As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 8(05):317-324, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0805046
Available online at: http://www.ijmtst.com/vol8issue05.html

Cyberbullying Detection Using Deep Learning Model

M.Gowtham Raj[1] | N.Madhu Balan[2] |T.Nitish[2] |P.Sibi Subhadeep[2] |M.Tharun[2]

[1]Assistant Professor Department of Information Technology, Adhiyamaan College of Engineering.
[2]Department of Information Technology, Adhiyamaan College of Engineering.
Corresponding Author Email Id: nitishthiyagu05@gmail.com

To Cite this Article
M.Gowtham Raj, N.Madhu Balan, T.Nitish, P.Sibi Subhadeep and M.Tharun[2]. Cyberbullying Detection Using Deep Learning Model. International Journal for Modern Trends in Science and Technology 2022, 8(05), pp. 317-324. https://doi.org/10.46501/IJMTST0805046

Article Info
Received: 07 April 2022; Accepted: 12 May 2022; Published: 15 May 2022.

ABSTRACT

Cyberbullying is bullying that takes place over digital devices like cell phones, computers, and tablets. Cyberbullying can occur through SMS, Text, and apps, or online in social media, forums, or gaming where people can view, participate in, or share content. Cyberbullying includes sending, posting, or sharing negative, harmful, false, or mean content about someone else. It can include sharing personal or private information about someone else causing embarrassment or humiliation. The content an individual shares online both their personal content as well as any negative, mean, or hurtful content – creates a kind of permanent public record of their views, activities, and behaviour. To avoid or detecting cyberbullying attacks, many existing approaches in the literature incorporate Machine Learning and Natural Language Processing text classification models without considering the sentence semantics. The main goal of this project is to overcome that issue. This project proposed a model LSTM - CNN architecture for detecting cyberbullying attacks and it used word2vec to train the custom of word embeddings. This model is used to classify tweets or comments as bullying or non-bullying based on the toxicity score. LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. A convolutional neural network (CNN) is a type of artificial neural network and it has a convolutional layer to extract information by a larger piece of text and by using this model LSTM- CNN achieve a higher accuracy in analysis, classification and detecting the cyberbullying attacks on posts and comments.

1. INTRODUCTION

1.1 Overview

The term social networking refers to the use of internet-based social media sites to stay connected with friends, family, colleagues, customers, or clients. Social networking can have a social purpose, a business purpose, or both, through sites like Facebook, Twitter, LinkedIn, and Instagram. Social networking is also a significant base for marketers seeking to engage customers. Facebook remains the largest and most popular social network, with 2.8 billion people using the platform on a monthly basis, as of Dec. 31, 2020. Instagram, Facebook Messenger, Twitter, and Pinterest are the next most popular, according to Statista. Social networking involves the development and maintenance of personal and business relationships using technology. This is done through the use of social networking sites, such as Facebook, Instagram, and Twitter. These sites allow people and corporations to connect with one another so they can develop relationships and so they can share information, ideas, and messages.

- Social networking is the use of internet-based social media platforms to stay connected with friends, family, or peers. The most popular social networking

sites in the U.S. include Facebook, Instagram, and Twitter. Marketers use social networking for increasing brand recognition and encouraging brand loyalty. Social media can help connect people and businesses and can help promote brand awareness. There are disadvantages related to social media, including the spread of misinformation and the high cost of using and maintaining social network profiles.

Social networking can have a big impact on the spread of misinformation and it can spread like wildfire. This became increasingly prevalent after 2012. Information starts as rumours, which spread faster than facts. One study found that misinformation is 70% more likely to be shared than factual information on Twitter.

Networking on social media can have just as much of a detrimental impact on companies. Criticism of a brand can spread very quickly on social media. This can create a virtual headache for a company's public relations (PR) department.

Although social networking itself is free, building and maintaining a company profile takes hours each week. Costs for those hours add up quickly. Businesses need many followers before a social media marketing campaign starts generating a positive return on investment (ROI). For example, submitting a post to 15 followers does not have the same effect as submitting the post to 15,000 followers. In simple words, wilful and repeated harm inflicted through the use of computers, cell phones, or other electronic devices is called cyberbullying.

Cyberbullying is mainly defined as intentional acts of antagonism carried out by a group or single person using electronic communication, many times or over time against victims who cannot easily defend themselves.

