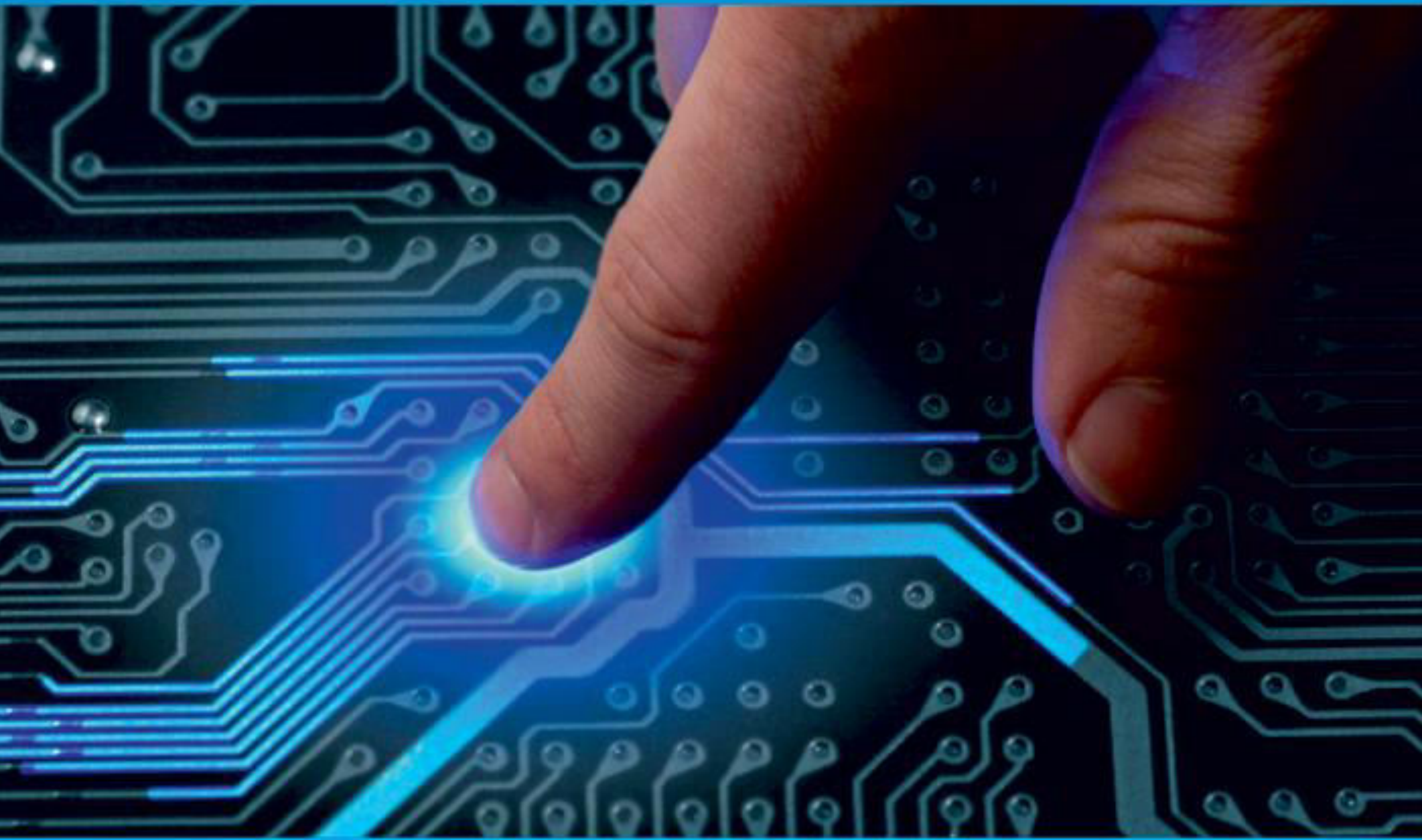
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005320 |

Detection of Cyberbullying tweets using Machine Learning

G. Ramadevi, S.Sai Charan, M.Pedda Bharath, M.Siva Punna Rao, Sk.Mehaboob

Asst. Professor, Dept. of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, A.P., India

Student, Dept. of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, A.P., India

Student, Dept. of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, A.P., India

Student, Dept. of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, A.P., India

Student, Dept. of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, A.P., India

ABSTRACT: 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. Both Naive Bayes and SVM (Support Vector Machine) were able to detect the true positives with 71.25% and 52.70% accuracy respectively. But SVM outperforms Naive Bayes of similar work on the same dataset. Also, Twitter API is used to fetch tweets, and tweets are passed to the model to detect whether the tweets are bullying or not.

