Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, August : 2023

Machine Learning Detection of Cyberbullying on Social Media

D. Bujjibabu 1,Gogineni Vanaja 2

#1Professor in the department of MCA at QIS College of Engineering and Technology (Autonomous), Vengamukkapalem, Prakasam (DT)
#2MCA Student in the department of MCA at QIS College of Engineering and Technology (Autonomous), Vengamukkapalem, Prakasam (DT)

ABSTRACT_ Cyberbullying is a significant issue on the internet that affects both adults and teenagers. Mistakes like despair and suicide have resulted from it.A increasing demand exists for the regulation of material on social media platforms. The work that follows builds a model based on the detection of cyberbullying in text data using natural language processing and machine learning utilising data from two different types of cyberbullying, hate speech tweets from Twittter and comments based on personal assaults from Wikipedia forums. To determine the most effective method, three feature extraction techniques and four classifiers are examined. The model provides accuracy levels above 90% for data from Tweets, and accuracy levels above 80% for data from Wikipedia.

1.INTRODUCTIOHN

Technology has become an essential part of our lives more than ever before. as the internet has developed. These days, social media is very popular. However, as with everything else, misusers will emerge occasionally late or early, but there will absolutely be one. Nowadays, cyberbullying is common. Social networking websites are useful for interpersonal communication. Despite the fact that social networking has become more common over time, most people use it in unethical and immoral ways to spread negativity. This occurs frequently between adolescents and young adults. Bullying each other online is one of their negative behaviors. In the online environment, it is difficult to determine whether someone is speaking for fun or with ulterior motives. They will often laugh it off with a simple joke like "or don't take it so seriously." The use of technology to bully, threaten, shame, or harm another person is known as cyberbullying. This fight over the internet frequently leads to threats in real life for one person. Suicide has been attempted by some people. At the outset, such activities must be stopped. If an individual's tweet or post is found to be offensive, for instance, his or her account may be terminated or suspended for a specific period of time. What exactly is cyberbullying then?

Cyberbullying is provocation, compromising, humiliating or focusing on somebody to have a good time or even by very much arranged implies

Explores on Cyberbullying Episodes show that 11.4% of 720 youthful people groups reviewed in the NCT DELHI were survivors of cyberbullying in a 2018 study by Kid Right and You, a NGO in India, and close to half of them didn't specify it to their educators, guardians or gatekeepers. 22.8% matured 13-18 who involved the web for around 3 hours daily were helpless against Cyberbullying while 28% of individuals who use web over 4 hours daily were casualties. There are a lot of other reports that tell us that cyberbullying has a big impact on people of all ages, and children between the ages of 13 and 20 face a lot of challenges in terms of their health, mental fitness, and ability to make decisions in any situation. According to the researchers, this issue should be taken seriously by every nation and resolved. Numerous child suicides in Russia and other nations occurred in 2016 as a result of the Blue Whale Challenge incident. It was a game that was played on a variety of social networks and involved a relationship between a participant and an administrator. Participants are given certain tasks for fifty days. At first, they are simple, like getting up at 4:30 in the morning or watching a horror movie. However, over time, they progressed to self-harm, which led to suicides. Later, it was discovered that the administrators were between the ages of 12 and 14..

2.LITERATURE SURVEY

2.1) Representation Learning: A Review and New Perspectives

AUTHORS: Y. Bengio, A. Courville, and P. Vincent

The success of machine learning algorithms generally depends on data representation, and we hypothesize that this is because different representations can entangle and hide more or less the different explanatory factors of variation behind the data. Although specific domain knowledge can be used to help design representations, learning with generic priors can also be used, and the quest for AI is motivating the design of more powerful representation-learning algorithms implementing such priors. This paper reviews recent work in the area of unsupervised feature learning and deep learning, covering advances in probabilistic models, auto-encoders, manifold learning, and deep networks. This motivates longer-term unanswered questions about the appropriate objectives

for learning good representations, for computing representations (i.e., inference), and the geometrical connections between representation learning, density estimation and manifold learning.