Cyberbullying is bullying with the use of digital technologies. It can take place on social media, messaging platforms, gaming platforms and mobile phones. It is repeated behaviour, aimed at scaring, angering or shaming those who are targeted. Examples include:

- spreading lies about or posting embarrassing photos of someone on social media. sending hurtful messages or threats via messaging platforms. impersonating someone and sending mean messages to others on their behalf. Face-to-face bullying and cyberbullying can often happen alongside each other. But cyberbullying leaves a digital footprint – a record that can prove useful and provide evidence to help stop the abuse.

For example, all friends joke around with each other, but sometimes it's hard to tell if someone is just having fun or trying to hurt you, especially online. Sometimes they'll laugh it off with a "just kidding," or "don't take it so seriously."

But if you feel hurt or think others are laughing at you instead of with you, then the joke has gone too far. If it continues even after you've asked the person to stop and you are still feeling upset about it, then this could be bullying. And when the bullying takes place online, it can result in unwanted attention from a wide range of people including strangers.

1.2 Problems Identified

Cyberbullying is the harassment that takes place in digital devices such as mobile phones, computers and tablets. The means used to harass victims are very diverse: text messages, applications, social media, forums or interactive games. It may be considered cyberbullying to send, post or share offensive, harmful, false, or cruel content about another person, in order to humiliate or embarrass them. In many cases these type of acts are illegal and therefore denounced. One of the things that complicates these type of situations that occur through the Internet, is the anonymity this environment allows. Since this facilitates cyberbullying can cover almost all areas of the victim's life, that is: educational environment, work, social or loving life. When the identity of the harasser is not known, even if the facts are reported, in many cases it is not enough to open an investigation, identify it and pay for the crime committed.

Peculiarities of cyberbullying

- **Persistence:** digital devices and the Internet facilitate immediate communication during 24 hours, so that victims do not have a minute of tranquility.
- **The permanence:** The information disseminated through the Internet is permanent and public domain, unless it is reported or eliminated (long and complex process). Three types of cyberbullying can be identified:

1. **Cyberbullying:**

When the harassment occurs between adults. That is, both the victim and the cyberbully are of legal age.

2. **Cyber sexual harassment:**

It's when the purpose of cyberbullying is sexual.

3. **Cyberbullying at school:**

The harassment occurs among underage. That is, both the victim and the cyberbully are children or young people. This type of harassment usually begins in the educational field.

Children and young people of school age are the potential targets of cyberbullying, so it is very important to educate them about the risks they face in social media and control its activity on them.

A. *Possible consequences of cyberbullying:*

These are the most serious consequences, derived from this type of harassment. The severity of the same may vary, depending on the profile of the victim, the type of harassment and the support received. They can manifest as:

- Depression
- Distrust towards anyone
- Isolation
- Alterations in sleep
- Eating disorders …
- *Physical:*

This type of consequences arise when the harassment extends to the physical world, and the victim suffers some type of aggression from his stalker. To avoid these types of consequences, it is advisable to never go to meetings requested by the harasser.

- *Sexual:*

When it comes to a sexual cyberbullying, the most common consequences are: sextortion or grooming. Whose objective in both cases is to force the victims to perform sexual acts without their consent and then spread it. In the case of grooming, the harasser is an adult, so he could be convicted of sexual abuse. In contrast in the case of sextortion, the stalker can be of any age. But in both cases you must go to the authorities and report the facts.

Effects of cyberbullying:

When bullying happens online it can feel as if you're being attacked everywhere, even inside your own home. It can seem like there's no escape. The effects can last a long time and affect a person in many ways:

- **Mentally** — feeling upset, embarrassed, stupid, even angry
- **Emotionally** — feeling ashamed or losing interest in the things you love
- **Physically** — tired (loss of sleep), or experiencing symptoms like stomach aches and headaches

All those who have a mobile, a tablet or any computer device and share any content on social media such as: Twitter, WhatsApp, Snapchat, Facebook, YouTube, etc., can be a victim of cyberbullying. The feeling of being laughed at or harassed by others, can prevent people from speaking up or trying to deal with the problem. In extreme cases, cyberbullying can even lead to people taking their own lives. Cyberbullying can affect us in many ways. So it must need to take some actions to avoid from cyberbullying so there is a crisis-specific immediate necessity to develop a proper detecting system to monitor the social media platforms and detecting the cyberbullying attacks.

