International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 9 Issue 6 ǁ 2021 ǁ PP. 06-13

Detection of Cyberbullying Incidents on Instagram Social Network

Pratik Ziman [1],Chinmay Gaikwad [2], Anita Mhatre[3]
1 BE Student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, Maharashtra
2 BE Student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, Maharashtra
3 Professor, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, Maharashtra

Abstract
Worldwide accessibility to the Internet has incredibly reshaped our perception of the world. One of the children of the World Wide Web is Social Media (SM), which is present in many forms: online game platforms, dating apps, forums, online news services, and social networks. Different social networks aim at different objectives: opinion transmission (Instagram, Instagram, Facebook, etc.), business contacts (LinkedIn), image sharing (Instagram), video transmission (YouTube), dating (Metric), and so on. However, they all have one thing in common: they aim to connect people. Among the many existing social networks, Instagram currently ranks as one of the leading platforms and allows users to post their pictures & videos and other users can see and can comment on them. In recent years, social networks (and especially Instagram) have been used to spread hate messages. Hate speech refers to a kind of speech that denigrates a person or multiple persons based on their membership to a group, usually defined by race, ethnicity, sexual orientation, gender identity, disability, religion, political affiliation, or views.We study detection of Cyberbullying in photo sharing networks, with an eye on developing early-warning mechanisms for the prediction of comments on posted images vulnerable to attacks. Given the overwhelming increase in media accompanying text in online social networks, we investigate use of posted images and captions for improved detection of bullying in response to shared content. We validate our approaches on a dataset of over 3000 images along with peer-generated comments posted on the Instagram photo-sharing network, running comprehensive experiments using a variety of classifiers and feature sets. In addition to standard image and text features, we leverage several novel features including topics determined from image captions and a pre-trained convolutional neural network on image pixels. We identify the importance of these advanced features in assisting detection of Cyberbullying in posted comments. We also provide results on classification of images and captions themselves as potential targets for cyber bullies.
Keywords: *Social Networks; Denigrates; Affiliation; Cyberbullying,datascraping,sentiment analysis,data visualization;*

Date of Submission: 03-06-2021 Date of acceptance: 17-06-2021

I. INTRODUCTION

1.1 OVERVIEW

A growing body of research into Cyberbullying in online social networks has been catalyzed by increasing prevalence and deepening consequences of this type of abuse. To date, automated detection of Cyberbullying has focused on analyses of text in which bullying is suspected to be present. However, given the increase in media accompanying text in online social networks, an increasing number of Cyberbullying incidents are linked with photos and media content, which are often used as targets for harassment and stalking.

For instance, in Instagram, a highly popular online photo sharing platform, bullying is becoming a serious concern. Recent statistics indicate that anywhere between 9% and 25% of users claim to have been bullied on Instagram, with the problem even more prevalent on Instagram and Facebook. Considering the pervasiveness and danger increasingly represented by bullying online, bully detection is of interest to a cross-sectional community of social and computer scientists. In particular, detecting instances of Cyberbullying through analysis of media content is an important and challenging task, as the connection between a bullied image and its context is unclear. Yet, insight into the characteristics of shared content may prove extremely useful in eventual development of warning mechanisms designed to prevent Cyberbullying.

In this work, we develop methods for detecting Cyberbullying in commentaries following shared images on Instagram. In addition to image-specific and text features extracted from comments and from image captions, we leverage several novel features including topics determined from image captions and outputs of a pre-trained convolutional neural network applied to image pixels. We identify the importance of these advanced features in detecting occurrences of Cyberbullying in posted comments. We also provide results on

classification of images and captions themselves as potential targets for cyber bullies. Leveraging features of the posted images and captions as well as the comments themselves, we are able to classify comments that contain bullying with accuracy. Moreover, we lay the foundation for identifying posted content which may be particularly vulnerable to bullying, noting the difficulty of the problem space and suggesting pointers for next steps.