2.2) Users of the world, unite! The challenges and opportunities of Social Media

AUTHORS: A. M. Kaplan and M. Haenlein

The concept of social media is top of the agenda for many business executives today. Decision makers, as well as consultants, try to identify ways in which firms can make profitable use of applications such as Wikipedia, YouTube, Facebook, Second Life, and Twitter. Yet despite this interest, there seems to be very limited understanding of what the term ‘‘social media’’ exactly means; this article intends to provide some clarification. We begin by describing the concept of social media, and discuss how it differs from related concepts such as Web 2.0 and User Generated Content. Based on this definition, we then provide a classification of social media which groups applications currently subsumed under the generalized term into more specific categories by characteristic: collaborative projects, blogs, content communities, social networking sites, virtual game worlds, and virtual social worlds. Finally, we present 10 pieces of advice for companies which decide to utilize social media.

2.3) Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth

AUTHORS: R. M. Kowalski, G. W. Giumetti, A. N. Schroeder, and M. R. Lattanner

Although the Internet has transformed the way our world operates, it has also served as a venue for cyberbullying, a serious form of misbehaviour among youth. With many of today's youth experiencing acts of cyberbullying, a growing body of literature has begun to document the prevalence, predictors, and outcomes of this behaviour, but the literature is highly fragmented and lacks theoretical focus. Therefore, our purpose in the present article is to provide a critical review of the existing cyberbullying research. The general aggression model is proposed as a useful theoretical framework from which to understand this phenomenon. Additionally, results from a meta-analytic review are presented to highlight the size of the relationships between cyberbullying and traditional bullying, as well as relationships

between cyberbullying and other meaningful behavioural and psychological variables. Mixed effects meta-analysis results indicate that among the strongest associations with cyberbullying perpetration were normative beliefs about aggression and moral disengagement, and the strongest associations with cyberbullying victimization were stress and suicidal ideation. Several methodological and sample characteristics served as moderators of these relationships. Limitations of the meta-analysis include issues dealing with causality or directionality of these associations as well as generalizability for those meta-analytic estimates that are based on smaller sets of studies ($k < 5$). Finally, the present results uncover important areas for future research. We provide a relevant agenda, including the need for understanding the incremental impact of cyberbullying (over and above traditional bullying) on key behavioural and psychological outcomes

3.PROPOSED WORK

The suggested method for creating cyberbullying prediction models involves using a textual content classification approach that comprises creating computer learning classifiers from labelled textual content examples. A lexicon-based model that incorporates computing orientation for a record from the semantic orientation of phrases or phrases in the document is another option. In lexicon-based models, the lexicon is typically built manually or automatically by using seed phrases to expand the collection of words. However, literature rarely employs the lexicon-based strategy for cyberbullying prediction.

The main reason for this is the lack of structure in the texts on SM websites, which makes it difficult for the lexicon-based approach to fully understand cyberbullying. However, features that are widely employed as inputs to computer learning algorithms are extracted from lexicons. For instance, lexicon-based approaches are used as profane points in computer learning models, such as the usage of a profanity-based dictionary to track the variety of profane phrases in a message. The secret to accurate cyberbullying prediction is to have a set of extracted and manufactured elements.

3.1 IMPLEMENTAION

1) Dataset Upload & Analysis: using this module we will upload dataset and then perform analysis methods such as finding various cybercrime and its count and then clean dataset by removing missing values

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, August : 2023

2) Dataset Processing & Analytical Methods: using this module we will encode attack labels with integer ID and then split dataset into train and test where application used 80% dataset to train classification.

3) Run ML Model: using this module we will train classification algorithm with above 80% dataset and then build a prediction model

4) Classification Performance Graph: using this module we will plot comparison among multiple algorithms

5) Predict Output: using this module we will upload test dataset and then classification model will predict output based on input data