1.3 Artificial Intelligence

Artificial Intelligence (AI) is the field of computer science dedicated to solving cognitive problems commonly associated with human intelligence, such as learning, problem solving, and pattern recognition. Artificial Intelligence, often abbreviated as "AI", may connote robotics or futuristic scenes, AI goes well beyond the automatons of science fiction, into the non-fiction of modern-day advanced computer science. Professor Pedro Domingo's, a prominent researcher in this field, describes "five tribes" of machine learning, comprised of symbolists, with origins in logic and philosophy; connectionists, stemming from neuroscience; revolutionaries, relating to evolutionary biology; Bayesians, engaged with statistics and probability; and analogizes with origins in psychology. Recently, advances in the efficiency of statistical computation have led to Bayesians being successful at furthering the field in a number of areas, under the name "machine learning". Similarly, advances in network computation have led to connectionists furthering a subfield under the name "deep learning". Machine learning (ML) and deep learning (DL) are both computer science fields derived from the discipline of Artificial Intelligence.

Broadly, these techniques are separated into "supervised" and "unsupervised" learning techniques, where "supervised" uses training data that includes the

desired output, and "unsupervised" uses training data without the desired output.

AI becomes "smarter" and learns faster with more data, and every day, businesses are generating this fuel for running machine learning and deep learning solutions, whether collected and extracted from a data warehouse like Amazon Redshift, ground-trothed through the power of "the crowd" with Mechanical Turk, or dynamically mined through Kinesis Streams. Further, with the advent of IoT, sensor technology exponentially adds to the amount of data to be analysed -- data from sources and places and objects and events that have previously been nearly untouched.

1.3.1 AI in everyday life

Below are some AI applications that you may not realize are AI-powered:

Online shopping and advertising

Artificial intelligence is widely used to provide personalized recommendations to people, based for example on their previous searches and purchases or other online behaviour. AI is hugely important in commerce: optimizing products, planning inventory, logistics etc.

Web search

Search engines learn from the vast input of data, provided by their users to provide relevant search results.

Digital personal assistants

Smartphones use AI to provide services that are as relevant and personalized as possible. Virtual assistants answering questions, providing recommendations and helping organize daily routines have become ubiquitous.

Machine translations

Language translation software, either based on written or spoken text, relies on artificial intelligence to provide and improve translations. This also applies to functions such as automated subtitling.

Smart homes, cities and infrastructure

Smart thermostats learn from our behaviour to save energy, while developers of smart cities hope to regulate traffic to improve connectivity and reduce traffic jams.

Cars

While self-driving vehicles are not yet standard, cars already use AI-powered safety functions. The EU has for example helped to fund VI-DAS, automated sensors that detect possible dangerous situations and accidents.

Navigation is largely AI-powered.

Cybersecurity

AI systems can help recognize and fight cyberattacks and other cyber threats based on the continuous input of data, recognizing patterns and backtracking the attacks.

Artificial intelligence against Covid-19

In the case of Covid-19, AI has been used in thermal imaging in airports and elsewhere. In medicine it can help recognize infection from computerized tomography lung scans. It has also been used to provide data to track the spread of the disease.

Fighting disinformation

Certain AI applications can detect fake news and disinformation by mining social media information, looking for words that are sensational or alarming and identifying which online sources are deemed authoritative.

1.3.2 Deep Learning

Deep Learning is a branch of machine learning that involves layering algorithms in an effort to gain greater understanding of the data. The algorithms are no longer limited to create an explainable set of relationships as would a more basic regression. Instead, deep learning relies on these layers of non-linear algorithms to create distributed representations that interact based on a series of factors. Given large sets of training data, deep learning algorithms begin to be able to identify the relationships between elements. These relationships may be between shapes, colors, words, and more. From this, the system can then be used to create predictions. Within machine learning and artificial intelligence, the power of deep learning stems from the system being able to identify more relationships than humans could practically code in software, or relationships that humans may not even be able to perceive. After sufficient training, this allows the network of algorithms to begin to make predictions or interpretations of very complex data.

1.4 Scope of the project

In the present information age, there has been an increase in the use of social media platforms, as a result, there have been many cyberbullying instances occur in social media. Here our method is used to classify tweets as bullying or non-bullying based on the toxicity score.

The features of this system include visualization of statistics, search methods to find bullying attacks, and automatic generation of report. So if the person posting some hateful comments or post it detect and block that

person id so the people relieve the feeling of being laughed at or harassed by others, can prevent people from speaking up or trying to deal with the problem.

