out

and describe the purpose of each. Describe the services needed and the nature of communications. Refer to documents that describe detailed application programming interface protocols.

Identify data that will be shared across software components. If the data sharing mechanism must be implemented in a specific way (for example, use of a global data area in a multitasking operating system), specify this as an implementation constraint.

Communication Interface:

Describe the requirements associated with any communications functions required by this product, including e-mail, web browser, network server communications protocols, electronic forms, and so on. Define any pertinent message formatting. Identify any communication standards that will be used, such as FTP or HTTP. Specify any communication security or encryption issues, data transfer rates, and synchronization mechanisms.

Non-functional Requirements

Performance Requirements:

If there are performance requirements for the product under various circumstances, state them here and explain their rationale, to help the developers understand the intent and make suitable design choices. Specify the timing relationships for real time systems.
Make such requirements as specific as possible. You may need to state performance requirements for individual functional requirements or features.

Safety Requirements:

Specify those requirements that are concerned with possible loss, damage, or harm that could result from the use of the product. Define any safeguards or actions that must be taken, as well as actions that must be prevented. Refer to any external policies or regulations that state safety issues that affect the product's design or use. Define any safety certifications that must be satisfied.

Security Requirements:

Specify any requirements regarding security or privacy issues surrounding use of the product or protection of the data used or created by the product. Define any user identity authentication requirements. Refer to any external policies or regulations containing security issues that affect the product. Define any security or privacy certifications that must be satisfied.

Software Quality Attributes:

Specify any additional quality characteristics for the product that will be important to either the customers or the developers. Some to consider are: adaptability, availability, correctness, flexibility, interoperability, maintainability, portability, reliability, reusability, robustness, testability, and usability.

Write these to be specific, quantitative, and verifiable when possible. At the least, clarify the relative preferences for various attributes, such as ease of use over ease of learning.

Other Requirements:

- Linux Operating System/Windows
- Python Platform (Anaconda2, Spyder, Jupyter)
- NLTK package
- Modern Web Browser
- Numpy package
- Panda package

Chapter-4 : Motivation

Cyberbullying can take place on social media sites such as Facebook, Myspace, and Twitter. "By 2008, 93% of young people between the ages of 12 and 17 were online. In fact, youth spend more time with media than any single other activity besides sleeping." The last decade has witnessed a surge of cyberbullying, which is categorized as bullying that occurs through the use of electronic communication technologies, such as e-mail, instant messaging, social media, online gaming, or through digital messages or images sent to a cellular phone.

There are many risks attached to social media sites, and cyberbullying is one of the larger risks. One million children were harassed, threatened or subjected to other forms of cyberbullying on Facebook during the past year, while 90 percent of social-media-using teens who have witnessed online cruelty say they have ignored mean behavior on social media, and 35 percent have done so frequently.

Ninety-five percent of social-media-using teens who have witnessed cruel behavior on social networking sites say they

have seen others ignoring the mean behavior, and 55 percent have witnessed this frequently. Terms like "Facebook depression" have been coined specifically in regard to the result of extended social media use, with cyberbullying playing a large part in this.

Chapter-5 : Research Objective

Cyberbullying and social networking harassment are the two areas where textual patterns have been used to detect and filter unwanted messages [2]. It is increasing amongst teenagers as we can say that when someone harasses any other person on the internet or on any particular social media site it leads to some harmful bullying behaviour like posting rumours, threats etc.

However, bullying or harassment over the internet is very harmful to a lot of people as they go through mental damage as due to bullying some people threaten others to post their photos or nudes over the internet whereas to avoid this they have to fulfil the demands of the culprits. Cyberbullying is an illegal and unlawful activity. The internet is a place where cyberbullying is very common. Social media sites like – Facebook, Snap chat, Instagram and other common sites [10].

Cyberbullying is a serious issue as it is very difficult for the person who is suffering from it. Now it's the duty of that person who is going through this situation to know how to handle it up by sharing the problem with their close ones. Victims must be identified and family members should support them. Teachers should give awareness to the students about cyberbullying and

other negative effects of social media and students should feel free to consult their teacher and parents.

There are various laws established in the country because the increase in cybercrime country has launched many laws, rules and regulations. The rules are designed to protect the victim and keep the case information confidential and there are different government awareness programs which aware people about cyberbullying and people must cooperate and support the victims who are suffering through this situation[10].

Chapter-6 : Methodology

Here we are using the method in the following format.

1) First step is to select the dataset online from various sides. Dataset of cyberbullying mainly contains the user comments, posts, images, etc, . Here the dataset used already contains the decisions of the sentences which are bullying or non-bullying.

2) Second step is to preprocess the data so that only relevant information remains while classifying the data. It contains the removal of stop words, tokenization, lemmatization, a bag of words.

