WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Cyberbullying Detection On Social Media Using Machine Learning

Harsh Agarwal Jagruti Jadhav Komal Nagar Babita Jaybhaye Shrikant Dhamdhere

Parvatibai Genba moze college of engineering,Pune-411001,India

A B S T R A C T

The popularity of online social networks has created massive social interaction among their users, and this leads to a huge amount of user-generated communication data. In recent years, Cyberbullying has grown into a major problem with the growth of online interaction and social media. Cyberbullying has been recognized recently as a serious national health concern among online social network users and developing an efficient detection model holds tremendous practical significance. In this paper, we have proposed set of unique features derived from Twitter based on these features, we developed a supervised machine learning solution for detecting cyberbullying in the Twitter. An evaluation demonstrates that our developed detection model based on our proposed features, achieved results with an area under the precision of 0.963 and an f-measure of 0.904. These results indicate that the proposed model based on these features provides a feasible solution to detecting Cyberbullying in online interacting environments

Keywords:cyberbullying; machine learning; classifiers; Naive Bayes; support vector machine (SVM); cyber-aggressive; supervised; cybercrime; online-communication;twitter

1. Introduction

Cyberbullying is a deliberate and repetitive act to harm or humiliate someone using information and communication technologies such as mobile phones, emails and social media [1] [2]. It is often categorized into various forms, such as cyber harassment (i.e. repetitively harassing and threatening someone), denigration/slandering (i.e. sharing false information about someone), flaming (i.e. brief insulting online interactions) and happy slapping (i.e. recording a session while a person is being bullied for circulation purpose), among others [2]. Impacts of cyberbullying are detrimental in nature, ranging from emotional (anger, fear, self-blame etc.) to psychological (low self-esteem, depression, suicidal etc.) and physical (loss of sleep, headache, eating disorder etc.).

Despite the various prevention and intervention strategies, cyberbullying perpetration has not decreased in the last one decade [3]. Recent studies have looked into automatically detecting cyberbullying incidents, for instance, an affect analysis based on a lexicon and Support Vector Machine (SVM) was found to be effective in detecting cyberbullying, however the accuracy decreased when the size of the data increased, suggesting that SVM may not be ideal in dealing with frequent language ambiguities typical for cyberbullying [4]. [5] automatically collected data from an in-game chat (World of Tanks) and found cyberbullying to be a learned behaviour (i.e. new players are less likely to engage in cyberbullying).

Cyberbullying and its impact on social media: Cyberbullying is not just limited to creating a fake identity and publishing/posting some embarrassing photo or video, unpleasant rumours about someone but also giving them threats. The impacts of cyberbullying on social media are horrifying, sometimes leading tothe death of some unfortunate victims. The behaviour of the victims also changes due to this, which affects their Emotions, self-confidence and a sense of fear is also seen in such people.

2. Background

Cyberbullying – a social media problem:

The use of information and communication technologies, particularly social media has revolutionized the manner in which people communicate and form relationships with one another, with statistics around the world indicating a high prevalence rate of social media applications. For instance, according to

* *Corresponding author.*8668217369;
E-mail address:8668217369

the recent report by Pew Research Centre (2018), Instagram (75%) and Snapchat (73%) were found to be most popular among those between 18 and 24 years old whereas Facebook and YouTube were more popular among those older than 50 (i.e. 68%). This unfortunately, provides an avenue for anti-social behaviours such as misogyny [6] [7], sexual predation sexism [8] and cyberbullying perpetration [9] [10] [11]. Facebook, for instance, is one of the most popular social

Facebook, for instance, is one of the most popular social media platforms that allows its users to create their own profiles, upload their photos and videos, and send messages (both private and public). It has a wide reach, as any comments or posts can reach thousands of people, especially through "liking" and "sharing" mechanisms, and thus allowing cyberbullies to distribute nasty or unwanted information about their victims easily [12]. Instagram allows its users to share photos and videos, to follow others and support Stories. Like7 Facebook, it is also easy for one to set up new, anonymous profiles for cyberbullying perpetration. Thevelocity and size of the distribution mechanism allow hostile comments or humiliating images to go viral within hours [12].

2.1. Features

This section specifically focuses on the features incorporated in our cyberbullying detection model. It encompasses user personalities focusing on Big Five and Dark Triad models, sentiment, emotion and Twitter-based features.