1.5 Objective

The key objective of AI present is to develop a robust system with help of using deep learning methods, the system monitor social platforms comments and post using LSTM-CNN algorithm for detecting the cyberbullying.

2. SYSTEM ANALYSIS

2.1 Existing System

In existing system various types of techniques have been applied to detect cyberbullying. Previous works on computational studies of bullying have shown that the natural language processing (NLP) and machine learning (ML) are powerful tools to study bullying.

A. Logistic Regression and Random Forest: Logistic Regression was used for the detection of hyperaggressive comments on social media. Logistic Regression was applied to the dataset after feature extraction using TFIDF and N-gram. It resulted in an AUC score of 86.92 %. Random Forest classification was used to help in identifying cyberbullying on one of the affected social media, Twitter. The proposed method was applied to 50 groups of tweets using the training/ test split. Random Forest resulted in the highest F1-Score which is 0.90.

B. XGBoost: Feature extraction and detection of cyberbullying in Twitter messages with the help of Natural Language Processing tools and different Machine learning algorithms were analyzed and experimented with using extreme Gradient Boosting (XGBoost) model on word2vec features outperformed all Support Vector Machine (SVM), Logistic Regression, Random Forest for the classification of tweets into bullying and non-bullying with an F1 score of 70% and a precision of 68%.

C. Support vector machine: The author then tested the you tube dataset with 3 different Machine Learning Models: a Naive Bayes Classifier, Decision Tree Classifier(C4.5), and a Support Vector Machine(SVM) with a Linear Kernel. It was observed that clustering results for the hate posts turned out to have a lower precision in the you tube dataset when compared to the Form Spring tests, as textual analysis and syntactical features perform differently on both sides. When this hybrid approach was applied to the Twitter dataset, it resulted in a weak recall and F1 Score of the accuracy of the SVM and Naive Bayes that is 71.25% and 52.70%.

Disadvantages:

- Process of reporting such cases is long, tedious job. Difficult to track. Most of the cyberbullying cases go unreported. Low accuracy. Time consuming process. Problem is not automatically detected and not promptly report bullying message. Response time is slow. Basic features and common classifier accuracy is low. Data are manually labelled using online services or custom applications. Usually data limited only to a small percentage.

2.2 Proposed system:

Cyberbullying is mainly defined as intentional acts of antagonism carried out by a group or single person using electronic communication, many times or over time against victims who cannot easily defend themselves. Cyberbullying is bullying with the use of digital technologies. It can take place on social media, messaging platforms, gaming platforms and mobile phones. It is repeated behaviour, aimed at scaring, angering or shaming those who are targeted. This project proposed a model LSTM-CNN architecture is used for detecting cyberbullying attacks. Thus, this project have designed a method of automatically detecting the cyberbullying attack cases. Identifies the messages or comments or posts which the LSTM-CNN model predicts as offensive or negative then it blocks that person id, then the admin can create automated reports (the automated report generator creates a simple report for probable accusers) and send them to the concerned authorities or crime department without any problem.

Advantages:

- It successfullyclassifies the tweets in various classes. Auto report generator generates a simple report for probable accusers. Several analytics and report can be sent to the crime department. Accuracy is high. Foul language on any given page, removes it, and can highlight words as well, also censor NSFW images on the page. This method detects the offensive post or messages it block that user id. The "filtered content" is displayed at back to the page, in such a way preventing the display of explicit content.

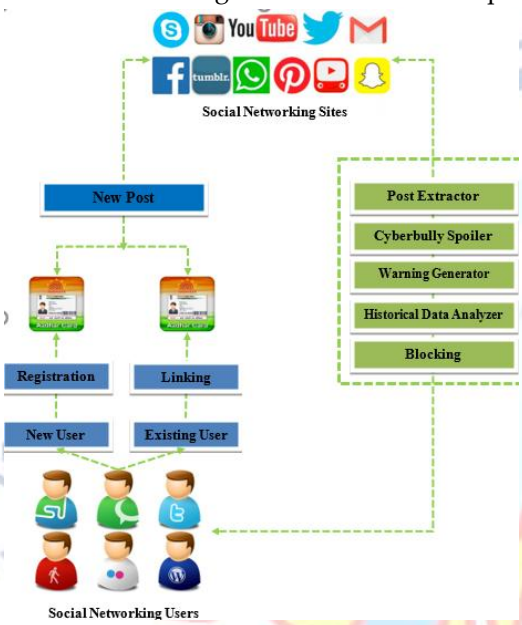
An automatically generate a report for each incident is also provided.