1.2 BACKGROUND

Cyberbullying is a growing problem affecting more than half of all American teens. The main goal of this paper is to investigate fundamentally new approaches to understand and automatically detect incidents of Cyberbullying over images in Instagram, a media-based mobile social network. To this end, we have collected a sample Instagram data set consisting of images and their associated comments, and designed a labeling study for Cyberbullying as well as image content using human labelers at the crowd-sourced Crowd flower Web site. An analysis of the labeled data is then presented, including a study of correlations between different features and Cyberbullying as well as cyber aggression. Using the labeled data, we further design and evaluate the accuracy of a classifier to automatically detect incidents of cyber-bullying.

1.3 PROBLEM STATEMENT

With the advancement of technology, craze of social networking platforms is proliferating. Online users now share their information with each other easily using computers, Mobile phones. However, this has led to the growth of cyber-criminal acts for example, cyber bullying which has become a worldwide epidemic. Cyber bullying is the use of electronic communication to bully a person by sending harmful messages using social media, instant messaging or through digital messages. It has emerged out as a platform for Insulting,

Humiliating a person who can affect the person either physically or emotionally and sometimes leading to suicidal attempts in the worst case. The main issue in preventing cyber bullying is detecting its occurrence so that an appropriate action can be taken at initial stages. To overcome this problem, many methods and techniques had been worked upon till now to control this problem. This paper is a survey covering cyber bullying and cyber bullying detection techniques. Through machine learning, we can detect language patterns used by bullies and their victims, and develop rules to automatically detect cyber bullying content.

Cyber bullying is the use of technology as a medium to bully someone. Although it has been an issue for many years, the recognition of its impact on young people has recently increased. Social networking sites provide a fertile medium for bullies, and teens and young adults who use these sites are vulnerable to attacks. Cyber bullying is when an intended person use the internet, cell phones, or other technological devices to send or post text or images intended to hurt, embarrass, threaten, torment, humiliate, or intimidate their victim. Cyber bullying is when a intended person use the internet, cell phones, or other technological devices to send or post text or images intended to hurt, embarrass, threaten, torment, humiliate, or intimidate their victim.

Cyber-aggression is defined as intentional harm carried out through electronic means to an individual or a group of individuals of any age, who perceive such acts as offensive, derogatory, harmful or unwanted. It uses of technology as a medium to bully someone. Although it has been an issue for many years, the recognition of its impact on young people has recently increased. Social networking sites provide a fertile medium for bullies, and teens and young adults who use these sites are vulnerable to attacks. Through machine learning, we can detect language patterns used by bullies and their victims, and develop rules to automatically detect cyber bullying content.

1.4 OBJECTIVE & PUROSE

1. Real time data

Accessing a comment form the users feed through the data scraping & performing the machine learning algorithm

2. Sentiment analysis

Sentiment Analysis refers to the automated techniques which extract the opinions from a specific piece of text written in natural language.

3. Classification of sentiment

In other words, sentiment analysis finds out whether the particular piece of text is positive, negative, or neutral.

1.5 SCOPE

Our project is mainly for detection of cyber bullying on Instagram system regarding it. Implementation of this project the person to who is facing the cyber bullying or who wants to analyse the post comments .

The main scope of project is to analyse the real time data taken from the user post and analysis the data into excel .csv file and further this file will be analysed using the textblob sentiment analysis.

1.6 SPECIFICATION

Selenium :-

Automation testing through Selenium and Python.

Selenium allows you to define tests and automatically detect the results of these tests on a pre-decided browser. A suite of Selenium functions enables you to create step-by-step interactions with a webpage and assess the response of a browser to various changes. You can then decide if the response of the browser is in line with what you expect.

This post assumes that you do not have prior knowledge of Selenium. However, basic knowledge of front-end concepts like DOM and familiarity with Python is expected.