KEYWORDS: Cyberbullying, Machine Learning, Classifiers, Naïve Bayes, Support Vector Machine, Twitter API.

I. INTRODUCTION

To date, Internet users from everywhere over the globe utilize and access forms of social network services and social media as a fundamental to their personal networking, transferring, and sharing of information within the communities.
The sites that provide “Online Social Networking” services assist users in forming a sway or perception, in maintaining and acquiring new relationships within the SNS. we will deduce that the web community may be a place for the SNS users to satisfy people allow users to satisfy people on the other sides of the globe in cyberspace, allowing users to explain their social networks clearly and maintain a connection and network with others, in the virtual world.
Social networks like microblogging could be a mass medium that exists within the kind of blogging. Twitter may be a microblogging service that evolved as a disruptive platform that's meant for the users to transmit their daily activities, feelings, and opinion by posting simple tweets (messages) within their friend’s circle. The topics range from the standard of living to current activities, personal opinions, experience sharing and other interests. The social networks, i.e. Twitter has was something that's indivisible to those 645,750,000 active Twitter users (Statistic Brain, 2014) and becoming a part of their life, where everyone can share the knowledge, as anyone with large amount of users can't live without Twitter.
With the recent popularity of Twitter, it's important to understand why and the way people use these tool, as Twitter can sometimes wont to abuse by irresponsible users to cyberbully and
post something bad and harm individual’s personally. The importance of those social network services has caused an increase in cyberbullying circumstances, particularly among the teenagers . Hence, it's important to spot the cyberbullying event and therefore the attacking messages in social media. Though cyberbullying won't cause any physical damage initially, however, it likely caused destructive psychological effects, like low self-esteem, mental depression, suicide consideration and even suicide (S. Hinduja and J. W. Patchin, 2010). A fatal cyberbullying incident had happened on MySpace SNS (Tavani, Herman. T. , 2013), whereby Megan Meier, a 13-year-old teen became increasingly distressed by the net harassment being directed at her, and eventually decided to finish her life by hanging herself in her bedroom in 2006. Hence, recognizing the cyberbullying event itself isn't efficient in combating cyberbullying as such, as we want to identify the 000 user of the cyberbully so as to arrest them for justice, and to forestall further similar cases to happen. It’s a motivation to make a web-based application, i.e. the Cyberbullying Detection System on Twitter, to effectively discover the cyberbullying related tweets from Twitter and providing

reasonable solution thereafter. With this technique, the users can identify the cyberbullying related tweets supported the keywords and populate it during a news feed form. By doing this, it allows users to see the identities of the cyberbullies and also the victims from the cyberbullying tweets. Also, it'll allow the users to come up with reports to higher authorities, i.e. police reports, based on case's severity and wishes.

In conclusion, with the arrival of this cyberbullying detection and solution system in Twitter, it will help the authorities to observe, regulate or a minimum of decrease the harassing incidents in cyberspace in Malaysia. With the implementation of this method, this can also help to lift the cyberbullying awareness among the Twitter users, and posting the tweets responsibly within the social media, as posting irritating tweets is prohibited and bullies are often convicted under the Computer Crimes Act, the legal code.

II. RELATED WORKS

The rise of social media platforms in recent years brought up huge information resources that involve new approaches to studying the respective data. Users interact with this system through the Web interface, mobile application, instant messaging(IM) agent, or sending SMS updates. There are several kinds of research being done to investigate the usage and the communities on Twitter. (Java, A., et al., 2007), investigate the motivation of research user's in adopting this specific microblogging platform, i.e. Twitter. As mentioned in this research, there are still shallow studies that have been done on this form of communication and information sharing, and hence, further studies on the topological and geographical structure of Twitter's social network have been carried out in this research in attempting to comprehend the user intentions and community structure in microblogging.

Cyberbullying can be defined as a type of harassment (or bullying) that takes place online, via e-mail, text messaging, or online forums, such as social networking sites. Social networks provide an ideal background for data gathering and information that might enable the criminals to execute their crime, for example, by determining one that is a vulnerable or 'suitable' victim. We categorized this kind of crime as cyber-related crime and we are expanding its definition to include cyberbullying as one of the serious offenses in the cyber realm as it has resulted in death(Tavani, Herman T., 2013).

A statistical report investigated by Cyber Security Malaysia in 2007 showed that 60 cases have been reported involving cyberbullying. Although the report illustrated some isolated cases, however, while the same survey also found that nine out of ten children in Malaysia have been exposed to negative experiences or elements from online use. According to the report by Cyber Security Malaysia, most cyber bullies and their victims have close contact including their close friends, ex-spouses, and former colleagues. Thus, the existing problem required serious attention and a solution. Cyberbullying is a serious sign and should be addressed by all parties and their concerns on the matter are necessary including parents, teachers, and the surrounding community at large.