2.2. User personalities

One of the most comprehensive and popular method to determine personality is based on the Big Five model [13] [14], which is a hierarchical organization of personality traits in terms of five basic dimensions/facets: Extraversion - the tendency of being outgoing, sociable, to be interested in other people, assertive, active, paying more attention to external events and excitement seeking Agreeableness - the tendency to be kind, friendly, gentle, getting along with others and being warm to other people Conscientiousness – it presents how much a person pays attention to others when making decisions Neuroticism - the tendency to be depressed, fearful and moody Openness - the tendency to be creative, perceptive, thoughtful, broad-minded, and willing to make adjustments in activities in accordance with new ideas.

2.3. Cyberbullying detection and machine learning

Machine learning, an application of artificial intelligence provides systems the ability to automatically learn and improve from experience without being explicitly programmed, often differentiated as supervised, unsupervised or semi-supervised algorithms [15].The supervised algorithms take a set of training instances to build a model that generates a desired prediction for an unseen instance (i.e. based on labelled/annotated data), whereas unsupervised algorithms do not depend on labelled data, and thus often used for clustering problems [15] As cyberbullying is deemed to be a classification problem (i.e. categorizing an instance as bully or non-bully), the supervised learning algorithms were adopted in the present study.

Studies on cyberbullying detection are mainly based on superintending algorithms such as Naïve Bayes, SVM, Decision Trees (J48), and Random Forest, often with performance comparisons made between several of these classifiers [10] [16]. Naïve Bayes is a Bayesian theorem algorithm and is well-known for its ability to classify texts based on a probability (i.e., outcomes are based on the highest probability). It is therefore, wellsuited for real-time predictions, text classifications and recommendation systems [17]. It assumes independence between predictors, that is, the presence/absence of a feature is unrelated to the presence/absence of any otherfeature. Therefore, in the context of tweets, each word or feature is considered as a unique variable by Naive Bayes to determine the probability of that word/feature. For instance, [18] proposed a model for detecting cyberbullying using Na¨ıve Bayes, whereby the presence of an offensive word indicates cyberbullying, and the absence indicates otherwise. The authors however, did not evaluate their proposed model, but other similar studies such as [19] reported an overall accuracy of 63% using Naïve Bayes to detectcyberbullying based on YouTube comments.

The present study adopted a similar approach whereby the presence of specific features (or combination of features) (e.g., high number of followers-following or a negative personality) may result in a tweet to be classified as a bully. J48 is a popular decision tree algorithm, which uses the depth-first strategy that considers all the possible tests to split the dataset before one with the highest information gain is selected [20]. The trees contain several nodes, that is, root (main node, no incoming edges), internal (with incoming and outgoing edges) and leaf (no outgoing edges). Both the root and internal nodes correspond to each feature tested whereas the leaf node is the final classification. Therefore, in the context of cyberbullying, features such as number of followers, popularity, positive sentiment can be used as the root or internal nodes, whereas bully or not-bully will be the corresponding leaf node. J48 is generally easy to use and relatively fast, however the preparation of large decision trees (i.e., large datasets with many features) are complex and time-consuming (Zhao and Zhang, 2008). [16] explored the social network graphs features, namely the relationships between users and related features (e.g., number of friends), and network embeddedness (i.e., relationship between users) etc. using J48, with results indicating an accuracy of 62.8

3. Related work

A previous study proposed an approach for offensive language detection that was equipped with a lexical syntactic feature and demonstrated a higher precision than the traditional learningbased approach [21]. A YouTube databased study [22] applied SVM to detect cyberbullying, and determined that incorporating userbased content improved the detection accuracy of SVM. Using data sets from Myspace, developed a genderbased cyberbullying detection approach that used the gender feature in enhancing the discrimination capacity of a classifier [23]. included age and gender as features in their approach; however, these features were limited to the information provided by users in their online profiles [24]. Moreover, most studies determined that only a few users provided complete information about themselves in their online profiles. Alternatively, the tweet contents of these users were analysed to determine their age and gender [23]; [25]. Several studies on cyberbullying detection utilizedprofane words as a feature, thereby significantly improving the model performance. A recent study [26] proposed a model for detecting cyberbullies in Myspace and recognizing the pairwise interactions between

users through which the influence of bullies could spread. Nalini and Sheela proposed an approach for detecting cyberbullying messages in Twitter by applying a feature selection weighting scheme [27]. included pronouns, skip-gram, TFeIDF, and N-grams as additional features in improving the overall classification accuracy of their model [28].