4.RESULTS AND DISCUSSIO

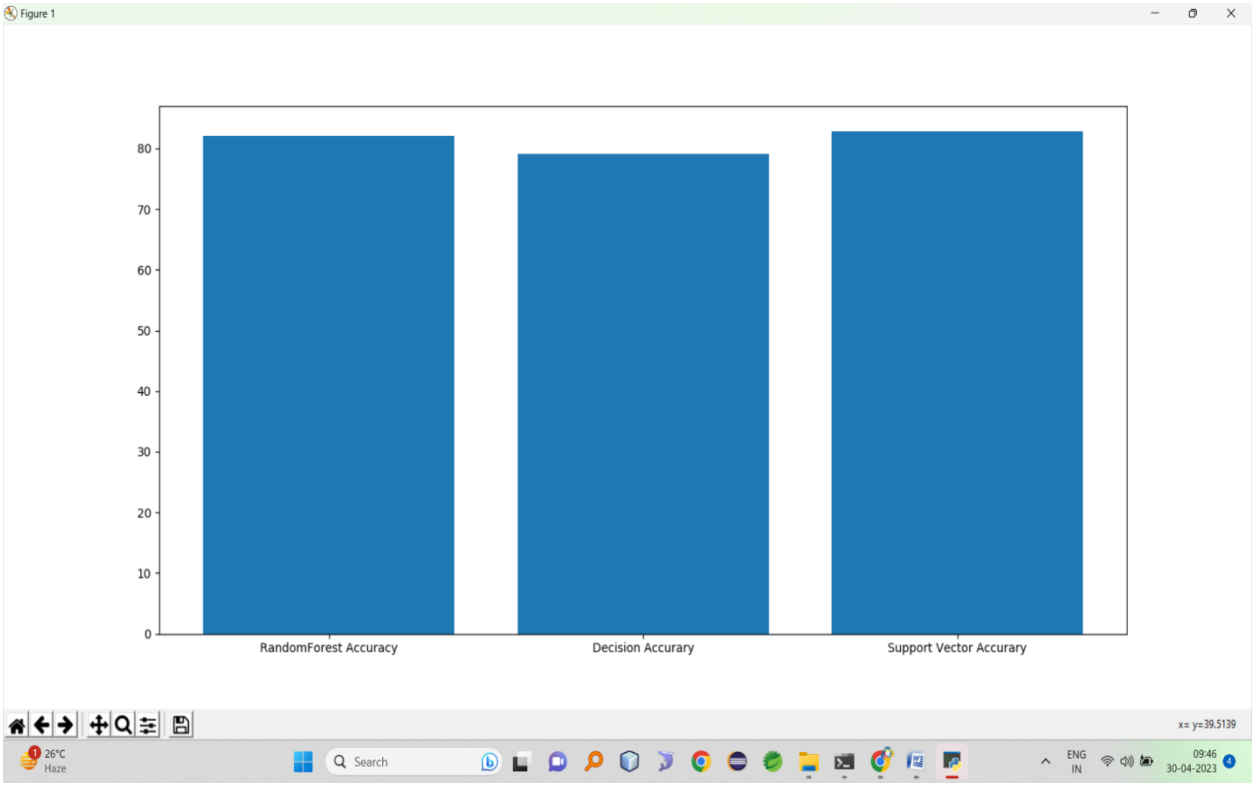

Fig 3: Accuracy Comparison

Accuracy Comparison

Sno	Algorithm Name	Accuracy	Efficiency
1	Random Forest	82%	82%

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, August : 2023

2	**Decision Tree**	**78%**	**78%**
3	**SVM**	**82%**	**82%**

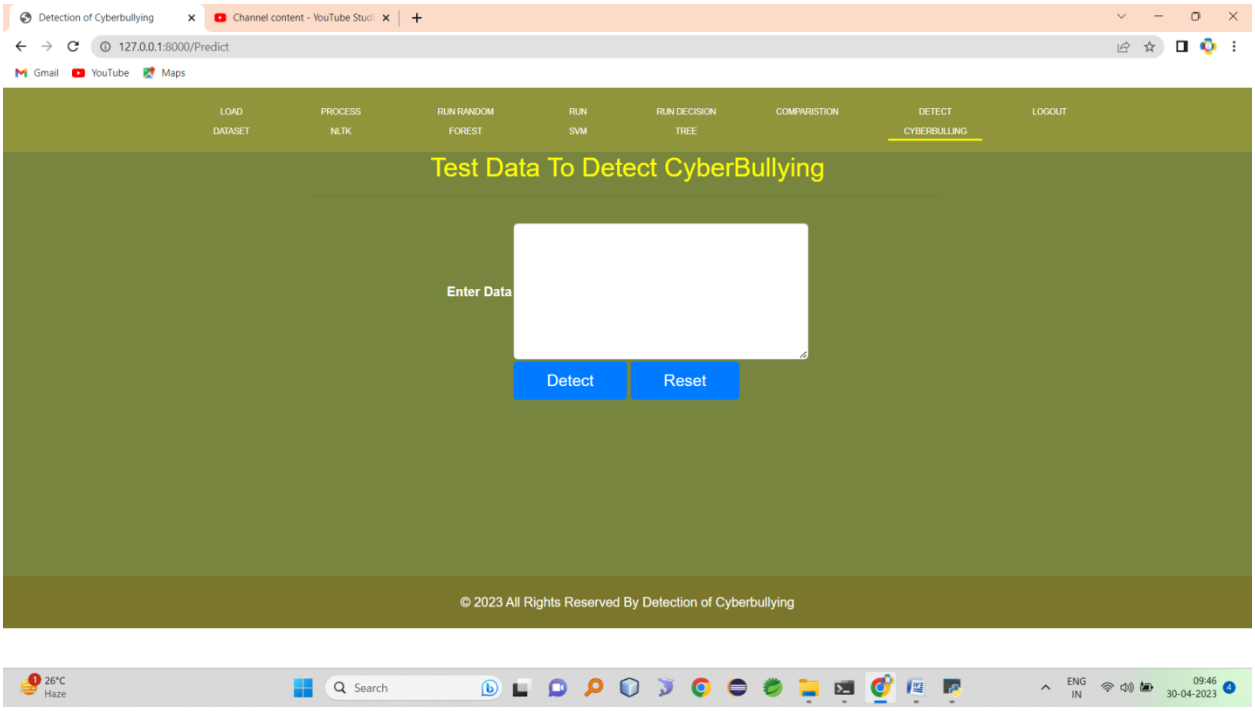

Fig 4: Input data

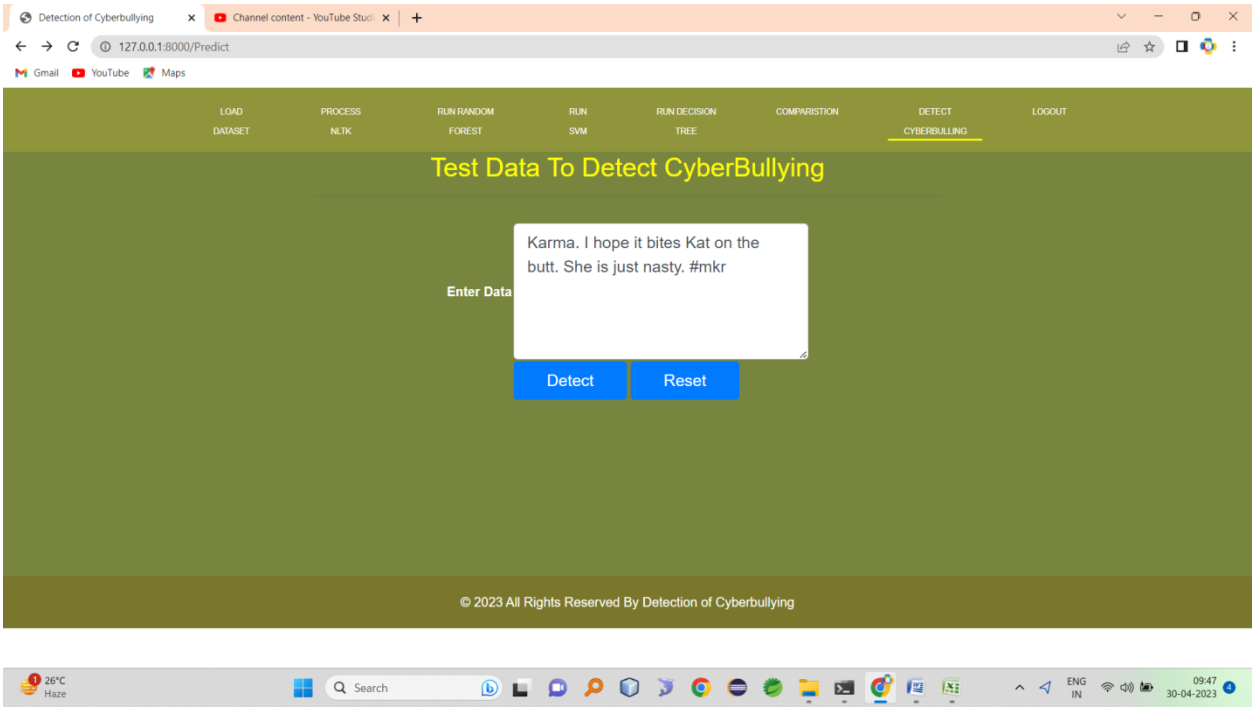

Fig 5: Predict output based on input data

5.CONCLUSION

There is a need to restrict the growth of cyberbullying because it can be dangerous and result in unfortunate events like suicide, depression, and other problems. Consequently, it is crucial to recognise cyberbullying on social media networks. More data and better-classified user information are now available for numerous other types of cyberattacks. On social media platforms, cyberbullying detection can be used to block users who attempt to engage in such behaviour. In this research, we suggested a detection architecture for cyberbullying to address the issue. We talked about the data architecture for hate speech on Twitter and personal attacks on Wikipedia. Natural Language Processing methods worked well for hate speech, with accuracy rates of over 90 percent Because tweets containing hate speech had vulgarity, which made it simple to spot, % using basic Machine learning techniques. Because of this, it performs better with BoW and Tf-Idf models than Word2Vec models. Although the three feature selection approaches worked similarly, it was challenging to identify personal assaults using the same model because the comments lacked a lot of learnable sentiment. When integrated with Multi Layered Perceptron's, Word2Vec models that exploit the context of features gave similar results in both datasets with significantly less features. Based on shifting nature

