International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Latent Semantic Analysis Based Sentimental Analysis of Tweets in Social Media for the Classification of Cyberbullying Text

[1]D. C. Joy Winnie Wise, [2]S. Ambareesh, [3]Ramesh Babu P., [4]D. Sugumar, [5]John Philip Bhimavarapu, [6]A. Senthil Kumar

Submitted: 04/10/2023 **Revised:** 22/11/2023 **Accepted:** 02/12/2023

Abstract: With wide spread of mobile technology, cyberbullying has developed as a substantial problem, particularly among adolescents. This is especially true in the case of adolescents. The fact that some people have chosen to end their own lives by committing suicide has also helped increase awareness of the issue among the broader population. Various methods are adopted to reduce the suicides and in broader sense, todays online media is highly prone to bullying that is termed as cyber bullying. Methods are adopted to detect the cyberbullying text, however most of them lacks clarity in detecting the accurate cyber bullying tweets. In this paper, Latent Semantic Analysis (LSA) based sentimental analysis of tweets in social media for the classification of cyberbullying text. The study uses LSA that helps in classifying the texts and help the user to post their opinions in social media without any online abuse. The simulation is conducted to test the efficacy of the classification model and the results show that the proposed method achieves higher rate of accuracy than other existing methods.

Keywords: *Latent semantic analysis, sentimental analysis, tweets, cyberbullying text.*

1. Introduction

In recent times, it is obvious that having access to technologies that are able to automatically spot potential behaviors associated with cyberbullying can be of tremendous assistance in preventing potentially harmful scenarios for the victim [1]. This can be of great assistance because it can automatically spot potential behaviors associated with cyberbullying [2]. In spite of the fact that cyberbullying has recently attracted a substantial amount of interest from a social standpoint, very few computer research on the subject have been carried out up to this time [3].

Research in psychology and sociology provide useful insights that can be used in the construction of models that can recognize instances of bullying. These insights can be utilized in the building of models [4]. There are a number of key ways in which traditional forms of bullying can be differentiated from cyberbullying. These differences can be seen when comparing cyberbullying to more traditional forms of bullying [5]. To begin, there is no past relationship between the bully and the person who is being bullied [6]. This means that there is no history between the two parties [7]. It is common for victims of harassment to be unaware of the identity of the person who is harassing them; as a result, it is substantially more challenging for these victims to face the person who is harassing them [8].

Sometimes, the bully will use aliases so that he does not have to come into direct touch with the victim, which gives him the opportunity to behave in a more hostile manner [9]. Sadly, victims do not always succeed in their efforts to have the aggressor behave in a less hostile manner or to calm him down [10]. The lack of both physical borders and temporal limits is also extremely significant, as the victim privacy can be invaded at any time, regardless of the circumstances. This is extremely crucial since the victim private can be breached [11].

In the context of the problem of cyberbullying, it is vital to conduct a re-evaluation of the significance of the component of repetition because of the one-of-a-kind qualities that are inherent to digital communication [12-15]. The research offers a method for the automatic identification of evidence of bullies, such as exchanges that take place online and entail the use of abusive

[1]*Professor, Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, Tamil Nadu, India.*
[2]*Associate Professor, School of Computer Science and Engineering, Jain University, Bangalore, Karnataka, India.*
[3]*Associate Professor, Department of Computer Science, College of Engineering and Technology, Wollega University, Nekemte, Oromia Region, Ethiopia.*
[4]*Associate Professor, Electronics and Communication Engineering, Karunya Institute of Technology and Sciences (Deemed to be University), Coimbatore, Tamil Nadu, India.*
[5]*Assistant Professor, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.*
[6]*Professor, Artificial Intelligence Department, Sri Vishnu Engineering College for Women, Bhimawaram, Andhra Pradesh, India.*
[1]*drdcjoywinniewise@gmail.com,* [2]*ambihce@gmail.com,*
[3]*drprb2009@gmail.com,* [4]*sugumar@karunya.edu,*
[5]*johnphilip@kluniversity.in,* [6]*drsenthilkumar.ai@svecw.edu.in*

language, which has the potential to lead to instances of cyberbullying. During the phases of data preparation and detection, the research will demonstrate that the most accurate results can be obtained by integrating natural language processing (NLP) approaches with unsupervised machine learning algorithms.