Some previous research has discussed cyberbullying in social media. Writing style, structure, and specific cyberbullying content as features to predict the user's potential to send out offensive messages. The technique that has been used to identify offensive language is the Lexical Syntactic Feature (LSF) approach and it is successful in detecting some offensive content in social media, which has achieved a precision of 98.24%, and recall of 94.34%, and also succeeds in detecting users who sent offensive messages, achieving precession of 77.9%, and recall of 77.8%(Chen et al. 2012).

With the rapid and wide coverage of Twitter, events can be discovered in an instant manner by monitoring and observing the incoming tweets. The event detection system, Twitter-based Event Detection and Analysis System (TEDAS), (R.Li, K.H.Lei, R.Khadiwala, Chang,2012) employs an adapted information retrieval architecture that covers online processing and an offline processing part. The offline processing is based on a fetcher accessing Twitter's API and a classifier to mark tweets as event-related or not event-related. Not only that, this system can help in identifying and examining events by exploring rich information from Twitter. From this research, there are three main functions proposed, which are detecting new events, ranking events based on their priority, and generating spatial and temporal patterns for the events detected. The TEDAS system is mainly focusing on Crime and Disaster-related Events (CDE), for instance, car accidents.

III. METHODOLOGY

The implementation of the Cyberbullying Detection System on Twitter is based on PHP andHTML with MySQL and Twitter API. This system will detect cyberbullying-related tweetsthat have matching keywords from the database. The conceptual model of cyberbullying detection system is shown. The conceptual model of the system describes the complete processof how bullying-related tweets can be identified and alert the associations, and others or guardians in monitoring their cyberbullying activities. Initially, the user needs to login into thesystem. The system has been coded with the account from the website. The API connects to Twitter when the sign is perfectly done.

IV. PROJECT AND SCOPE

In this Cyberbullying Detection System on Twitter, we've got to clarify the scope so as to accomplish this project

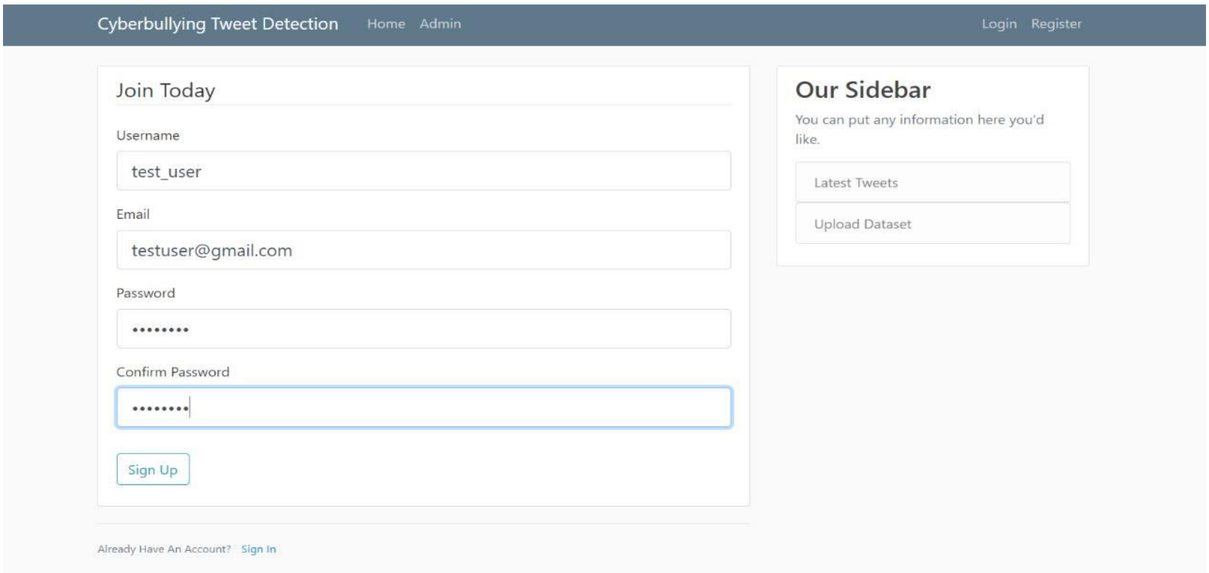
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005320 |

efficaciously. Since this is often a text-based cyberbullying tweets mining system, we've to obviously choose the kind of text or language that is going to be captured for this method to process. during this system, we only target the English language with proper and formal text. Thus, cyberbullying means won't be captured within the system.
Furthermore, since the system only specializes in the text posted by Twitter users. The thing we focus on during this study is the only text written by users. Besides that, supported that include 'gay', 'bitch', 'slag', 'homo', 'dike', and 'queer', may appear in our targeted tweets.
The social network platform that we are visiting further discuss and research are within the context of Twitter SNS. We are specializing in the tweets posted by the users on Twitter and capture the keywords written by the users for keyword matching so as to work out the cyberbullying event, the checking of the victims and cyberbullies, and their areas. The website is offering one of the best paraphrasing tools to rephrase sentences.