REFERENCES

[1] I. H. Ting, W. S. Liou, D. Liberona, S. L. Wang, and G. M. T. Bermudez, "Towards the detection of cyberbullying based on social network mining techniques," in Proceedings of 4th

International Conference on Behavioural, Economic, and Socio-Cultural Computing, BESC 2017, 2017, vol. 2018-January, doi: 10.1109/BESC.2017.8256403.

[2] P. Galán-García, J. G. de la Puerta, C. L. Gómez, I. Santos, and P. G. Bringas, "Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying," 2014, doi: 10.1007/978-3-319-01854-6_43.

[3] A. Mangaonkar, A. Hayrapetian, and R. Raje, "Collaborative detection of cyberbullying behavior in Twitter data," 2015, doi: 10.1109/EIT.2015.7293405.

[4] R. Zhao, A. Zhou, and K. Mao, "Automatic detection of cyberbullying on social networks based on bullying features," 2016, doi: 10.1145/2833312.2849567.

[5] V. Banerjee, J. Telavane, P. Gaikwad, and P. Vartak, "Detection of Cyberbullying Using Deep Neural Network," 2019, doi: 10.1109/ICACCS.2019.8728378.

[6] K. Reynolds, A. Kontostathis, and L. Edwards, "Using machine learning to detect cyberbullying," 2011, doi: 10.1109/ICMLA.2011.152.

[7] J. Yadav, D. Kumar, and D. Chauhan, "Cyberbullying Detection using Pre-Trained BERT Model," 2020, doi: 10.1109/ICESC48915.2020.9155700.

[8] M. Dadvar and K. Eckert, "Cyberbullying Detection in Social Networks Using Deep Learning Based Models; A Reproducibility Study," arXiv. 2018.

[9] S. Agrawal and A. Awekar, "Deep learning for detecting cyberbullying across multiple social media platforms," arXiv. 2018.

