iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43683

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Cyberbullying Detection using Natural Language Processing

Md. Habeeb Ur Rahman[1], Mudigonda Divya[2], B. Ramya Reddy[3], Dr. K Sateesh Kumar[4], P. Ramya Vani[5]
[1, 2, 3]*Students of Department of Electronics and Computer Engineering, Sreenidhi Institute of Science and Technology, Hyderabad*
[4, 5] *Assistant Professor, Department of Electronics and Computer Engineering, Sreenidhi Institute of Science and Technology, Hyderabad*

Abstract: ***Around the world, the use of the Internet and social media has increased exponentially, and they have become an integral part of daily life. It allows people to share their thoughts, feelings, and ideas with their loved ones through the Internet and social media. But with social networking sites becoming more popular, cyberbullying is on the rise. Using technology as a medium to bully someone is known as Cyberbullying. The Internet can be a source of abusive and harmful content and cause harm to others. Social networking sites provide a great medium for harassment, bullies, and youngsters who use these sites are vulnerable to attacks. Bullying can have long-term effects on adolescents' ability to socialize and build lasting friendships Victims of cyberbullying often feel humiliated. social media users often can hide their identity, which helps misuse the available features. The use of offensive language has become one of the most popular issues on social networking. Text containing any form of abusive conduct that displays acts intended to hurt others is offensive language. Cyberbullying frequently leads to serious mental and physical distress, particularly for women and children, and sometimes forces them to commit suicide.***
The purpose of this project is to develop a technique that is effective to detect and avoid cyberbullying on social networking sites we are using Natural Language Processing and other machine learning algorithms. The dataset that we used for this project was collected from Kaggle, it contains data from Twitter that is then labeled to train the algorithm. Several classifiers are used to train and recognize bullying actions. The evaluation of the proposed Model for cyberbullying dataset shows that Logistic Regression performs better and achieves good accuracy than SVM, Ransom forest, Naive-Bayes, and Xgboost algorithm.
Keywords: Cyberbullying, Machine learning, Natural language processing, Social media, Kaggle, Dataset.

I. INTRODUCTION

Social networking sites are great tools for connecting with people. However, as social networking has become widespread, people are finding illegal and unethical ways to use these communities. We see that people, especially teens and young adults, are finding new ways to bully one another over the Internet. Close to 25% of parents in a study conducted by Symantec reported that, to their knowledge, their child has been involved in a cyberbullying incident [1]. According to the Cambridge dictionary the term cyberbullying is defined as the activity of using the internet to harm or frighten another person, especially by sending them unpleasant messages. Bullying has always been a part of society. With the inception of the internet, it was only a matter of time until bullies found their way onto this new and opportunistic medium. Traditional bullying may end in physical damage as well as emotional and psychological damage, as opposed to cyberbullying, where it is all emotional and psychological [2]. Thus, the detection and prevention of cyberbullying are important to protect teenagers.

In this context, we suggest a cyberbullying detection model based on Natural Language Processing and Machine Learning that can detect whether a text relates to cyberbullying or not. We have investigated several Machine learning algorithms, including Naive Bayes, Support Vector Machine, Logistic Regression, Random Forest, and Xgboost in the proposed cyberbullying detection model. the cyberbullying detection framework consists of two major parts shown in 1. The first part is called NLP (Natural Language Processing) and the second part is named ML (Machine learning). In the first phase, datasets containing bullying texts, messages, and posts are collected and prepared for the machine learning algorithms using natural language processing. We conduct experiments with datasets collected from Kaggle which contains Twitter comments and posts. For performance analysis. The results indicate that the Logistic Regression performs better and achieves good accuracy than the SVM, Ransom forest, Naive-Bayes, and Xgboost algorithm.

The rest of the paper is organized as follows. Section II shows several related works. Section III describes the proposed approach. Section IV shows the experimental results and the evaluation of the proposed approach. Finally, Section V concludes the paper.