ChromeDriver :-

WebDriver is an open source tool for automated testing of webapps across many browsers. It provides capabilities for navigating to web pages, user input, JavaScript execution, and more. ChromeDriver is a standalone server that implements the W3C WebDriver standard. ChromeDriver is available for Chrome on Android and Chrome on Desktop

TextBlob :-

TextBlob aims to provide access to common text-processing operations through a familiar interface. You can treat TextBlob objects as if they were Python strings that learned how to do Natural Language Processing.

The sentiment property returns a namedtuple of theform Sentiment(polarity, subjectivity). The polarity score is a float within the range [-1.0, 1.0]. The subjectivity is a float within the range [0.0, 1.0] where 0.0 is very objective and 1.0 is very subjective.

Specification of project

Import data form Instagram

The project helps to collect the dataset from Instagram with the help of selenium and webdriver and save the data into excel file

Commests.csv file

This file contains the dataset of user post contain field like user name and comment

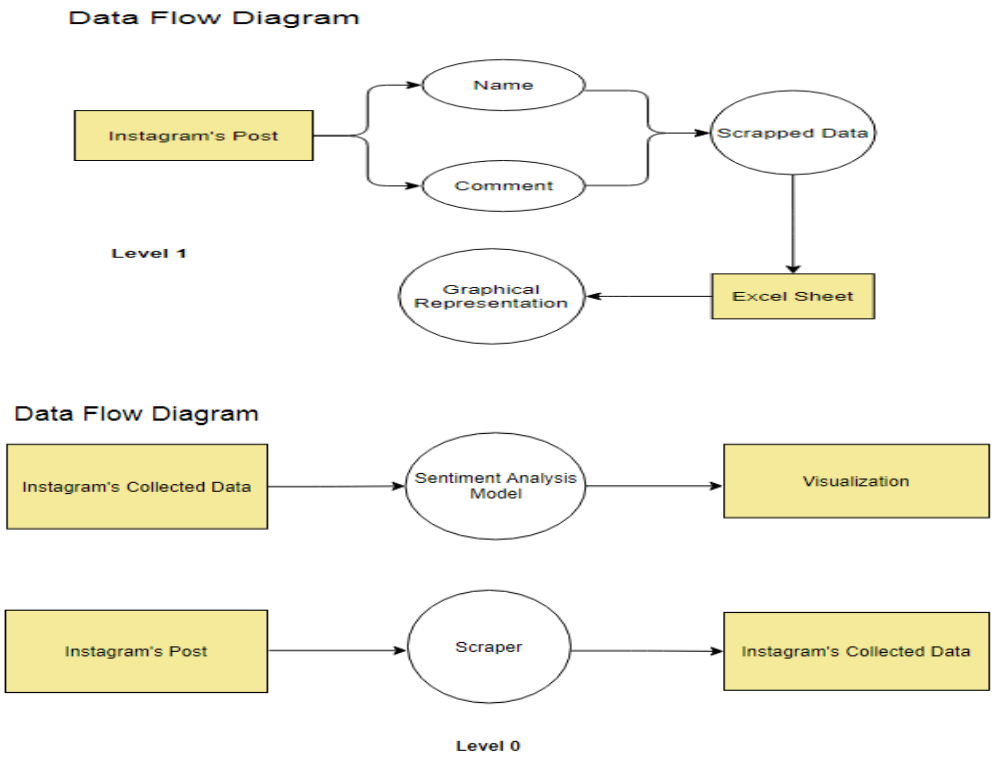
Analysis

After the collection of analysis the dataset pre-processing is done and later it is been classified into various category such as positive negative and neutral