2. Related Works

In this section, an in-depth discussion of prior research on mining Twitter data for sentiment analysis and opinions. The research was conducted by the authors of this article. Others were responsible for carrying out the research. In order to identify contributions and potential recommendations for the future, a comprehensive analysis of the approaches that have been offered in the existing body of research will be conducted. A particular emphasis will be placed on Twitter data and criminal activity conducted online.

The authors of [16] published a statistical investigation of the employment of Twitter social sentiment sensors in the identification of cyber-attacks based on the L1 regularization regression technique. Their findings were published in the context of identifying cyber-attacks.

The authors of [17] provided a time-series topic identification approach in order to infer possible themes of relevance over the course of time. By analyzing data from Twitter and applying this terminology, it was possible to forecast a rise in Chicago overall crime rate. This was done so that the authors could draw conclusions about the subjects. In place of depending on word dictionaries, the authors developed a dynamic vocabulary that is capable of recognizing emerging patterns in the problem space. The employment of content-based features led to an improvement in the overall prediction performance, as indicated by the findings of the study.

Sharma et al. [18] introduced a Sentiment Reasoning approach for conducting sentiment analysis on Twitter data in order to identify instances of cybercrime and potential hazards to cybersecurity. This was done in order to improve the effectiveness of the system. The author is of the opinion that countries in Asia, in comparison to other countries in the European Union, are more susceptible to the negative impacts of cyberattacks.

Founta et al. [19] presented a DL-based architecture with the intention of recognizing a wide variety of inappropriate online activities on Twitter. This was done specifically for Twitter. The proposed method was successful in producing outstanding results in terms of detection rate while simultaneously increasing AUC from 92% to 98%.

Al-Smadi et al. [20] conducted a sentiment analysis of reviews written in Arabic for hotels by utilizing a Recurrent Neutral Network (RNN) and Support Vector Machines (SVM). Research could potentially benefit from employing methodologies from the field of deep learning (DL). During both the training and the testing phases of the model, the DRNN demonstrated a much faster execution time compared to the other neural networks.

The method in [21] applied a DL technique that was based on a CNN, quite similar to the one that was used in the earlier study, in order to identify instances of cyberbullying that were found in Instagram postings. This was done in order to identify instances of cyberbullying that were found in Instagram postings. Following the discovery of bullying terms, these sentences are put through an additional round of in-depth analysis using the N.B. classifier, which ultimately leads to the successful identification of potential cyberbullying threats.

3. Proposed Method

In this study, the research presents a method for identifying instances of cyberbullying in social media that goes beyond relying solely on emotional analysis by also taking into account syntactic, semantic, and sarcastic aspects of the text that is in question. This method was developed as part of a larger investigation that was conducted by the authors of this study. As a result of our investigation into the subject, the research came up with this methodology as illustrated in Figure 1.

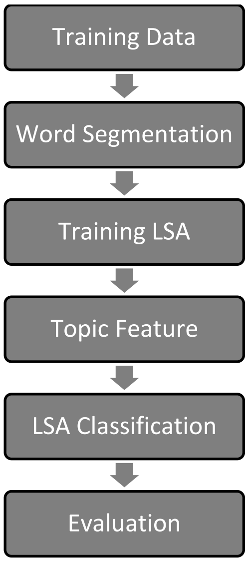

Fig 1: Proposed Method

In order to achieve this, the research will initially turn to more conventional methods of opinion mining, such as the contextual mining of text. The study is able to find and extract the relevant subjective information from the source material, and as a consequence, the research are able to get a more in-depth grasp of the author feeling, perspective, or opinion in relation to the subject at hand. The research next proceed to provide a selection of social aspects that have the capacity to significantly impact and drive efforts made to uncover instances of cyberbullying. These elements have been selected because they have the potential to significantly influence and drive these efforts.