©IJRASET: All Rights are Reserved

II. RELATED WORK

There are several works on machine learning-based cyber-bullying detection. A supervised machine learning algorithm was proposed using a bag-of-words approach to detect the sentiment and contextual features of a sentence [3]. This algorithm shows barely 61.9% of accuracy. Massachusetts Institute of Technology conducted a project called Ruminati [4] employing a support vector machine to detect cyberbullying of you-tube comments. The researcher combined detection with common sense reasoning by adding social parameters. The result of this project was improved to 66.7% accuracy for applying probabilistic modeling. Reynolds et al. [5] proposed a language-based cyberbullying detection method that shows 78.5% of accuracy. The author Nandhini et al. [6] have proposed a model that uses the Naïve Bayes machine learning approach and by their work, they achieved 91% accuracy and got their dataset from MySpace.com, and then they proposed another model [7] Naïve Bayes classifier and genetic operations (FuzGen) and they achieved 87% accuracy. While Chavan et al. [8] used two classifiers: logistic regression and support vector machine. The logistic regression achieved 73.76 accuracy and 60% recall and 64.4% Precision. While for the support vector machine they achieved 77.65% accuracy and 58% recall and 70% precision and they got their dataset from Kaggle. Furthermore, Ting et al. [9] proposed a technique based on SNM, they collected their data from social media and then used SNA measurements and sentiments as features. Seven experiments were made and they achieved around 97% precision and 71% recall. Furthermore, Harsh Dani et al. [10] introduced a new framework called SICD, they used KNN for classification. Finally, they achieved 0.6105 F1 score and 0.7539 AUC score. The authors Nobata et al. [11] showed that using abusive language has increased recently, They used a framework called Vowpal wabbit for classification, and they also developed a supervised classification methodology with NLP features that outperform the deep learning approach, The F-Score reached 0.817 using dataset collected from comments posted on Yahoo News and Finance.

III.PROPOSED APPROACH

The cyberbullying detection framework consists of two major parts. The first part is called NLP (Natural Language Processing) and the second part is ML (Machine learning) [12].

A. Methodology

1) Natural Language Processing (NLP) in Cyberbullying Detection

One direction in this field is to detect offensive content using Natural Language Processing (NLP). The most explanatory method for presenting what happens within a Natural Language Processing system is using the “levels of language” approach [13]. These levels are used by people to extract meaning from text or spoken languages. This leveling refers to language processing relying mainly on formal models or representations of knowledge related to these levels [14]. Moreover, language processing applications distinguish themselves from data processing systems by using the knowledge of the language. The analysis of natural language processing has the following levels:

- Phonology level (knowledge of linguistic sounds)
- Morphology level (knowledge of the meaningful components of words)
- Lexical level (deals with the lexical meaning of words and parts of speech analyses)
- Syntactic level (knowledge of the structural relationships between words)
- Semantic level (knowledge of meaning)
- Discourse level (knowledge about linguistic units more extensive than a single utterance)
- Pragmatic level (knowledge of the relationship of meaning to the goals and intentions of the speaker)

Yin et al. [16]; Reynolds et al. [17]; and Dinakar et al. [15] are the earliest researchers working in NLP cyberbullying detection, who investigated predictive strength n-grams, part-speech information (e.g., first and second pronoun), and sentiment information based on profanity lexicons for this task (with and without TF-IDF weighting). Similar features were also used for detecting events related to cyberbullying and fine-grained categories of text in [18] To conclude, some of the common word representation techniques used and proven to improve the classification accuracy [19] are Term Frequency (TF) [20], Term Frequency-Inverse Document Frequency (TF-IDF) [21], Global Vectors for Word Representation (GloVe) [22], and Word2Vec [23]. One of the main limitations of NLP is that of contextual expert knowledge. For instance, many dubious claims about the detection of sarcasm, but how one would detect sarcasm in a short post like “Great game!” responded to a defeat. Therefore, it is not about linguistics; it is about possessing knowledge relevant to the conversation.

2) Machine Learning in Cyberbullying Detection

Machine learning-based cyberbullying keywords is another direction of cyberbullying detection, which has been used widely by several researchers. Moreover, Machine learning (ML) is a branch of artificial intelligence technology that gives systems the capability to learn and develop automatically from experience without being specially programmed, often categorized as supervised, semi-supervised or unsupervised algorithms [24]. Several training instances in supervised algorithms are utilized to build a model that generates the desired prediction (i.e., based on annotated/labeled data). As cyberbullying is considered a classification issue (i.e., categorizing an instance as offensive or non-offensive), several supervised learning algorithms have been employed in this study for the further evolution of their classification accuracy and performance in detecting cyberbullying in SM, in particular on Twitter. The classifiers adopted in the current study are as follows:

3) Logistic Regression

Logistic regression is one of the well-known techniques introduced in the field of statistics by machine learning [25]. Logistic regression is an algorithm that constructs a separate hyper-plane between two datasets utilizing the logistic function [26]. The logistic regression algorithm takes features (inputs) and produces a forecast according to the probability of a class suitable for the input. For instance, if the likelihood is ≥0.5, the instance classification will be a positive class; otherwise, the prediction will be for the other class (negative class) [27], as given in Equation (1). In logistic regression was used in the implementation of predictive cyberbullying models.