V. RESULTS



Join Today

Username: test_user

Email: testuser@gmail.com

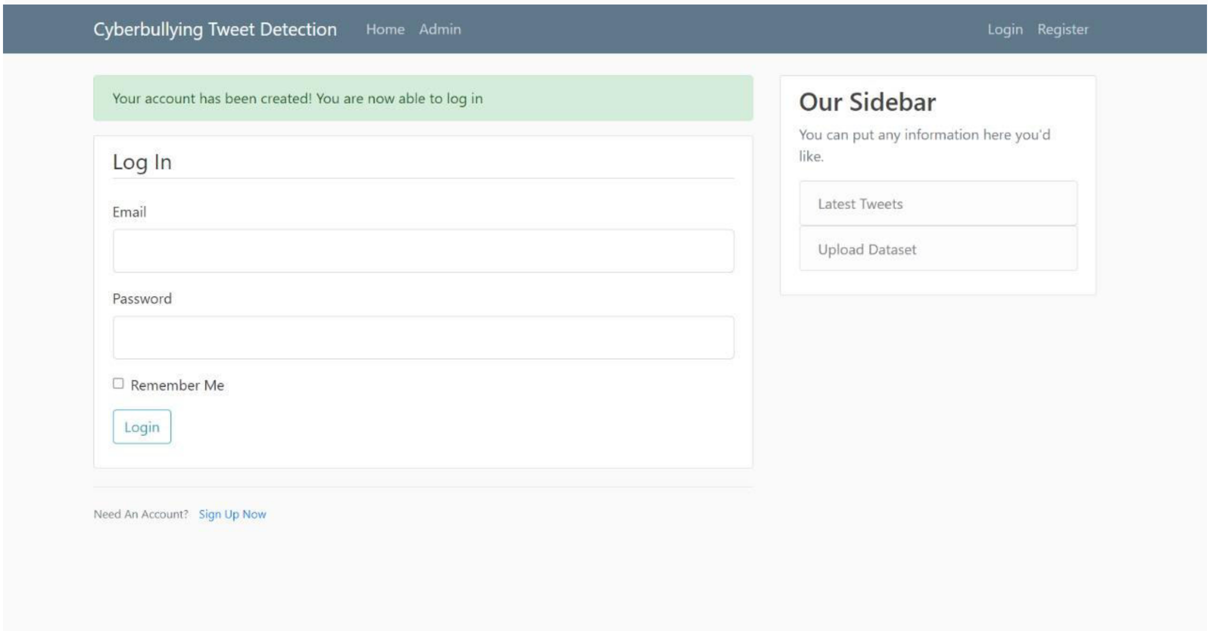
Password: ••••••••

Confirm Password: ••••••••

Already Have An Account? Sign In

Our Sidebar

You can put any information here you'd like.