Following an analysis of the various types of systems that are currently in use, each of these attributes has been classified in this manner so that it is possible to apply it to the text in order to generate a one-of-a-kind identifier for it, and this classification was arrived at following a literature review. When it comes to solving problems that include pattern recognition and classification, the efficiency of the algorithms relies heavily on the process of picking the features to be used in the analysis.

Emotional qualities are meant to provide an all-encompassing analysis of the overarching feeling (whether it be positive or negative) that a piece of textual content is trying to express to the reader. Our sentiment score algorithm has been trained to aim for an agreement rate of 80–85%, which is the normal rate at which human analysts tend to agree with one another. The findings of the research that has been presented led to the establishment of this objective.

When using sarcastic features, the study makes an effort to create incongruity with the surrounding context. Incongruence is the condition that arises when the actions of a person and the assertions that they make do not correspond with one another. It possible that only about half of the objects in a text are placed in what the research might refer to as their anticipated settings, while the other half might be completely incongruous with those settings. This could happen if, for example, the author forgot to move an object before placing it in the text. The fundamental character of the sarcastic sentence will not be picked up by sentiment analysis since it is being secret by the stuff that is surrounding it. Because of this, it can be very difficult to determine whether or not a comment constitutes cyberbullying because the genuine nature of the comment cannot be determined. When determining whether or not a piece of information is meant to be sardonic, the research take into consideration a variety of pragmatic characteristics, such as the use of emoticons and mentions, among other things.

The research not only monitor, evaluate, and assign a density to the number of such negative terms or insults that are present in a single sentence, but the research also take into consideration the syntactic qualities that the research have detected in the lists of insults. This is done in addition to monitoring and assigning a density to the number of such negative terms or insults that are present in a single sentence. Through the application of metrics such as density range, the research arrived at the conclusion that the entire sentence is bad, and the research stand by this assessment. The use of all capital letters in online hate speech is taken into consideration as a syntactic aspect because of the fact that it is commonly understood as a form of screaming or attacking. In a similar vein, while deriving syntactic features, both the use of special characters and the patterns generated by them are taken into consideration. This is because special characters can form patterns.

The semantic qualities of each word on its own or in combination with the properties of the other word to discover the lexical relationship between two words in a language. Examining the semantic properties of a word can help one arrive at a conclusion regarding the meaning of the word.

In this article, the research has looked for trigrams and bigrams that can be found in a number of different textual sources. It is common practice for people who engage in acts of online harassment to take into account not only the target denial of the statement but also the mapping of various pronouns that can be used to refer to that target, either implicitly or overtly. This is done in order to circumvent the target ability to challenge the statement. This is done in order to increase the possibility that the target would view the harassment as being aimed at them. The target is the person who is the subject of the harassment.

The social traits of a bully or a bullied person are the qualities of the bully or the bullied person social behavior in general. It will not be adequate to discover the nature of the text that is being shared by merely reading the post that is being shared. The research looked for characteristics that are shared among bullies and discovered that there are a lot of them. The research have given some consideration to the question of whether or not hate speech can be utilized in a manner that specifically identifies the victim. When compiling information on the history of the post, the research also takes into account the previous interactions that have taken place between the bully and the victim. It is possible for a person author profile to reveal whether or not they have engaged in disruptive activity on other social networking networks.

With the help of transformers, which are one of the essential components, the research were able to construct an approach that identifies instances of cyberbullying. Transformers are beneficial for resolving a wide number of challenges that are linked with natural language processing. This is mostly due to the fact that they are able to take sequential data as input. Translation and text summarization are two examples of the difficulties encountered in this process. The BERT is a relatively recent breakthrough that has made great steps toward enhancing the work of natural language processing. These improvements can be attributed to the fact that the BERT has been implemented.