$$h\theta\ (x) = 1/1 + e\ {-\theta Tx}$$

4) Logistic Light Gradient Boosting Machine

LightGBM is one of the powerful boosting algorithms in machine learning, and it is known as a gradient boosting framework that uses a tree-based learning algorithm [28]. However, it performs better compared to XGBoost and CatBoost [29]. Gradient-based One-side Sampling (GOSS) is used in LightGBM to classify the observations used to compute the separation. The LightGBM has the primary advantage of modifying the training algorithm, which significantly increases the process, and leads in many cases to a more efficient model. LightGBM has been used in many classification fields, such as online behavior detection [30] and anomalies detection in big accounting data [31]. However, LightGBM was not commonly used in the area of cyberbullying detection. Thus, in this study, we attempt to explore LightGBM in cyberbullying detection to evaluate its classification accuracy.

$$\theta = \theta - \eta \cdot \nabla\theta\ J\ (\theta; x(i); y(i)),$$

5) Random Forest

Random Forest (RF) classifier is an ensemble algorithm [32] that matches multiple decision-tab classifiers on different data sub-samples, using average data to enhance predictive accuracy and control of fitting [33]. Ensemble algorithms combine more than one algorithm of the same or different kinds for classifying data. RF was commonly used in the literature for the development of cyberbullying prediction models; examples are the studies conducted by. Consequently, RF consists of several trees used randomly to pick the variables for the classifier data. In the following four simplified steps, the construction of the RF takes place. In the training data, N is the number of examples (cases), and M is the number of attributes in the classifier.

6) Multinomial Naive Bayes

Multinomial Naive Bayes (Multinomial NB) is widely used for document/text classification problems. However, in the cyberbullying detection field, NB was the most commonly used to implement cyberbullying prediction models, such as in [34] and, NB classifiers were developed by applying the theorem of Bayes among features. This model assumes that a parametric model produces the text and makes use of training data to determine Bayes-optimal parameter estimates of the model. With those approximations, it categorizes produced test data [35]. NB classifiers can accommodate an arbitrary number of separate continuous or categorical functions. Assuming the functions are distinct, a task for estimating high-dimensional density is reduced to estimating one-dimensional kernel density. The NB algorithm is a learning algorithm based on the Bayes theorem's use with strong (naive) assumptions of independence. Therefore, in [36], NB was discussed in detail

7) Support Vector Machine Classifier

Support Vector Machine (SVM) is a supervised machine learning classifier widely utilized in text classification [37]. SVM turns the original feature space into a user-defined kernel-based higher-dimensional space and then seeks support vectors for optimizing the distance (margin) between two categories. SVM originally approximates a hyperplane separating the two categories.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

SVM accordingly selects samples from both categories, which are nearest to the hyperplane, referred to as support vectors [38]. SVM seeks to efficiently distinguish the two categories (e.g., positive and negative). If the dataset is separable by nonlinear

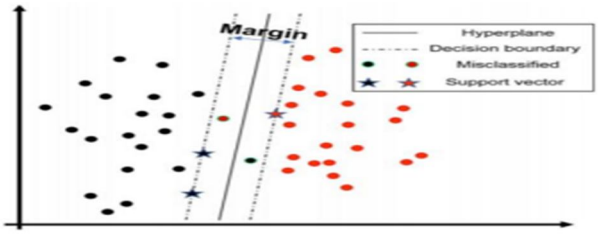


Figure 1: SVM classifier implementation for a dataset

boundaries, specific kernels are implemented in the SVM to turn the function space appropriately. Soft margin is utilized to prevent overfitting by giving less weighting to classification errors along the decision boundaries for a dataset that is not easily separable [39]. In this research, we utilize SVM with a linear kernel for the basis function. Figure 1 shows the SVM classifier implementation for a dataset with two features and two categories where all samples for the training are depicted as circles or stars. Support vectors (referred to as stars) are for each of the two categories from the training samples, meaning that they are nearest to the hyperplane among the other training samples. Two results of the training were misclassified because they were on the wrong side of the hyperplane. Therefore, SVM was used to construct cyberbullying prediction models in [40] and found to be effective and efficient. However, the work in [37] reported that the accuracy decreased when the data size increased, suggesting that SVM may not be ideal for dealing with frequent language ambiguities typical of cyberbullying.

B. Implementation

This section describes the implementation Model for cyberbullying detection on Twitter, its visualization, and the proposed methodology for conducting sentiment analysis on the dataset selected, as well as discussing the evaluation metrics of each classifier used.

1) Dataset

Detecting cyberbullying in social media through cyberbullying keywords and using machine learning for detection are theoretical and practical challenges. From a practical perspective, the researchers are still attempting to detect and classify offensive content based on the learning model. However, the classification accuracy and the implementation of the right model remain a critical challenge to construct an effective and efficient cyberbullying detection model. In this study, we used a Dataset collected from Kaggle that consists of posts and comments taken from social media platform called Twitter. A global dataset of 26,835 tweets to evaluate five classifiers that are commonly used in cyberbully content detection. Therefore, our dataset is taken from two sources [8,45]; and has been divided into two parts. The first part contains 70% of the tweets used for training purposes, and the other part contains 30% used for predications purposes. The evolution of each classifier will be conducted based on the performance metrics. The data set contains labeled posts, these are labeled as offensive and Non-offensive. The distribution of offensive and non-offensive data is 54.7% and 45.3% respectively as shown in fig 2.