3) Next step is to use the Naive Bayes Classifier. In that first divide data as it will easy for applying the in the classifier. After applying it will give the prediction of the bullying and non-bullying sentences.

Chapter-7 : System Design

7.1 Input Screen Design

Here we are using the python programming language for the prediction of cyberbullying using the Naïve Bayes Classifier. Using Python, you can program machines to analyze text from surveys, social media mentions, product reviews, and more. Python is an interpreted, high-level and general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented, and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library.

For text mining nltk is used. NLTK stands for *Natural Language Toolkit*. This is a suite of libraries and programs for

symbolic and statistical NLP for English. It ships with graphical demonstrations and sample data. First getting to see the light in 2001, NLTK hopes to support research and teaching in NLP and other areas closely related.

NLTK is a powerful Python package that provides a set of diverse natural languages algorithms. It is free, opensource, easy to use, large community, and well documented. NLTK consists of the most common algorithms such as tokenizing, part-of-speech tagging, stemming, sentiment analysis, topic segmentation, and named entity recognition. NLTK helps the computer to analysis, preprocess, and understand the written text.

With the help of nltk.tokenize.word_tokenize**()** method, we are able to extract the tokens from string of characters by using tokenize.word_tokenize() method. It actually returns the syllables from a single word. A single word can contain one or two syllables.

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 valuable processing time. For this, we can remove them easily, by storing a list of words that you consider to stop words. NLTK(Natural Language Toolkit) in python has a list of stopwords stored in 16 different languages. You can find them in the nltk_data directory.

Lemmatization is the process of grouping together the different inflected forms of a word so they can be analysed as a single item. Lemmatization is similar to stemming but it brings context to the words. So it links words with similar meaning to one word. Text preprocessing includes both Stemming as well as Lemmatization. Many times people find these two terms confusing. Some treat these two as same. Actually, lemmatization is preferred over Stemming because lemmatization does morphological analysis of the words.

The multinomial Naive Bayes classifier is suitable for classification with discrete features (e.g., word counts for text classification). The multinomial distribution normally requires integer feature counts. However, in practice, fractional counts such as tf-idf may also work.

7.2 Input Code Design

```
#Import libraries

import pandas as pd

import numpy as np

import nltk

from nltk.corpus import stopwords

from nltk.tokenize import word_tokenize

import string

from nltk.stem import WordNetLemmatizer

from nltk.collocations import BigramCollocationFinder

from nltk.metrics import BigramAssocMeasures

nltk.download('punkt')

nltk.download('stopwords')

nltk.download('wordnet')
```

```
from sklearn.naive_bayes import MultinomialNB

from sklearn.metrics import roc_curve, auc, roc_auc_score



--------------------------------------------------------------------



#Load dataset

import pandas as pd

df1 = pd.read_csv("cleanprojectdataset.csv")



--------------------------------------------------------------------



print(df1)

--------------------------------------------------------------------



#Tokenize words and labels into lists

Tweet = []
```

```
Labels = []


for row in df1['Tweet']:
    #tokenize words
    words = word_tokenize(row)
    #remove punctuations
    clean_words = [word.lower() for word in words if word not in set(string.punctuation)]
    #remove stop words
    english_stops = set(stopwords.words('english'))
    characters_to_remove = ["""",'``',"rt","https","’","“","”","\u200b","--","n't","'s","...","//t.c" ]
    clean_words = [word for word in clean_words if word not in english_stops]
    clean_words = [word for word in clean_words if word not in set(characters_to_remove)]
```

```
    #Lematise words
    wordnet_lemmatizer = WordNetLemmatizer()
    lemma_list = [wordnet_lemmatizer.lemmatize(word) for word in clean_words]
    Tweet.append(lemma_list)

    for row in df1["Text Label"]:
        Labels.append(row)

---------------------------------------------------------------------

#combine them to create bag of words
combined = zip(Tweet, Labels)
---------------------------------------------------------------------

#Create bag of words and dictionary object
```

```
def bag_of_words(words):
    return dict([(word, True) for word in words])
```

```
#Key, Value Pair into new list for modeling
Final_Data = []
for r, v in combined:
    bag_of_words(r)
    Final_Data.append((bag_of_words(r),v))
```

```
#random shuffle
import random
random.shuffle(Final_Data)
```

```
print(len(Final_Data))

----------------------------------------------------------------------

train_set, test_set = Final_Data[0:747], Final_Data[747:]

import nltk
import collections
from nltk.metrics.scores import (accuracy, precision, recall, f_measure)
nb_classifier = nltk.NaiveBayesClassifier.train(train_set)
nb_classifier.show_most_informative_features(10)

from nltk.classify.util import accuracy
print(accuracy(nb_classifier, test_set))
```

```
refsets = collections.defaultdict(set)

testsets = collections.defaultdict(set)


for i, (Final_Data, label) in enumerate(test_set):
    refsets[label].add(i)
    observed = nb_classifier.classify(Final_Data)
    testsets[observed].add(i)


print('bullying precision:', precision(refsets['Bullying'], testsets['Bullying']))

print('bullying recall:', recall(refsets['Bullying'], testsets['Bullying']))

print('bullying F-measure:', f_measure(refsets['Bullying'], testsets['Bullying']))

print('not-bullying precision:', precision(refsets['Non-Bullying'], testsets['Non-Bullying']))
```

```
print('not-bullying recall:', recall(refsets['Non-Bullying'], testsets['Non-Bullying']))

print('not-bullying F-measure:', f_measure(refsets['Non-Bullying'], testsets['Non-Bullying']))

----------------------------------------------------------------------
```