Architecture Diagram

3. MODULES

1. Social Networking Web App

Build a social networking service is an online platform



which people use to build social networks or social relationships with other people who share similar personal or career interests, activities, backgrounds or real-life connections. Social networking services vary in format and the number of features.The classification model has been exposed as a REST API which was consumed by a Web application built using Python's Flask framework. The main features include an Admin dashboard for visualization of cyberbullying activities, an option to search tweets, and automatic generation and emailing of reports of cyberbullying activity.

2. Aadhar User Account Management

2.1. New User

Create user account with Aadhaar number.

2.2. Existing User

The existing users of Facebook will also have to upload a scanned copy of their Aadhar Card. If they fail to do so, their profile will be suspended within the next 15 days.

3. **Cyberbullying Analysis API:** In this module we developed the API for cyberbullying analytics on chat or post user data. It focuses on keywords and analyzes chat or post according to a two-pole scale (positive and negative).

4. Cyberbullying Data Set Annotation

4.1. Training Phase: We used cyberbullying data from Kaggle. The dataset in consisted of two labels, positive and negative, while was composed of three labels of positive, neutral, and negative. Furthermore, the dataset in was composed of five labels of positive, somewhat positive, neutral, somewhat negative, and negative.

4.2. Testing Phase: In this module, user login and give comments about the post and also chat with the other social networking users.

5. **Pre-processing:** The preprocessing step is essential in cyberbullying detection. It consists of both cleaning of texts (e.g., removal of stop words and punctuation marks), as well as spam content removal. In the proposed model, it has been applied to remove and clean unwanted noise in text detection. For example, stop words, special characters, and repeated words were removed. Then, the stemming for the remaining words to their original roots has been applied as a result of this preprocessing, and the dataset containing clean tweets is produced for the proposed model to be run and predicted.

6. 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. Therefore, TF-IDF is one of the most commonly used feature extraction techniques in text detection. 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. 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 vector word. The Skip-gram model is based on a corpus of terms w and meaning. The aim is to increase the likelihood of:

$$\underset{\theta}{\operatorname{argmax}} \prod_{w \in T} [\prod_{c \in C} p(c \mid w; \theta)],$$

where T refers to text, and _ is a parameter of p (c |w; 0). Figure 4 illustrates the Word2Vec model architecture, where 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. The Word2Vec technique implements both training models. The basic idea behind the two training models is that either a word is utilized to predict the context of it or the other way around—to use the context to predict a current word. Utilizing TF-IDF is weighted by its relative frequency instead of merely counting the words, which would overemphasize frequent words. The TF-IDF features notify the model if a word appears more often in a statement than the entire text corpus does typically. Prior work has found TF-IDF features useful for cyberbullying detection in SM. As with BOW, the TF-IDF vocabulary is constructed during model training and then reused for test prediction. Both BOW and TF-IDF are considered to be simple, proven methods for classifying text. In Equation, the mathematical representation by TF-IDF of the weight of a term in a document is given.

$$W(d,t) = TF(d,t) * \log\left(\frac{N}{df(t)}\right),$$

In this case, N is the number of documents and d f(t) is the number of documents in the corpus
containing the word t. In Equation above, the first term enhances the recall, while the second term enhances the word embedding accuracy.

7. **LSTM CNN Classification:** For cyberbullying detection in online social media sites, four Deep Neural Network (DNN) based models (i.e., BLSTM, LSTM, CNN, and Attention-based BLSTM) were developed and the proposed model systematically detected cases of cyberbullying on numerous Social media platform.The final model architecture consisted of an initial LSTM layer that receives word embeddings for each token in the comment text as input. The embeddings from the embedding layer were preloaded and the layer wasn't trainable. The intuition is that LSTM's output tokens will store information of previous tokens along with the initial token; In other words, the LSTM layer is generating a new encoding for the original input. The output of the LSTM layer, which has 128 cells, is then fed into a convolution layer, having 64 cells and a filter size of 3 which we expect will extract local features. Finally, the convolution layer's output will be pooled to a smaller dimension and ultimately outputted using the sigmoid activation. The result of the model is an aggregated toxicity percentage (score) of the text input, with all the labels contributing equally to the final toxicity percentage.
8. **Prediction:** In this module, Finally, we used matrix factorization to predict cyberbullying words or text and block the post or the user.
9. **Admin dashboard: :** The admin dashboard features a graphical visualization of the results generated by the model. The graphs displayed are, namely, monthly and weekly distribution of cyberbullying activities, sentiment analysis distribution of bullying and non-bullying reports, website-wise distribution of cyberbullying instances, list of recent instances of cyberbullying, and worldwide distribution of cyberbullying instances.