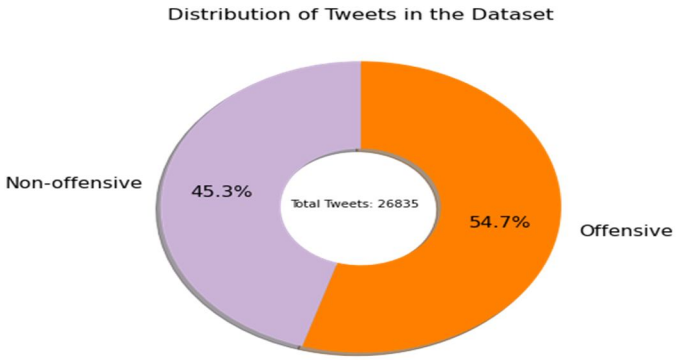


Figure 2: Distribution of Tweets in the Database

2) Model Overview

Figure 3 illustrates the proposed model of cyberbullying detection, which has four phases: the preprocessing phase, the feature extraction phase, the classification phase, and the evaluation phase.

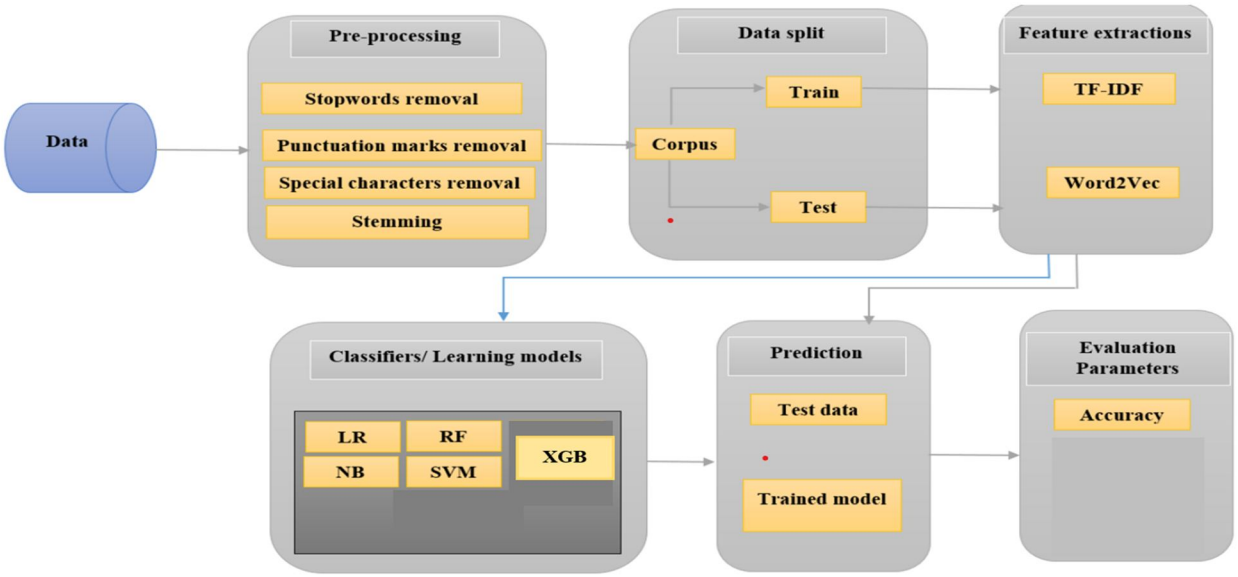


Figure 3: Proposed system for Cyberbullying Detection

3) Feature Extraction

Feature extraction is a critical step for text classification in cyberbullying. In the proposed model, we have used TF-IDF and Word2Vec techniques for feature extraction. TF-IDF is a combination of TF and IDF (term frequency-inverse document frequency), and this algorithm is based on word statistics for text feature extraction. This model considers only the expressions of words that are the same in all texts [72]. Therefore, TF-IDF is one of the most commonly used feature extraction techniques in text detection [16]. Word2Vec is a two-layer neural net that "vectorizes" words to process text. Its input is a corpus of text, and its output is a set of vectors: attribute vectors representing words in that structure [49]. The Word2Vec method uses two hidden layers of shallow neural networks, continuous bag-of-words (CBOW), and the Skip-gram model to construct a high-dimensional for each word [15]. The Skip-gram model is based on a corpus of terms w and meaning c. The aim is to increase the likelihood of:

$$\text{argmax } \theta \; Y \; w \in T \; [\; Y \; c \in c \; p(c \mid w; \theta)],$$

where T refers to text, and θ is a parameter of p (c |w; θ). Figure 4 illustrates the Word2Vec model architecture, where the CBOW model attempts to find a word based on previous terms, while Skip-gram attempts to find terms that could fall in the vicinity of each word.