4. Materials and methods

4.1. Experimental setup

Stepwise Procedure of SVM utilized in detecting the cyberbullying Steps:

1. For a particular location, a limited number of tweets will be fetched through dataset
2. The Data Pre-processing, Data Extraction will be performed on the fetched Tweets
3. Pre-processed tweets will be passed to SVM and Naïve Bayes model (see Developing the Model section) to calculate the probabilities of fetched tweets to check whether a fetched tweet is bullying or not.
4. If the probability of fetched tweet lies in the range of 0.5 to 1, then the tweet will not be considered as a bullied tweet.
5. Again, the list of tweets will be passed to the SVM and Naive Bayes model to predict the results of the tweets.
6. And again, the average probability of that tweet will be calculated and if it lies above 0.5 then it will be considered as a bullied tweet and it will be recorded in our database. If the average probability is less than 0.5 then the record will be removed from the database.

4.2. Developing the model

The entire model is divided into 3 major steps: Pre-processing, the algorithm, and feature extraction.

Pre-processing

The Natural Language Toolkit (NLTK) is used for the pre-processing of data. NLTK is used for tokenization of text patterns, to remove stop words from the text, etc. Tokenization: In tokenization, the input text is split as the separated words and words are appended to the list. Firstly, TweetTokenizer is used to tokenized text into the sentences. Then 4 different tokenizers are used to tokenize the sentences into the words:

o TweetTokenizer
o WordPunctTokenizer
o TreebankWordTokenizer
o PunctWordTokenizer

Lowering Text: It lowers all the letters of the words from the tokenization list. Example: Before lowering “Hey There” after lowering “hey there”. Removing Stop words: This is the most important part of the pre-processing. Stop words are useless words in the data. Stop words can be get rid of very easily using NLTK. In this stage stop words like -https, -, are removed from the text.
Wordnet lemmatize: Wordnet lemmatize finds the synonyms of a word, meaning and many more and links them to the one word

Feature Extraction:

In this step, the proposed model has transformed the data in a suitable form which is passed to the machine learning algorithms. The frequency dictionary is used to extract the features of the given data. Features of the data are extracted and put them in a list of features. Also, the frequency of word defining polarity (i.e., the text is Bullying or Non-Bullying) of each text is extracted and stored in the list of features

Algorithm Selection:

To detect social media bullying automatically, supervised Binary classification machine learning algorithms like SVM with linear kernel and Naive Bayes is used. The reason behind this is both SVM and Naive Bayes calculate the probabilities for each class (i.e. probabilities of Bullying and Non-Bullying tweets). Both SVM and NB algorithms are used for the classification of the two-cluster. Both the machine learning models were evaluated on the same dataset. But SVM outperformed Naive Bayes of similar work on the same dataset. Classification report [9] is also evaluated. The accuracy, recall, f-score, and precision are also calculated

Precision = TP / (TP+FP)

Recall =TP/(TP+FN)

F-Score = 2*(Precision*Recall) / (Precision + Recall)

Where ,

TP = True positive numbers

TN = True negative numbers

FN = False negative numbers

FP = False positive numbers

Support Vector Machine

Support Vector Machine is a supervised classification machine learning algorithm. SVM can be used for both regression and classification. SVM also calculates the probabilities for each category. SVM with non-Linear Kernel is used as our data is linearly separable.

HYPERPLANE: The main aim of the SVM is to find the hyperplane which divides the dataset into two categories. Many hyperplanes separate two categories of the data points. The main aim of the SVM is to find the hyperplane with a maximum margin. For 2 attributes hyperplane is justa line. As the number of features increases, it is very difficult to imagine the hyperplanes' dimension. In our model, as there are only 2 classes, i.e., Bullying and Non-Bullying hyperplane was just a line. SUPPORT VECTORS: Data points that are closer to the hyperplane are called Support Vectors. To maximize the marginal distance between classifiers support vectors are used and if delete this support vector it will change the hyperplanes' position.