[10] Y. N. Silva, C. Rich, and D. Hall, "BullyBlocker: Towards the identification of cyberbullying in social networking sites," 2016, doi: 10.1109/ASONAM.2016.7752420

AUTHOR PROFILES

Dr.D.Bujji Babu, currently working as a Professor and Head in the Department of Master of Computer Applications, QIS College of Engineering and Technology, Ongole, Andhra Pradesh. He did his M.Tech(CSE) from JNTUK, Kakinada and Ph.D(CSE) from Acharya Nagarjuna University. He published more than 50 research papers in reputed peer reviewed Scopus indexed journals. He also attended and presented research papers in different national and international journals and the proceedings were indexed IEEE, Springer Link series. He visited the countries Kuching, Malaysia for attending and presenting his research articles. 3 Patent journals are published and in pipeline for grant. He wrote more than an dozen of monographs and published by the Technical Publishers. He Published Two Course content modules for the students of Acharya Nagarjuna University. He is the recognized research supervisor under JNTUK, Kakinada and guided several UG and PG Projects, currently supervising 3

research scholar under JNTUK. His area of interest is Software Engineering, Data Mining, Data Science, Big Data and Programming Languages. He is the Principal Investigator for the DST Sponsored Project and Co-PI for another DST Project.

MS. G. Vanaja as MCA Student in the department of Qis College of engineering and Technology (autonomous), Vengamukkapalem, Prakasam(DT).She has Completed B.Sc. in Computer Science from Sri Prathiba Degree College.Her areas of interests are Cloud Computing & Machine learning.