4) Classification Techniques

In this study, various classifiers have been used to classify whether the tweet is cyberbullying or non-cyber bullying. The classifier models constructed are LR, Light LGBM, SGD, RF, AdaBoost, naïve Bayes, and SVM. The effectiveness of a proposed model was examined in this study by utilizing serval evaluation measures to evaluate how successfully the model can differentiate cyberbullying from non-cyberbullying. In this study, seven machine learning algorithms have been constructed, namely, LR, Light LGBM, SGD, RF, AdaBoost, Naive Bayes, and SVM. It is essential to review standard assessment metrics in the research community to understand the performance of conflicting models. The most widely used criteria for evaluating SM platforms (e.g., Twitter) with cyberbullying classifiers are as follows: Accuracy Accuracy calculates the ratio of the actual detected cases to the overall cases, and it has been utilized to evaluate models of cyberbullying predictions in [60,65,79]. Therefore, it can be calculated as follows:

$$\text{Accuracy} = (tp + tn) \; (tp + fp + tn + fn)$$

where tp means true positive, tn is a true negative, fp denotes false positive, and fn is a false negative.

IV. RESULTS

The proposed model utilizes the selected five ML classifiers with feature extraction techniques. These techniques are set empirically to achieve higher accuracy. For instance, LR achieved the best accuracy in our dataset, where the classification accuracy 94%. RF and XgBoost have achieved almost the same accuracy 87.6% and 86.9% respectively, but RF performs better than XgBoost. Multinomial NB has achieved low accuracy with a detection rate of 84.1% and we can notice that the excellent recall levels out the low precision. Finally, SVM has achieved the lowest accuracy in our dataset, as shown in Table 1. Nevertheless, it achieved the best recall compared to the rest of the classifiers implemented in the current research. Furthermore, some studies have looked at the automatic cyberbullying detection incidents; for example, an effect analysis based on lexicon and SVM was found to be effective in detecting cyberbullying. However, the accuracy decreased when data size increased, suggesting that SVM may not be ideal for dealing with common language ambiguities typical of cyberbullying [61]. This proves that the low accuracy achieved by SVM is due to the large dataset used in this research.

This research computed the five classifiers' performances using the F-measure metric, as shown in Figure 4. Furthermore, the performances of all ML classifiers are enhanced by producing additional data utilizing data Future Internet 2020, 12, 187 14 of 20 synthesizing techniques. Multinomial NB assumes that every function is independent, but this is not true in real situations [115].

Table 1: Accuracy of Algorithms

	Algorithms	Accuracy
0	Naive-Bayes	0.791330
1	SVM	0.789219
2	Logistic Regression	0.940149
3	Random forest	0.591976
4	Xgboost	0.755186

Therefore, it does not outperform LR in our research as well. As stated in [116], LR performs well for the binary classification problem and works better as data size increases. LR updates several parameters iteratively and tries to eliminate the error. Simultaneously, SGD uses a single sample and uses a similar approximation to update the parameters. Therefore, SGD performs almost like LR, but the error is not as reduced as in LR [92]. Consequently, it is not surprising that LR also outperforms the other classifiers in our study.