Your account has been created! You are now able to log in

Log In

Email

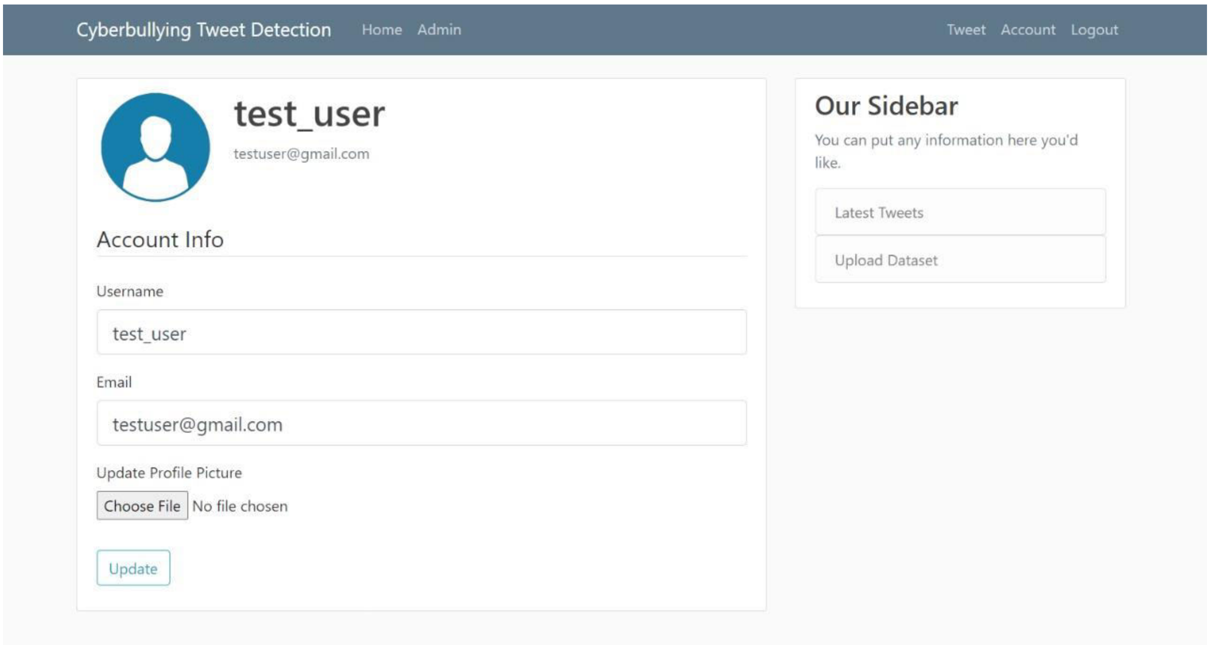
Password

☐ Remember Me

Need An Account? Sign Up Now

Our Sidebar

You can put any information here you'd like.



test_user

testuser@gmail.com

Account Info

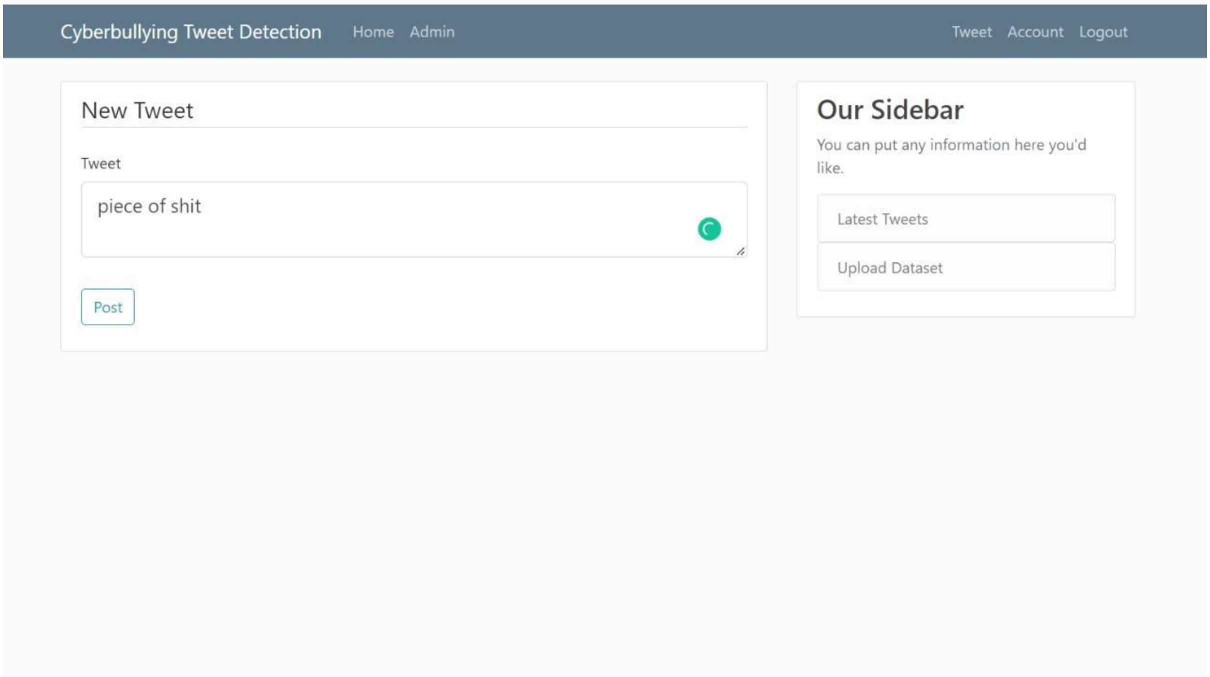
Username: test_user

Email: testuser@gmail.com

Update Profile Picture: No file chosen

Our Sidebar

You can put any information here you'd like.