7.3 Report

- First import pandas and numpy using import statement as give it alias as pd and np respectively.

- Import nltk and other library from it like stopwords, word_tokenize, WordNetLemmatizer.

- Download ‘punkt’, ‘stopwords’, ‘wordnet’ for supporting files.

- import MultinomialNB from sklearn.naive_bayes for use of naïve bayes Classifier.

- After that import roc_curve, auc, roc_auc_score

In [1]:
```
import pandas as pd
import numpy as np
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
import string
from nltk.stem import WordNetLemmatizer
nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import roc_curve, auc, roc_auc_score
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Megha\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\Megha\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\Megha\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
```

- Read the dataset file using pd.read_csv.

- Replace the ‘Text Label’ column in dataset, bullying with 1 and non-bullying with 0.

```
In [3]: #Load dataset
        import pandas as pd
        df1 = pd.read_csv("cleanprojectdataset.csv")
        df1['Text Label'] = np.where(df1['Text Label']=='Bullying',1,0)
```

- Print the dataset after replacing ‘Text Label’ column.

```
In [4]: print(df1)
```

```
                                                  Tweet  Text Label
0     .omg why are poc wearing fugly blue contacts s...           0
1     .Sorry but most of the runners popular right n...           0
2     .those jeans are hideous, and I?m afraid he?s ...           0
3     .I had to dress up for a presentation in class...           0
4     .Am I the only one who thinks justin bieber is...           0
...                                                 ...         ...
1060  No we are not, But you are a race baiting libt...           1
1061  you wont get anyone for this challenge., after...           1
1062  I will follow you if you are not a libtard,Mus...           1
1063  michaelianblack Ur a child, an ostrich w/ your...           1
1064  FoxNews. not to all the ppl I know that live t...           1

[1065 rows x 2 columns]
```

- Use word_tokenize() for tokenization.
- Remove the punctuation from every Tweet.
- Remove Stop word using stopwords.word()
- Perform Lemmitization using wordnet_lemitizer().

In [5]:
```
#Tokenize words and labels into lists

Tweet = []
Labels = []

for row in df1['Tweet']:
    #tokenize words
    words = word_tokenize(row)
    #remove punctuations
    clean_words = [word.lower() for word in words if word not in set(string.punctuation)]
    #remove stop words
    english_stops = set(stopwords.words('english'))
    characters_to_remove = ["''","``","rt","https",")","(","”","\u200b","--","n't","'s","...","//t.c" ]
    clean_words = [word for word in clean_words if word not in english_stops]
    clean_words = [word for word in clean_words if word not in set(characters_to_remove)]
    #Lematise words
    wordnet_lemmatizer = WordNetLemmatizer()
    lemma_list = [wordnet_lemmatizer.lemmatize(word) for word in clean_words]
    Tweet.append(lemma_list)

    for row in df1["Text Label"]:
        Labels.append(row)
```

- Combine both Tweet and Labels using zip() method.

```
In [6]: #combine them to create bag of words
        combined = zip(Tweet, Labels)
```

- Create bag of words using Dictionary.

```
In [8]: #Create bag of words and dictionary object
        def bag_of_words(words):
            return dict([(word, True) for word in words])
```

- Create one Final_Data named list and append bag of words and Labels in that list.

```
In [8]: #Key, Value Pair into new list for modeling
        Final_Data = []
        for r, v in combined:
            bag_of_words(r)
            Final_Data.append((bag_of_words(r),v))
```

- Print the length of the Dataset.

```
In [9]: #random shuffle
        import random
        random.shuffle(Final_Data)
        print(len(Final_Data))
        #print(Final_Data)

        1065
```