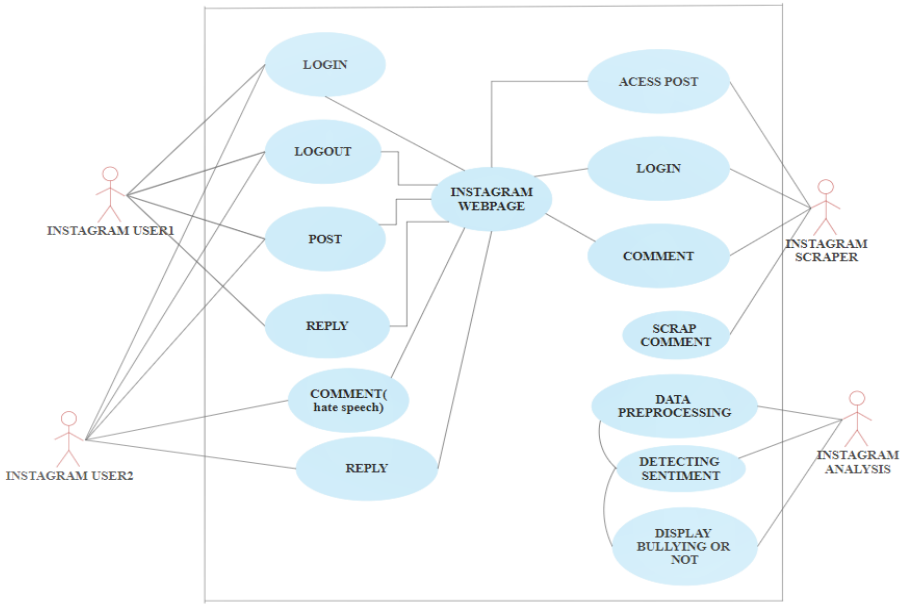
Graphical representations

The classified data is further fed to the data visualization section where the data is visualized

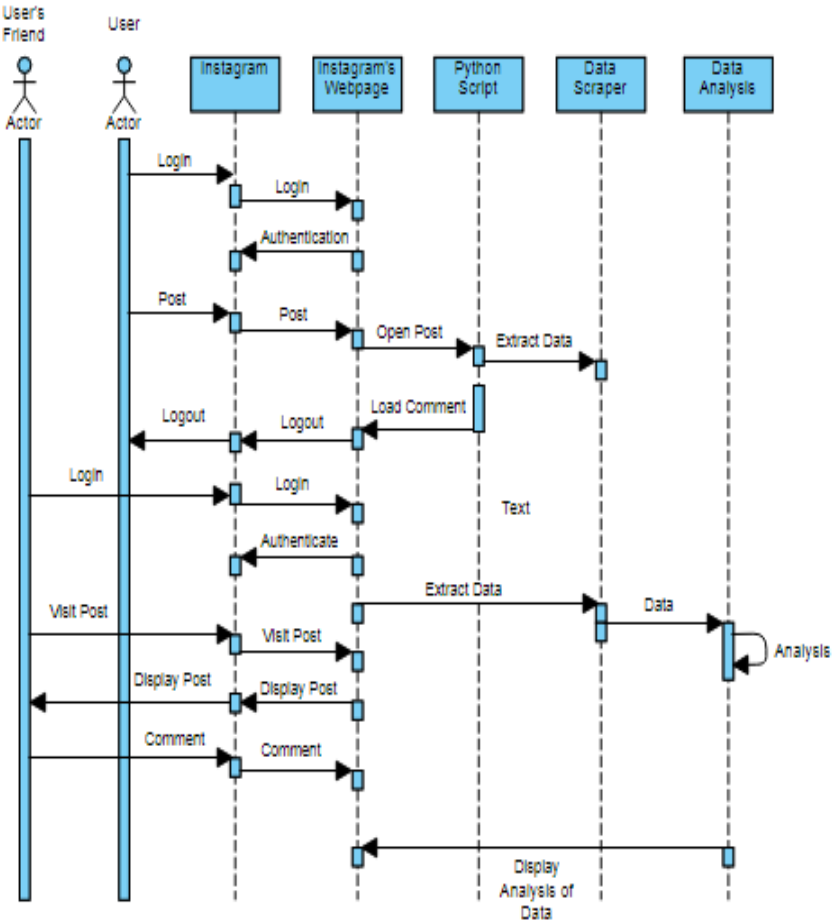
DFD



USECASE DIAGRAM



SEQUENCE DIAGRAM



Sequence Diagram

II. RESULTS AND DISCUSSION

DATA COLLECTION

Data is stored in a excel sheet the field contains two column first column as name which contain the dataset of the person name or userid who has commented on a particular post second field contains the dataset of the comment on that post As Shown Below:

[4]	145	145	real_praveen	Don't prove our epics with these human's word ...
	146	146	shivammishra6033	Well ending is beginning somehow it's a nature...
	147	147	_kirti_rao_official	Why did brahma create a woman and had lust for...
	148	148	reys_awareness	True
	149	149	memersoumik0	Jai Shree Krishna IN
	150	150	arnonee22	Jay Shri RadhaKrishna
	151	151	007virajmalik	Geeta English me likhi he kya
	152	152	arpittiwari1998	Jai shri ram
	157	157	just_smile_10_10	
	158	158	happymind44	Jitne bhi parsl sbko mere pas bhejdo sbko fres...
	159	159	ehsanullahnaim	Missing him
	160	160	dz_chinu_bad_girl_dz	
	161	161	normeenormason	"I keep myself busy with the things that I do,...
	162	162	sampathkumar__js	@sushantsinghrajput bahi
	164	164	ravi_beniwal_143	Miss you
	165	165	itz_sakshi.joshi01	susant sir
	166	166	official_atul.yadav	miss uu sir
	167	167	happymind44	Kal Tere Bhai choda muje maze leke
	168	168	happymind44	Puri night
	170	170	__rakesh_marathe__	
	171	171	bshsvshhshsoso	Hi friends, Please don't ignore One urgent hel...
	173	173	bikram_r0y	Miss you Dada, rest in peace
	174	174	happymind44	Tere Bhai ki sari strongnes he mere andr
	175	175	suchita.b.modi	There is something in you that makes people st...
	176	176	happymind44	Muje choda he bhot maze se hmara Jake khatm hu...
	177	177	prayik2	bad to see this