Naive Bayes

Naive Bayes is a supervised probabilistic machine learning algorithm that can be used for classification. Bayes Theorem Formula: Naive Bayes models are used recommendation systems, sentiment analysis, and spam filtering. Naive Bayes algorithms are very easy to implement. Types of Classifiers: Gaussian Naive Bayes

Bernoulli Naive Bayes

Multinomial Naive Bayes

Since our data is not discrete Gaussian Naïve Bayes approach is used.

4.3. Proposed method

This section proposes the methodology and framework used for classification of comments. The steps involved are Normalization, standard Feature extraction, feature selection and finally classification.

Normalization: The Data set we have used contains list of comments and respective labels. These should be converted into feature vector which are used by our machine- learning algorithms. For this we use different Natural language processing techniques to obtain an accurate representation of the comments in feature vector form. We use various techniques based on our observations.

Removing unwanted strings: For the comments to be used by machine-learning algorithms they should be in standard form. Raw comments present in dataset which contains many unwanted strings like and many such encoding parts should be removed.

Hence the first step is to pre-process the comments by removing unwanted strings, hyphens and punctuations Correcting words: One of the reasons comments are classified as insulting is the presence of profane or abusive words. The total number of bad words present in comments is taken as one of the features. A dictionary of 500 bad words is compiled, which also includes variations of words (online forums sometimes use special characters to build an insulting word.

When we encounter such words, the dictionary helps to convert them into natural form. Also, Stemming is applied to capture bad word variations that are not contained in dictionary. Stemming reduces a word toits core root, for example embarrassing is reduced to embarrass. Here it is noted that stemming is only applied to bad word dictionary not on the dataset used, as it will lead to information loss. Again, a small dictionary and a spell checker is used to convert all variations of "you"," you're" (e.g., u, ur etc) which are present in the dataset as participant use them as part of flexible language. Following Standard Feature Extraction: To train machine learning algorithms, strings should be converted in feature vector. We use frequency dictionary The process occurs in following steps and used it class as a pair for unique identification

Counting: Count the number of times each of these tokens occurs in a tweet along with it class it being added to a class which is 3X1 matrix.

Additional Features: Capturing pronouns: It is been observed that cyberaggressive comments which are directed towards peers are perceived more negatively and results in cyberbullying. Comments containing a pronoun like 'you' followed by an insulting or profane words are peer directed comments which are taken as negative and teens get frustrated after encountering such comments. So, to detect such comments we have used the count of pronouns

as one of the features for detecting cyberbullying. To extract this feature, we calculate feature for pronoun present in comment. This feature is our strong hypothesis which greatly increases the accuracy and helps in detecting cyber-aggressive comments.

Feature Selection: The machine learning algorithms cannot handle all the features so we created whole new feature set consisting of three parameters positive count, negative count and bias term. These would reduce the time complexity and space complexity.

• Chi-Square Method: chi square (X2) method is commonly used for selecting best features. This metric calculates the cost of a feature using the value of the chi-squared statistics with respect to class. Classification: Once the features are built, we extract the best features using chi-squared test and apply the machine learning algorithms to train models on it. We have used SVM and logistic regression on our feature data. A brief summary of these algorithms is given below.

• Support vector machine (SVM): This algorithm maps the training data into feature space using kernel functions and then separates the dataset using large hyperplane. We have used non-linear kernel Rbf function

. • Logistic Regression: This algorithm provides probabilistic approach to data. The outcome are probabilities modelled as a function of predicted variables, using a logistic function.

5. Detection of cyberbullying incidents

Since effective prediction enables better targeted detection, we were interested in applying a similar methodology as in the prediction section to the training and testing of a detector. This distinction is of course the detection algorithm benefits from having access to the text comments from the discussion. In this section we only work on the data with non-zero negativity. The idea comes from having a first layer predictor with near to zero false positive.

In addition, a variety of non-text features were evaluated, including those features extracted from user behaviour (number of shared media objects, following, followers), media properties (likes, post time, caption) and image content. For example, we investigate the feature corresponding to the number of words. However, adding this feature does not provide any value to the classifier performance. It was observed that the number of words is considerably higher for examples of cyberbullying. The reason is the high correlation between the number of words and a set of variables with positive coefficients, namely “bitch”, “fuck”, “gay”, “hate”, “shut”, “suck”, “ugly”.