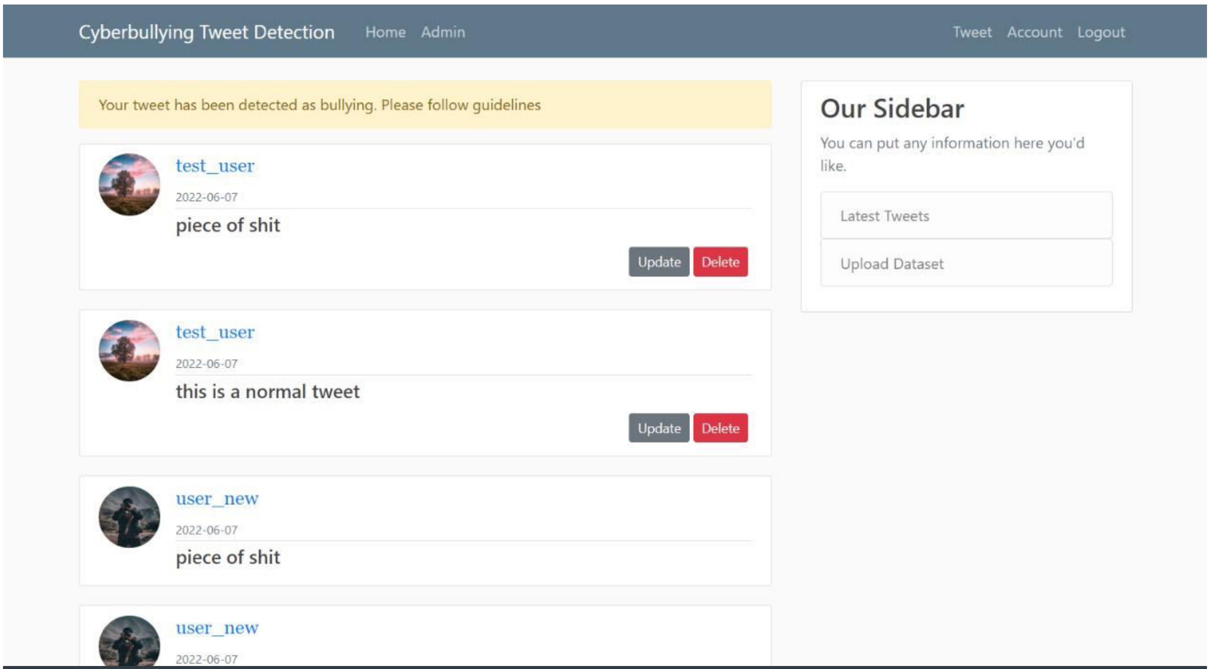


New Tweet

Tweet: piece of shit

Our Sidebar

You can put any information here you'd like.



Your tweet has been detected as bullying. Please follow guidelines

test_user
2022-06-07
piece of shit

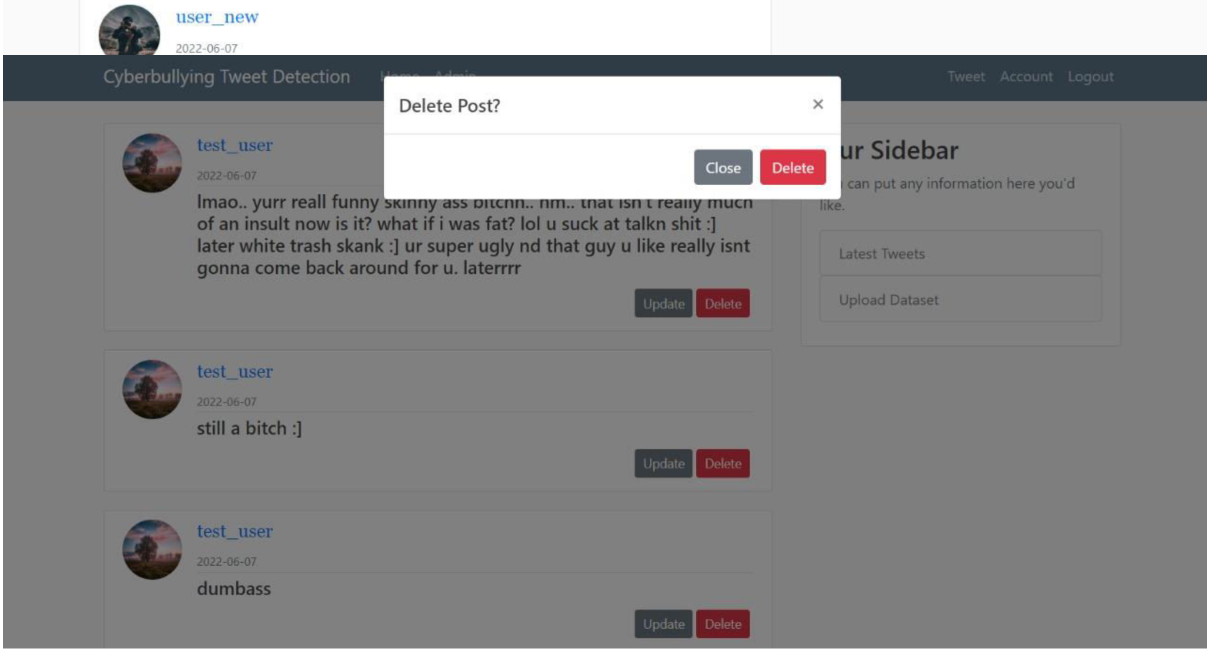
test_user
2022-06-07
this is a normal tweet