Fig:- data collected from various posts

SCREENSHOTS

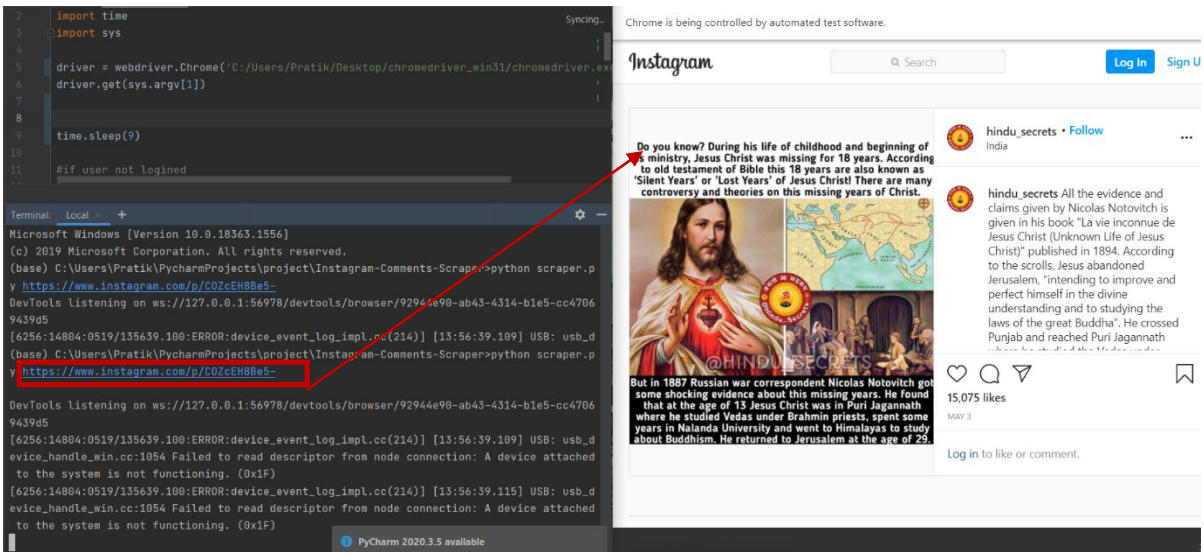


Fig: opening post in chromedriver

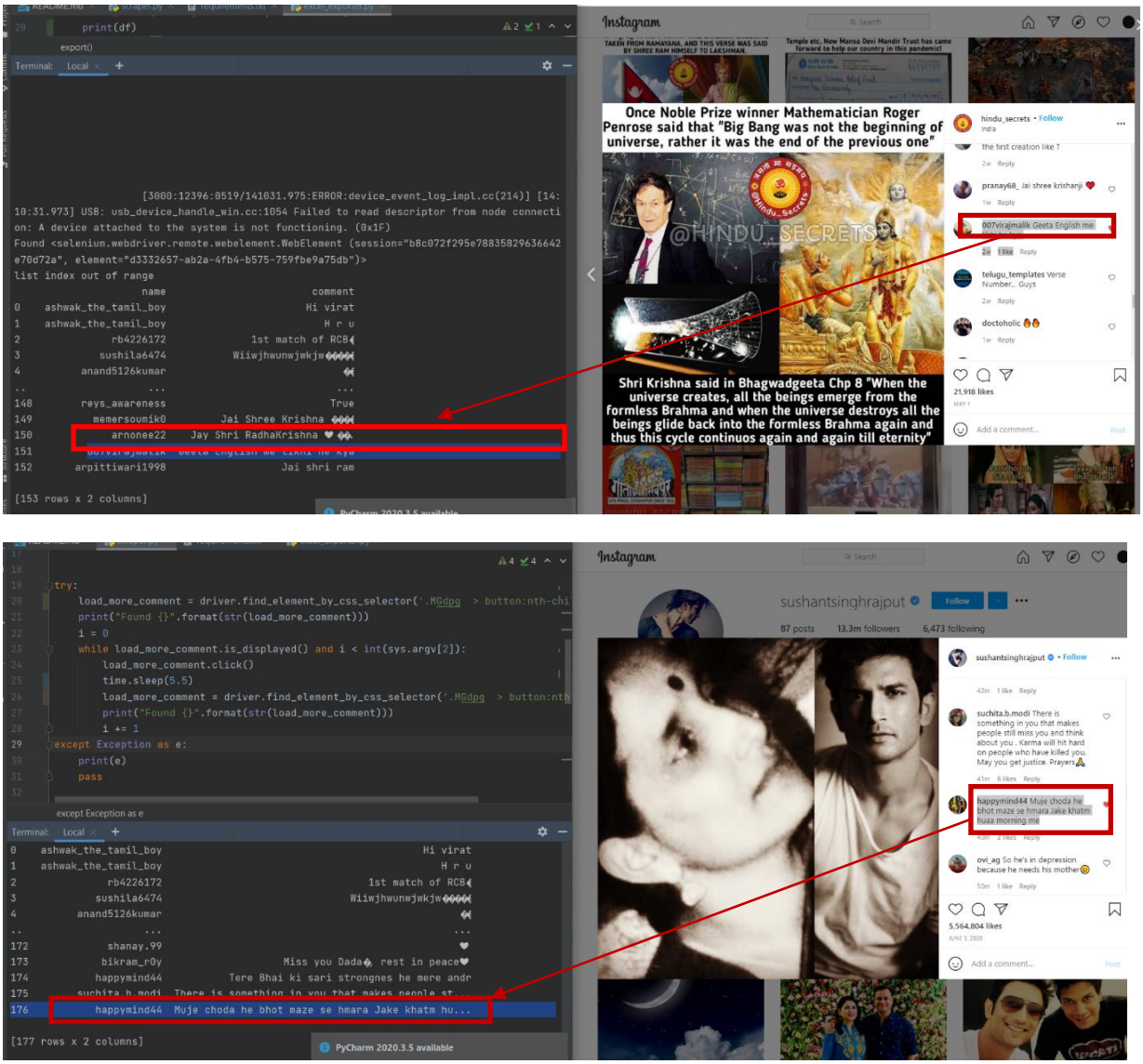


Fig: collecting data from post

[16]