Similarly, we considered the “time interval” variable, i.e., the mean time between posts. This variable also has high correlation with cyberbullying indicator words and does not add to the classifier performance. Both of these supports our correlation analysis for “time interval” and “word count”. Another feature related to the media session is the number of likes the image has received, however it does not provide any improvement with a very small coefficient in the model. Another theme for future work is to obtain greater detail from the labelling surveys. Our experience was that streamlining the survey improved the response rate, quality and speed. However, we desire more detailed labelling, 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. Finally, we can cascade our predictor with a more complicated detection algorithm to make examining cyberbullying-prone media sessions more scalable.

6. Discussion

While this paper has introduced prediction of cyberbullying in a media-based mobile social network, there remain a number of areas for improvement. One theme for future work is to improve the performance of our classifier and used it to various media related cyberbullying activities. New algorithms should be considered, such as RNN and LSTM. More input features should be evaluated, such as new image features, mobile sensor data, etc. Incorporating image features needs to be automated by applying image recognition algorithms. Temporal behaviour of comments for a posted media should be taken into account in designing the detectionclassifier. In this work we have only considered the image content and image and user metadata for prediction of cyberbullying. However, based on the improvement seen in using a small number of text comments, we think that considering the commenting history of users in previously shared media can prove to be useful.

Another theme for future work is to obtain greater detail from the labelling surveys. Our experience was that streamlining the survey improved the response rate, quality and speed. However, we desire more detailed labelling, 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. Finally, we can cascade our predictor with a more complicated detection algorithm to make examining cyberbullying-prone media sessions more scalable.

7. Conclusion

An approach is proposed for detecting and preventing cyberbullying using Supervised Binary classification Machine Learning algorithms. Our model is evaluated on both Support Vector Machine and Naive Bayes, also for feature extraction, used the Frequency word dictionary. As the results show us that the accuracy for detecting cyberbullying content has also been great for Support Vector Machine non-linear of around 90.4% which is better than our [29]. Our model will help people from the attacks of social media bullies.

Acknowledgements

Acknowledgements and Reference heading should be left justified, bold, with the first letter capitalized but have no numbers. Text below continues as normal.

Appendix A. An example appendix

Authors including an appendix section should do so before References section. Multiple appendices should all have headings in the style used above. They will automatically be ordered A, B, C etc.

A.1. Example of a sub-heading within an appendix

There is also the option to include a subheading within the Appendix if you wish.