user_new
2022-06-07
piece of shit

user_new
2022-06-07

Our Sidebar

You can put any information here you'd like.



Delete Post?

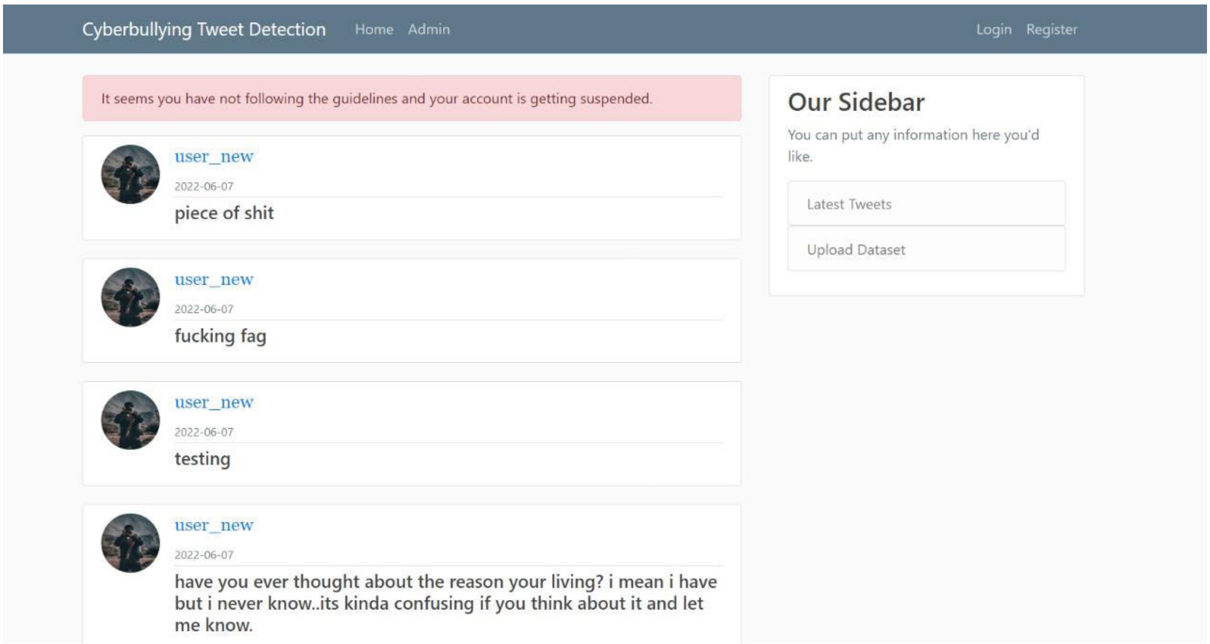
test_user
2022-06-07
lmao.. yurr reall funny skinny ass bitchh.. hm.. that isn t really much of an insult now is it? what if i was fat? lol u suck at talkn shit :] later white trash skank :] ur super ugly nd that guy u like really isnt gonna come back around for u. laterrrr

test_user
2022-06-07
still a bitch :]

test_user
2022-06-07
dumbass

ur Sidebar

can put any information here you'd like.



It seems you have not following the guidelines and your account is getting suspended.