	Unnamed: 0	sentence	sentiment	polarity	Sentiment_Type
0	0	NaN	0.000000	0.000000	Neutral
1	1	you are great	0.800000	0.750000	not_bullying
2	2	bhai	0.000000	0.000000	Neutral
3	3	hii	0.000000	0.000000	Neutral
4	4	amazing jo bhi mujhe follow kare ga main use l...	0.383333	0.616667	not_bullying
5	5	NaN	0.000000	0.000000	Neutral
6	6	my best team	1.000000	0.300000	not_bullying
7	7	virat kohli you are my best player my favorite...	0.750000	0.650000	not_bullying
8	8	mai jab bhi match dekhta hu ipl toh rcb ka jar...	0.000000	0.000000	Neutral
9	9	NaN	0.000000	0.000000	Neutral
10	10	NaN	0.000000	0.000000	Neutral
11	11	maths ke liye nobel prize nhi hota hai	0.000000	0.000000	Neutral
12	12	ek universe keh khtam honeh keh baadh bhrahma ...	0.000000	0.100000	Neutral
13	13	give exact verse where sri krsna says that in ...	0.250000	0.250000	not_bullying
14	14	it s brahman not brahma.. lord brahma is one o...	0.000000	0.600000	Neutral
15	15	he who meditates on me as th.supreme personali...	0.250000	0.611111	not_bullying
16	16	yea.. it s true.. that line is more bigger tha...	0.283333	0.550000	not_bullying
17	17	hinduism is a very complex philosophy and only...	-0.022500	0.680000	bullying
18	18	jai shree krishna	0.000000	0.000000	Neutral
19	19	but what was the first creation like	0.250000	0.333333	not_bullying
20	20	if they said it s proven fact if we say that i...	0.000000	0.000000	Neutral
21	21	nice caption though	0.600000	1.000000	not_bullying
22	22	why tho like why the continues cycle anyone kn...	0.000000	0.000000	Neutral
23	23	sounds like when every matter existing will be...	-0.086667	0.333333	bullying
24	24	NaN	0.000000	0.000000	Neutral