- Import train_test_split to split data into train and test.

```
In [10]: from sklearn.model_selection import train_test_split

         X_train, X_test, y_train, y_test = train_test_split(df1["Tweet"],
                                                             df1["Text Label"],
                                                             random_state=0)
```

```
In [11]: def answer_one():

             return len(df1[df1['Text Label']==1])/len(df1['Text Label'])*100

         answer_one()

Out[11]: 40.093896713615024
```

```
In [8]: from sklearn.feature_extraction.text import CountVectorizer

        def answer_two():

            # List all tokens and their counts as a dictionary:
            vocabulary = CountVectorizer().fit(X_train).vocabulary_

            # You want only the keys, i.e, the words:
            vocabulary = [x for x in vocabulary.keys()]

            # Store the lengths in a separate list:
            len_vocabulary = [len(x) for x in vocabulary]

            # Use the index of the longest token:
            return vocabulary[np.argmax(len_vocabulary)]

        answer_two()

Out[8]: 'healtheworldin5words'
```

- Use the MultinominalNB() for prediction of Bullying statement over the Non-Bullying Statement.

In [9]:
```
def answer_three():

    cv = CountVectorizer().fit(X_train)

    # Transform both X_train and X_test with the same CV object:
    X_train_cv = cv.transform(X_train)
    X_test_cv = cv.transform(X_test)

    # Classifier for prediction:
    clf = MultinomialNB(alpha=0.1)
    clf.fit(X_train_cv, y_train)
    preds_test = clf.predict(X_test_cv)

    return roc_auc_score(y_test, preds_test)

answer_three()
```

Out[9]:
```
0.6522959183673469
```

Chapter-8 : Conclusion

The rapid increase of social networks has shown a consistent growth in cyberbullying activities. Cyberbullying has become a major social problem. Cyberbullying has become an important area of research due to its impact on society. Various researches try to recognize the reason of cyberbullying and its aftereffect. But only a few try to enhance software to prohibit cyberbullying. Robust and selective representation of learning of text messages is crucial for consistent detection system.

Machine Learning representation and authentication makes automatic revelation of bullied messages in online media possible and ensures building a relevant and clear social media environment. The Email based cyber stalking is also a huge problem. Email based cyber stalking detection involves two phases; the first is to analyze and detect cyber stalking emails and the second phase is to verify the proof for finding out the cyber stalkers as a prohibition and detection mechanism.

We have used the Naïve Bayes classifier for easy use. It is not time consuming. It is the best algorithm for retrieving probability between every class of dataset. The best thing about Naïve Bayes Classifier is it is comparing each and every class with each other.

Chapter-9 : References

1) H.M. Chandana, Karnik Pooja Jagdish, Anna Mary, Bhawana Dorbi, Naveen Chandra Gowda "Cognitive Way of Detecting Cyberbullying in Chatbots" International Journal of Advanced Research in Computer Science Volume 9, Special Issue No. 3, May 2018

2) Arnisha Akhter, Uzzal K. Acharjee, Md Masbaul A. Polash "Cyber Bullying Detection and Classification using Multinomial Naïve Bayes and Fuzzy Logic" https://www.researchgate.net/publication/337144755

3) Rekha Sugandhi, Anurag Pande, Siddhant Chawla, Abhishek Agrawal, Husen Bhagat "Methods for Detection of Cyberbullying: A Survey"

4) Hariani, Imam Riadi "Detection Of Cyberbullying On Social Media Using Data Mining Techniques"

International Journal of Computer Science and Information Security (IJCSIS), Vol. 15, No. 3, March 2017

5) Vandana Nandakumar, Binsu C. Kovoor, Sreeja M. U. "CYBERBULLYING REVELATION IN TWITTER DATA USING NAÏVE BAYES CLASSIFIER ALGORITHM" International Journal of Advanced Research in Computer Science, Volume 9, No. 1, January-February 2018

6) Rekha Sugandhi, Anurag Pande, Siddhant Chawla, Abhishek Agrawal, Husen Bhagat; Methods for detection of cyberbullying : A survey, In: *Intelligent Systems Design and Applications (ISDA)*, 2015.

7) Paul Bocij, Victims of Cyberstalking: An Exploratory Study of Harassment Perpetrated via the Internet, vol 8, number 10-6, 2003.

8) Agate, J., Ledward, J., 2013. Social media: how the net is closing in on cyber bullies. Entertain. Law Rev. 24 (8), 263-268.

9) Qing Li; Bullying in the new playground: Research into cyberbullying and cyber victimization; In: *Australian Journal of Education Technology*, 2007, 23(4), 435-454.

10) https://en.wikipedia.org/wiki/Cyberbullying

11) https://www.ndtv.com/india-news/digital-india-sees-63-5-increase-in-cyber-crime-cases-shows-data-2302958

12) https://timesofindia.indiatimes.com/readersblog/riyable/effect-of-cyberbullying-5884/

13) https://www.brainspire.com/blog/what-are-the-effects-of-technology-on-human-interaction

14) https://www.linguamatics.com/