10. **Report generation:** A report is generated to document the incident and can be sent to the authorities for further actions. It is auto-generated using Python docxtpl in .docx format by the application and consists of the following details:

- Name, role, and signature of the admin
- Date and time of the incident
- Description of the incident, containing the URL to the tweet
- Name of the people involved Follow up action

Upon generation, the report can be auto mailed to the pertinent authorities. These details can also be edited by the admin before mailing.

11. **Performance Analysis:** In the module experiments were performed on a dataset obtained by extracting product reviews from Amazon.com. Considering reviews of one product at a time sentiment of the reviews were classified into four categories namely True

positive, False positive, True negative, False negative.

4. CONCLUSION:

Although the virtual revolution and the upward thrust of social media enabled incredible advances in verbal exchange structures and social interactions, a much wider proliferation of dangerous conduct known as bullying has additionally emerged. This article gives a novel framework of BullyNet to become aware of bully customers from the Twitter social network. We done giant studies on mining SNs for higher know-how of the relationships among customers in social media, to construct an SN primarily based totally on bullying tendencies. We discovered that with the aid of using building conversations primarily based totally at the context in addition to content, we may want to effectively become aware of the feelings and the conduct at the back of bullying. In our experimental study, the assessment of our proposed centrality measures to come across bullies from SN, and we performed around 80% accuracy with 81% precision in figuring out bullies for diverse cases. There are nonetheless numerous open questions deserving further investigation. First,our technique makes a specialty of extracting emotions and conduct from texts and emojis in tweets. However,it'd be exciting to analyze photographs and videos, for the reason that many customers use them to bully others. Second, it does now no longer distinguish among bully and competitive customers. Devising new algorithms or strategies to differentiate bullies from aggressors could show vital in higher identification of cyberbullies. Another subject matter of hobby could be to study the connection among communication graph dynamics and geographic vicinity and the way those dynamics are affected with the aid of using the geographic dispersion of the customers? Are the proximity growth the bullying conduct?

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] L. Cheng, J. Li, Y. N. Silva, D. L. Hall, and H. Liu, "XBully: Cyberbullying detection within a multi-modal context," in Proc. 12th ACM Int. Conf. Web Search Data Mining, Jan. 2019, pp. 339–347.

[2] C. Van Hee et al., "Automatic Detection of Cyberbullying in Social Media Text." 2018.

[3] M. Rezvan, S. Shekarpour, L. Balasuriya, K. Thirunarayan, V. Shalin, and A. Sheth, "A Quality Type-aware Annotated Corpus and Lexicon for Harassment Research," pp. 33–36, 2018.

[4] A. H. Alduailej and M. B. Khan, "The challenge of cyberbullying and its automatic detection in Arabic text," 2017 Int. Conf. Comput. Appl. ICCA 2017, pp. 389–394, 2017.

[5] A. Power, A. Keane, B. Nolan, and B. O. Neill, "A lexical database for public textual cyberbullying detection," Rev. Lenguas Para Fines Específicos, vol. 2, pp. 157–186, 2017.

[6] M. Drahošová and P. Balco, "ScienceDirect The analysis of advantages and disadvantages of use of social media the analysis of advantages and disadvantages of use of social media in European Union in European Union," Procedia Comput. Sci., vol. 109, pp. 1005–1009, 2017.

[7] R. Zhao, A. Zhou, and K. Mao, "Automatic detection of cyberbullying on social networks based on bullying features," in Proc. 17th Int. Conf. Distribution. Computer. Network, Jan. 2016, pp. 1–6.

[8] V. K. Singh, Q. Huang, and P. K. Atrey, "Cyberbullying detection using probabilistic socio-textual information fusion," in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), Aug. 2016, pp. 884–887.

[9] A. Squicciarini, S. Rajtmajer, Y. Liu, and C. Griffin, "Identification and characterization of cyberbullying dynamics in an online social network," in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining, Aug. 2015, pp. 280–285.

[10] 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," Logic J. IGPL, vol. 24, no. 1, pp. 42–53, 2015.