Fig: classification into sentence type

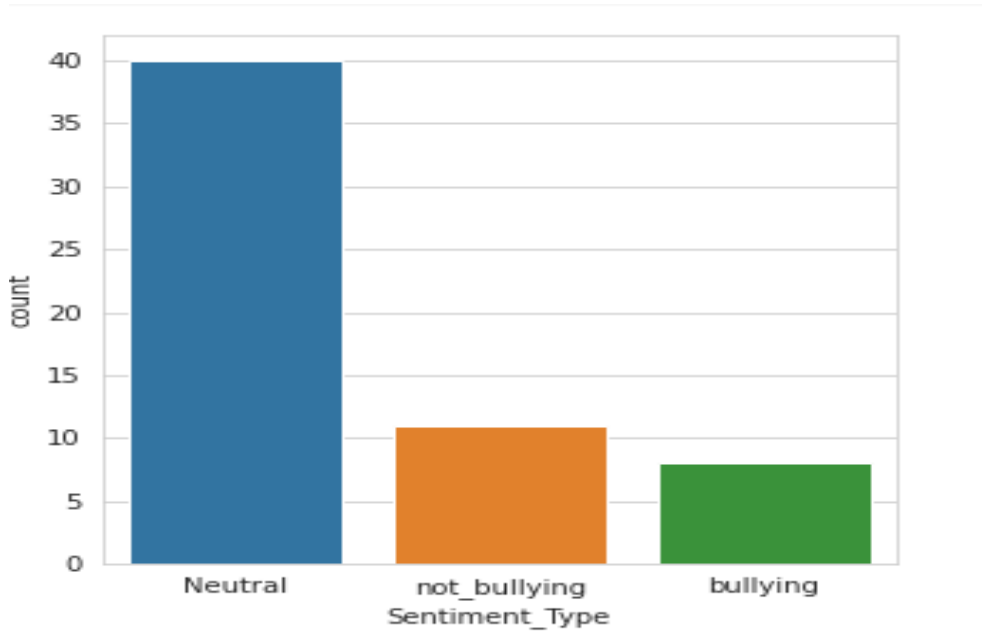


Fig: Graphical representation of data

III. CONCLUSION

Cyberbullying is a serious issue, and likely any form of bullying it can have long term effects on its victims. Our project will grow and help individuals to be aware of Cyber bullies. Parents, teachers and children must work together to prevent Cyberbullying and to make Internet a safer place for all.

FUTURE SCOPE

One theme for future work is to improve the performance of our classifier by adding more input features, such as new image features, temporal behavior of commenting, mobile sensor data, etc. A limitation of our current classifier is that it is designed only for highly negative media sessions. A more general classifier that can apply to all media sessions is needed. This will also require us to enlarge our labeled data set substantially.

Incorporating image features needs to be automated by applying image recognition algorithms. We plan to explore this research direction as well. We have applied a majority vote definition in designing our classifier. Another definition to consider is when at least one labeller has declared that he/she thinks this media session constitutes Cyberbullying. New classifiers will have to be designed for this definition. We also plan to consider designing classifiers for cyber aggression in addition to Cyberbullying, and to investigate those media sessions that represent the former but not the latter behavior. Another theme for future work is to obtain greater detail from the labeling surveys. Our experience was that streamlining the survey improved the response rate, quality and speed. However, we desire more detailed labeling, such as for different roles in Cyberbullying – identifying and differentiating the role of a victim's defender, who may also spew negativity, from a victim's bully or bullies.

REFERENCES

[1]. Peter K Smith, Jess Mahdavi, Manuel Carvalho, Sonja Fisher, Shanette Russell, And Neil Tippett. Cyberbullying: Its Nature And Impact In Secondary School Pupils. Journal Of Child Psychology And Psychiatry, 49(4):376–385, 2008.

[2]. Kellyreynolds, Aprilkontostathis, And lyn need wards . Using machine Learning To Detect Cyberbullying. In 2011 10th International Conference On Machine Learning And Applications And Workshops, Volume 2, Pages 241–244. Ieee, 2011.

[3]. B Nandhini And Ji Sheeba. Cyberbullying Detection And Classification Using Information Retrieval Algorithm. In Proceedings Of The 2015 International Conference On Advanced Research In Computer Science Engineering & Technology (Icarc set 2015), Page 20. Acm, 2015.

[4]. B Sri Nandhini And Ji Sheeba. Online Social Network Bullying Detection Using Intelligence Techniques. Procedia Computer Science, 45:485–492, 2015.

[5]. Walisa Romsaiyud, Kodchakorn Na Nakornphanom, Pimpaka Prasertsilp, Piyaporn Nurarak, And Pirom Konglerd. Automated Cyberbullying Detection Using Clustering Appearance Patterns. In Knowledge And Smart Technology (Kst), 2017 9th International Conference On, Pages 242– 247. Ieee, 2017.

[6]. Shane Murnion, William J Buchanan, Adrian Smales, And Gordon Russell. Machine Learning And Semantic Analysis Of In-Game Chat For Cyberbullying. Computers & Security, 76:197–213, 2018.

[7]. Sani Muhamad Isa, Livia Ashianti, Et Al. Cyberbullying Classification Usingtextmining. Ininformatics and computational sciences (Icicos), 2017 1st International Conference On, Pages 241–246. Ieee, 2017.

[8]. Karthikdinakar,Biragojones,Catherinehavasi,Henrylieberman,And Rosalind Picard. Common Sense Reasoning For Detection, Prevention, And Mitigation Of Cyberbullying. Acm Transactions On Interactive Intelligent Systems (Tiis), 2(3):18, 2012.