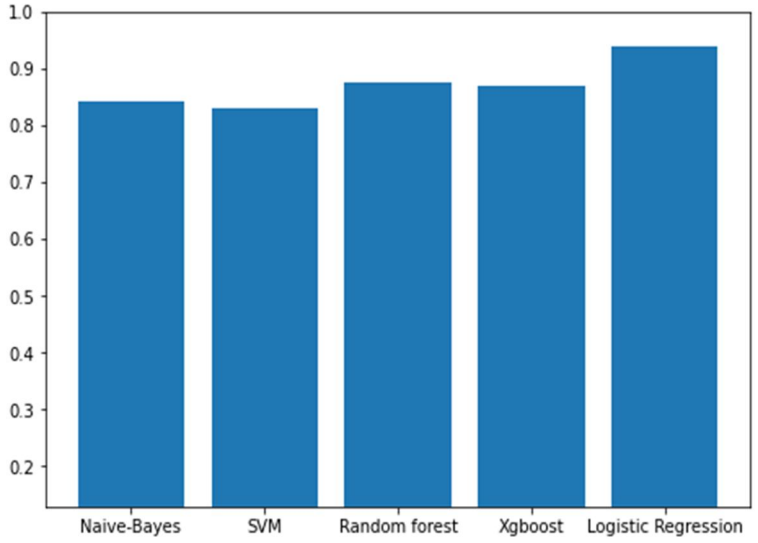


Figure 4: Accuracies of different Algorithms

V. CONCLUSION

Cyberbullying has become a severe problem in modern societies. This paper proposed a cyber-bully detection model whereby several classifiers based on NLP(TF-IDF) and Word2Vec feature extraction have been used. Furthermore, various methods of text classification based on machine learning were investigated. The experiments were conducted on a global Twitter dataset. The experimental results indicate that LR achieved the best accuracy in our dataset, where the classification accuracy and 94.01%. This means that LR performs better than other classifiers. Moreover, during the experiments, it was observed that LR performs better as data size increases and obtains the best prediction time compared to other classifiers used in this study. The feature extraction is a critical aspect of machine learning to enhance the detection accuracy. In this paper, we did not investigate many feature extraction techniques. Thus, one of the improvements is to incorporate and test different feature extractions to improve the detection rate of both classifiers LR and SGD. Another limitation that we are working on is building a real-time cyberbully detection platform, which will be useful to instantly detect and prevent the cyberbullying. Another research direction is working on cyberbully detection in various languages, mainly in Telugu and Hindi contexts.