user_new
2022-06-07
piece of shit

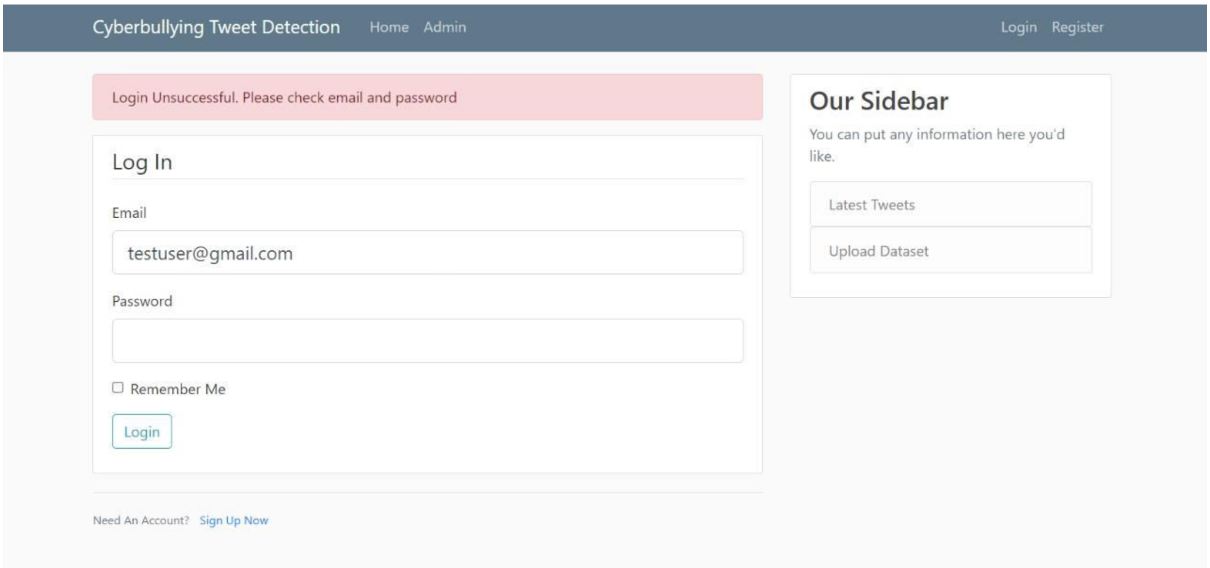
user_new
2022-06-07
fucking fag

user_new
2022-06-07
testing

user_new
2022-06-07
have you ever thought about the reason your living? i mean i have but i never know..its kinda confusing if you think about it and let me know.

Our Sidebar

You can put any information here you'd like.



Login Unsuccessful. Please check email and password

Log In

Email: testuser@gmail.com

Password:

□ Remember Me

Need An Account? Sign Up Now

Our Sidebar

You can put any information here you'd like.

VI. SOFTWARE TESTING

Software testing is the method of verification and evaluation that a software outcome or application does what it is supposed to do. The benefits of testing include preventing bugs, reducing development costs, and improving performance. Evaluating the software against various test case scenarios is called testing the software.

Software Testing is a method to check whether the actual software product matches all needs and is less prone to errors. It involves the execution of software/system components using manual or automated tools to evaluate one or more properties of interest. The goal of testing is to the identification of errors, spaces, or missing constraints in contrast to actual requirements.

Software testing can be stated as the process of checking and testing whether software or application has fewer defects, updated to the technical requirements as needed by its design and development and the user requirements are met effectively and efficiently by handling all the exceptional and boundary cases.

It mainly aims at measuring the specification, functionality, and performance of a software program or application.

VII. CONCLUSIONS

An approach is proposed for detecting and preventing Twitter cyberbullying using SupervisedBinary classification Machine Learning algorithms. Our model is evaluated on both Support Vector Machine and Naive Bayes, also for feature extraction, used the TFIDF vectorizer. As the results show us that the accuracy for detecting cyberbullying content has also been great for Support Vector Machine at around 84.05% which is better than Naive Bayes. Our model will help people from the attacks of social media bullies.

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 10t h 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 detect ion using machine learning on Bangla text ", 10th Internat ional Conference on Electrical and Computer Engineering, pages 385-388, IEEE Xplore, 2018

[5]Nekt ariaPot ha and ManolisMaragoudakis. " Cyberbullying detectionusing time series modeling", In 2014IEEE 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

[7] Vikas S Chavan, SS Shylaja. "Machine learning approach for detect ion of cyber-aggressive comments by peers on social media network". In Advances in computing, communications, and informatics (ICACCI), 2015 International Conference on, pages 2354–2358. IEEE,2015

[8] Walisa Romsaiyud, Kodchakorn na Nakornphanom, Pimpaka Prasertsilp, P iyaporn Nurarak, and Pirom Konglerd, "Automated cyberbullying detect ion using clustering appearance patterns", In Knowledge and Smart Technology (KST), 2017 9th International Conference on, pages 242–247. IEEE, 2017.

[9] https://muthu.co/understanding-the-classification-report-in-sklearn/

[10] https://developer.twit ter.com/en/apps

[11] https://text-processing.com/demo/tokenize/

[12].https://towardsdatascience.com/support-vector-machine-introduction-tomachine-learning-algorithms-934a444fca47

[13] https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

doi crossref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 ☎ 6381 907 438 ✉ ijircce@gmail.com

Scan to save the contact details

www.ijircce.com