REFERENCES

[1] Vimala Balakrishnan. Cyberbullying among young adults in malaysia: The roles of gender, age and internet frequency. Computers in Human Behavior, 46:149–157, 2015.
[2] Robin M Kowalski, Gary W Giumetti, Amber N Schroeder, and Heather H Reese. Cyber bullying among college students: Evidence from multiple domains of college life. In Misbehavior online in higher education. Emerald Group Publishing Limited, 2012.
[3] Sameer Hinduja and Justin W Patchin. Bullying beyond the schoolyard: Preventing and responding to cyberbullying. Corwin press, 2014.
[4] Michal Ptaszynski, Pawel Dybala, TatsuakiMatsuba, Fumito Masui, RafalRzepka, and Kenji Araki. Machine learning and affect analysis against cyber-bullying. the 36th AISB, pages 7–16, 2010.
[5] 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.
[6] Maria Anzovino. Misogyny Detection on Social Media: A Methodological Approach. PhD thesis, Master's Thesis, Department of Informatics, Systems and Communication, 2018.
[7] Lu Cheng, Ruocheng Guo, and Huan Liu. Robust cyberbullying detection with causal interpretation. In Companion Proceedings of The 2019 World Wide Web Conference, pages 169–175, 2019.
[8] Simona Frenda, Somnath Banerjee, Paolo Rosso, and Viviana Patti. Do linguistic features help deep learning? the case of aggressiveness in mexican tweets. Computaci´on y Sistemas, 24(2):633–643, 2020.
[9] Robin M Kowalski, Susan P Limber, and Annie McCord. A developmental approach to cyberbullying: Prevalence and protective factors. Aggression and Violent Behavior, 45:20–32, 2019.
[10] DespoinaChatzakou, Nicolas Kourtellis, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Athena Vakali. Mean birds: Detecting aggression and bullying on twitter. In Proceedings of the 2017 ACM on web science conference, pages 13–22, 2017.
[11] HomaHosseinmardi, Sabrina Arredondo Mattson, Rahat Ibn Rafiq, Richard Han, Qin Lv, and Shivakant Mishra. Detection of cyberbullying incidents on the instagram social network. arXiv preprint arXiv:1503.03909, 2015.
[12] Mei Sze Choo. Cyberbullying on Facebook and psychosocial adjustment in Malaysian adolescents. PhD thesis, [Honolulu]:[University of Hawaii at Manoa],[December 2016], 2016.
[13] Robert R McCrae and Oliver P John. An introduction to the five-factor model and its applications. Journal of personality, 60(2):175–215, 1992.
[14] Oliver P John, Sanjay Srivastava, et al. The big five trait taxonomy: History, measurement, and theoretical perspectives. Handbook of personality: Theory and research, 2(1999):102–138, 1999.
[15] Qianyun Jiang, Fengqing Zhao, XiaochunXie, Xingchao Wang, Jia Nie, Li Lei, and Pengcheng Wang. Difficulties in emotion regulation and cyberbullying among chinese adolescents: a mediation model of loneliness and depression. Journal of interpersonal violence, 37(1-2):NP1105– NP1124, 2022.
[16] Qianjia Huang, Vivek Kumar Singh, and Pradeep Kumar Atrey. Cyber bullying detection using social and textual analysis. In Proceedings of the 3rd International Workshop on Socially-aware Multimedia, pages 3–6, 2014.
[17] Nazanin Alavi, TarasReshetukha, Eric Prost, Kristen Antoniak, Charmy Patel, Saad Sajid, and Dianne Groll. Relationship between bullying and suicidal behaviour in youth presenting to the emergency department. Journal of the Canadian Academy of Child and Adolescent Psychiatry, 26(2):70, 2017.
[18] A Saravanaraj, JI Sheeba, and S Pradeep Devaneyan. Automatic detection of cyberbullying from twitter. IRACST-International J. Comput. Sci. Inf. Technol. Secur, 6(2016):2249–9555, 2016.
[19] Karthik Dinakar, Roi Reichart, and Henry Lieberman. Modeling the detection of textual cyberbullying. In Proceedings of the International AAAI Conference on Web and Social Media, volume 5, pages 11–17, 2011.
[20] Steven L Salzberg. C4. 5: Programs for machine learning by j. ross quinlan. morgankaufmann publishers, inc., 1993, 1994.
[21] Ying Chen, Yilu Zhou, Sencun Zhu, and Heng Xu. Detecting offensive language in social media to protect adolescent online safety. In 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing, pages 71–80. IEEE, 2012.
[22] MaralDadvar, Rudolf Berend Trieschnigg, and Franciska MG de Jong. Expert knowledge for automatic detection of bullies in social networks. In 25th Benelux Conference on Artificial Intelligence, BNAIC 2013, pages 57–64. Delft University of Technology, 2013.
[23] MaralDadvar, FMG de Jong, Roeland Ordelman, and DolfTrieschnigg. Improved cyberbullying detection using gender information. In Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012). University of Ghent, 2012.
[24] MaralDadvar, DolfTrieschnigg, Roeland Ordelman, and Franciska de Jong. Improving cyberbullying detection with user context. In European Conference on Information Retrieval, pages 693–696. Springer, 2013.

[25] Claudia Peersman, Walter Daelemans, and Leona Van Vaerenbergh. Predicting age and gender in online social networks. In Proceedings of the 3rd international workshop on Search and mining usergenerated contents, pages 37–44, 2011.

[26] Anna Squicciarini, Sarah Rajtmajer, Y Liu, and Christopher Griffin. Identification and characterization of cyberbullying dynamics in an online social network. In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, pages 280–285, 2015.

[27] K Nalini and L Sheela. Classification of tweets using text classifier to detect cyber bullying. In Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2, pages 637–645. Springer, 2015.

[28] Vikas S Chavan and SS Shylaja. Machine learning approach for detection of cyber-aggressive comments by peers on social media network. In 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pages 2354–2358. IEEE, 2015.

[29] Rahul Ramesh Dalvi, Sudhanshu Baliram Chavan, and Aparna Halbe. Detecting a twitter cyberbullying using machine learning. In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pages 297–301. IEEE, 2020