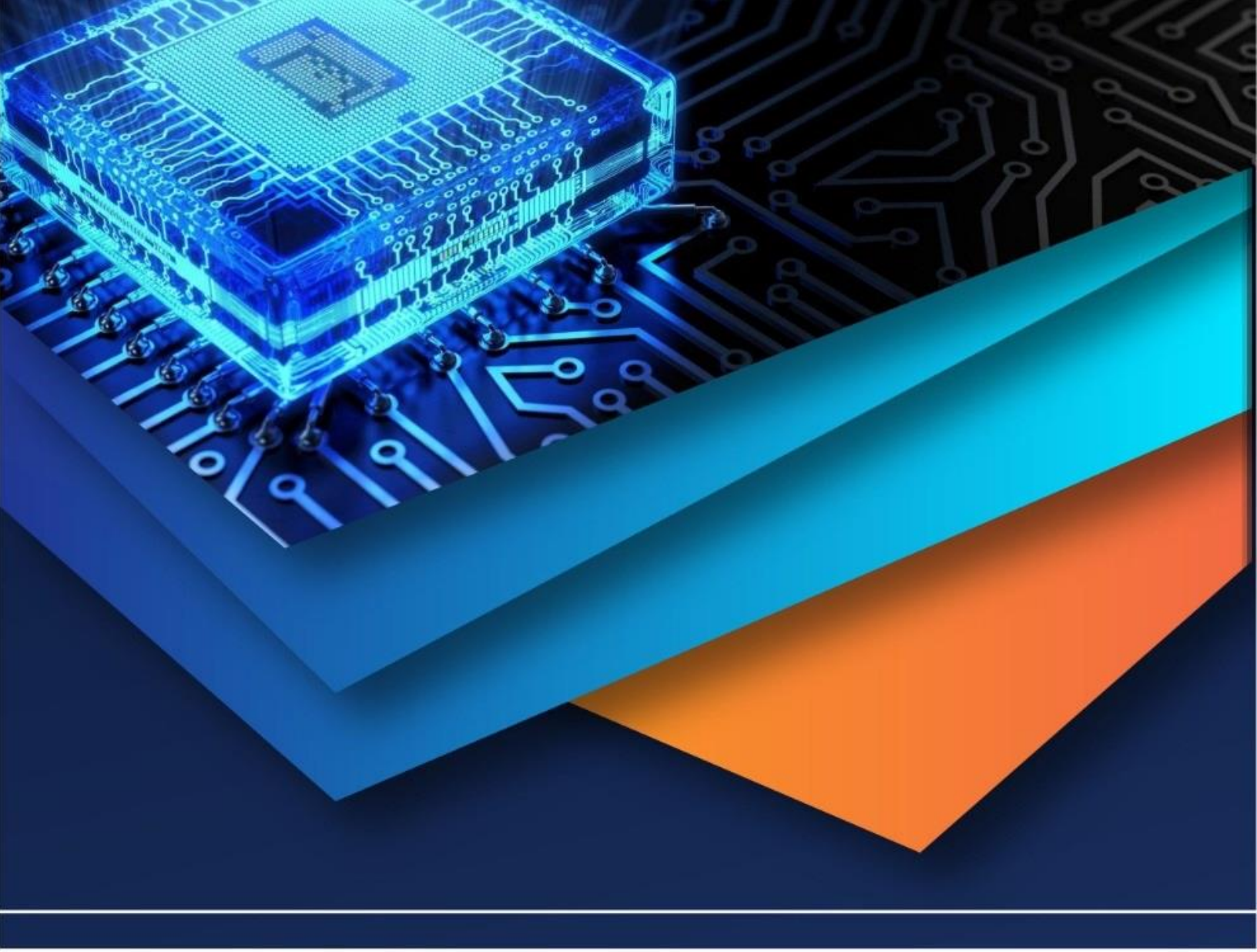
REFERENCES

[1] D. Poeter. (2011) Study: A Quarter of Parents Say Their Child Involved in Cyberbullying. pcmag.com. [Online]. Available: http://www.pcmag.com/article2/0,2817,2388540,00.asp.
[2] Hani J, Nashaat M, Ahmed M, Emad Z, Amer E, Mohammed A. Social media cyberbullying detection using machine learning. Int. J. Adv. Comput. Sci. Appl. 2019;10(5):703-7.
[3] Michele Di Capua, Emanuel Di Nardo, and Alfredo Petrosino. Un-supervised cyberbullying detection in social networks. In Pattern Recognition (ICPR), 2016 23rd International Conference on, pages 432–437. IEEE, 2016.
[4] K. Dinakar, R. Reichart, and H. Lieberman, "Modeling the detection of textual cyberbullying," in In Proceedings of the Social Mobile Web. Citeseer, 2011.
[5] K. Reynolds, A. Kontostathis, and L. Edwards, "Using machine learning to detect cyberbullying," in 2011 10th International Conference on Machine learning and applications and workshops, vol. 2. IEEE, 2011, pp. 241–244.
[6] 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 (ICARCSET 2015), page 20. ACM, 2015.
[7] B Sri Nandhini and JI Sheeba. Online social network bullying detection using intelligence techniques. Procedia Computer Science, 45:485–492, 2015.
[8] Vikas S Chavan and SS Shylaja. Machine learning approach for detection 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.
[9] I-Hsien Ting, Wun Sheng Liou, Dario Liberona, Shyue-Liang Wang, and Giovanny Mauricio Tarazona Bermudez. Towards the detection of cyberbullying based on social network mining techniques. In Behavioral, Economic, Socio-cultural Computing (BESC), 2017 International Conference on, pages 1–2. IEEE, 2017.
[10] Harsh Dani, Jundong Li, and Huan Liu. Sentiment informed cyberbullying detection in social media. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pages 52– 67. Springer, 2017.
[11] Chikashi Nobata, Joel Tetreault, Achint Thomas, Yashar Mehdad, and Yi Chang. Abusive language detection in online user content. In Proceedings of the 25th international conference on world wide web, pages 145–153. International World Wide Web Conferences Steering Committee, 2016.
[12] Islam, M. M., Uddin, M. A., Islam, L., Akter, A., Sharmin, S., & Acharjee, U. K. (2020). Cyberbullying Detection on Social Networks Using Machine Learning Approaches. 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE).
[13] Louppe, G. Understanding random forests: From theory to practice. arXiv 2014, arXiv:1407.7502.
[14] Novalita, N.; Herdiani, A.; Lukmana, I.; Puspandari, D. Cyberbullying identification on twitter using random forest classifier. J. Physics Conf. Ser. 2019, 1192, 012029. [CrossRef].
[15] García-Recuero, Á. Discouraging Abusive Behavior in Privacy-Preserving Online Social Networking Applications. In Proceedings of the 25th International Conference Companion on World Wide Web—WWW '16 Companion, Montreal, QC, Canada, 11–15 April 2016; Association for Computing Machinery (ACM): New York, NY, USA, 2016; pp. 305–309.
[16] Chatterjee, R.; Datta, A.; Sanyal, D.K. Ensemble Learning Approach to Motor Imagery EEG Signal Classification. In Machine Learning in Bio-Signal Analysis and Diagnostic Imaging; Elsevier BV: Amsterdam, The Netherlands, 2019; pp. 183–208.
[17] Misra, S.; Li, H. Noninvasive fracture characterization based on the classification of sonic wave travel times. In Machine Learning for Subsurface Characterization; Elsevier BV: Amsterdam, The Netherlands, 2020; pp. 243–287.
[18] Ibn Rafiq, R.; Hosseinmardi, H.; Han, R.; Lv, Q.; Mishra, S.; Mattson, S.A. Careful What You Share in Six Seconds.ss2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015—ASONAM '15; Association for Computing Machinery (ACM): New York, NY, USA, 2015; pp. 617–622.
[19] Tarwani, S.; Jethanandani, M.; Kant, V. Cyberbullying Detection in Hindi-English Code-Mixed Language Using Sentiment Classification. In Communications in Computer and Information Science; Springer Science and Business Media LLC: Singapore, 2019; pp. 543–551.
[20] Raza, M.O.; Memon, M.; Bhatti, S.; Bux, R. Detecting Cyberbullying in Social Commentary Using Supervised Machine Learning. In Advances in Intelligent Systems and Computing; Springer Science and Business Media LLC: Singapore, 2020; pp. 621–630.
[21] Galán-García, P.; De La Puerta, J.G.; Gómez, C.L.; Santos, I.; Bringas, P.G. Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying. Log. J. IGPL 2015, 24, jzv048. [CrossRef]
[22] Akhter, A.; Uzzal, K.A.; Polash, M.A. Cyber Bullying Detection and Classification using Multinomial Naïve Bayes and Fuzzy Logic. Int. J. Math. Sci. Comput. 2019, 5, 1–12. [CrossRef]
[23] Nandakumar, V. Cyberbullying revelation in Twitter data using naïve Bayes classifier algorithm. Int. J. Adv. Res. Comput. Sci. 2018, 9, 510–513. [CrossRef]

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[24] Dinakar, K.; Reichart, R.; Lieberman, H. Modeling the detection of textual cyberbullying. In Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, Barcelona, Spain, 17–21 July 2011.
[25] Snakenborg, J.; Van Acker, R.; Gable, R.A. Cyberbullying: Prevention and Intervention to Protect Our Children and Youth. Prev. Sch. Fail. Altern. Educ. Child. Youth 2011, 55, 88–95. [CrossRef]
[26] Patchin, J.W.; Hinduja, S. Traditional and Nontraditional Bullying Among Youth: A Test of General Strain Theory. Youth Soc. 2011, 43, 727–751. [CrossRef]
[27] Tenenbaum, L.S.; Varjas, K.; Meyers, J.; Parris, L. Coping strategies and perceived effectiveness in fourth through eighth grade victims of bullying. Sch. Psychol. Int. 2011, 32, 263–287. [CrossRef]
[28] Ybarra, M.L.; Mitchell, K.J.; Wolak, J.; Finkelhor, D. Examining Characteristics and Associated Distress Related to Internet Harassment: Findings from the Second Youth Internet Safety Survey. Pediatrics 2006, 118, e1169–e1177. [CrossRef] [PubMed]
[29] Smith, P.K.; Mahdavi, J.; Carvalho, M.; Fisher, S.; Russell, S.; Tippett, N. Cyberbullying: Its nature and impact in secondary school pupils. J. Child Psychol. Psychiatry 2008, 49, 376–385. [CrossRef]
[30] Bosse, T.; Stam, S. A Normative Agent System to Prevent Cyberbullying. In Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Lyon, France, 22–27 August 2011; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2011; Volume 2, pp. 425–430.
[31] Reynolds, K.; Kontostathis, A.; Edwards, L. Using Machine Learning to Detect Cyberbullying. In Proceedings of the 2011 10th International Conference on Machine Learning and Applications and Workshops, Honolulu, HI, USA, 18–21 December 2011; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2011; Volume 2, pp. 241–244.
[32] Salminen, J.; Hopf, M.; Chowdhury, S.A.; Jung, S.-G.; Almerekhi, H.; Jansen, B.J. Developing an online hate classifier for multiple social media platforms. Hum. Cent. Comput. Inf. Sci. 2020, 10, 1–34. [CrossRef]
[33] Dinakar, K.; Jones, B.; Havasi, C.; Lieberman, H.; Picard, R. Common Sense Reasoning for Detection, Prevention, and Mitigation of Cyberbullying. ACM Trans. Interact. Intell. Syst. 2012, 2, 1–30. [CrossRef]
[34] Hinduja, S.; Patchin, J.W. Cyberbullying: An Exploratory Analysis of Factors Related to Offending and Victimization. Deviant Behav. 2008, 29, 129–156. [CrossRef
[35] Notar, C.E.; Padgett, S.; Roden, J. Cyberbullying: Resources for Intervention and Prevention. Univers. J. Educ. Res. 2013, 1, 133–145.
[36] Fanti, K.A.; Demetriou, A.G.; Hawa, V.V. A longitudinal study of cyberbullying: Examining riskand protective factors. Eur. J. Dev. Psychol. 2012, 9, 168–181. [CrossRef]
[37] Joachims, T. Text categorization with Support Vector Machines: Learning with many relevant features. In the Computer Vision—ECCV 2018; Springer Science and Business Media LLC: Berlin, Germany, 1998; pp. 137–142.
[38] Ybarra, M.L.; Mitchell, K.J. Prevalence and Frequency of Internet Harassment Instigation: Implications for Adolescent Health. J. Adolesc. Health 2007, 41, 189–195. [CrossRef]
[39] Havas, J.; De Nooijer, J.; Crutzen, R.; Feron, F.J.M. Adolescents' views about an internet platform for adolescents with mental health problems. Health Educ. 2011, 111, 164–176. [CrossRef]
[40] Nonauharcelement.Education.gouv.fr. Non Au Harcèlement AppelezLe3020.2020.Availableonline:https://www.nonauharcelement.education.gouv.fr/ (accessed on 18 August 2020).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)