The model has been pre-trained on unlabeled texts written in both the left and right directions so that it can determine the meaning of texts written in either direction. This enables the model to read texts written in either direction. The model can be used in either direction.

Due to the fact that it uses semi-supervised learning, BERT is a viable model for jobs involving natural language processing. If the research adapt the LSA design to include a layer that is suitable to the job at hand, the research may be able to use this model to develop a cutting-edge machine learning model that can be applied to a specific problem. If the research do this, the research will hopefully be successful. The LSA is a paradigm that investigates the meaning of words by taking into account both the left and the right contexts in which they are used. Because LSA is a bidirectional model, this is the result.

If the research look at the context of the word bat in the first sentence, beginning with the leftmost section of the sentence and continuing all the way through to the end of the sentence, it is obvious that the term refers to the animal that is active at night. In contrast, the word bat in the second sentence brings to mind the cricket bat given the context that has been presented up to that point. Because of this, it can be challenging for a machine to establish the accurate meaning of a word if it does not have access to both the context in which the term is used and the intention of the person who is speaking it.

Pre-processing

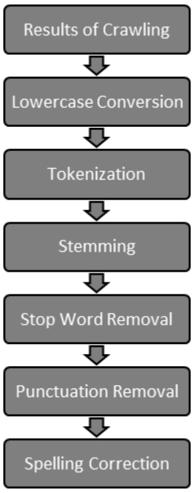

Fig 2: Preprocessing

Figure 2 provides an illustration of the steps that are involved in the preparation of data. The preprocessing method included the following six steps: converting all uppercase letters to lowercase, tokenizing each word, stemming each word, removing stopwords, removing punctuation, and correcting any spelling errors that may have occurred.

At the beginning of the process, which is referred to as the Convert into Lowercase phase, the content of the customer review is altered so that it uses only lowercase letters. The next stage is called tokenization, and it entails splitting up the text into its component words. This stage follows the stage before tokenization. The third step is termed summarizing, and it involves condensing the information that was obtained from the feedback provided by prior clients. In the fourth phase, which is known as Remove Punctuation, any punctuation that may have been included in the customer review will be eliminated. This step is referred to as Remove All Punctuation. In the third phase, which is known as spelling correction, inaccurate customer reviews are changed to reflect the right spelling.

Natural language processing, which is sometimes referred to by its acronym NLP, was used as a tool to assist in the labeling of various aspects of speech (POS tagging). When preprocessing and labeling at the point of sale are done together, the need for inaccurate language is cut down significantly. In order to properly tag each list of words later on, the research began by preprocessing each word list with POS tagging. This allowed us to tag each list of words in the appropriate manner. In order to determine the allocation of word parts, the POS tagger that is included within the Natural Language Toolkit Library was applied (NLTK). The research assigned each sentence using POS tagging, and then the research determined the degree of similarity between each sentence single noun and its infinitive by calculating the degree of similarity between the two.

LSA Classification

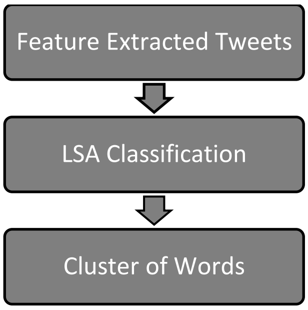

Fig 3: Classification

The procedure for discovering cryptic themes inside a document that also has a term list is depicted in Figure 3, which provides an example of the method. The research assembled a vocabulary of terminology by compiling all of the available product reviews and categorizing them according to certain terms. Every single document had an evaluation of a company, service, or product labeled Identity Document, which was located in the file header. On the label, the serial number was altered in such a way that it would correlate with the count that was determined through evaluation.

The research are going to use WordNet, which is a database of English words as well as the semantic and syntactic relationships between them, in order to improve the quality of the output the research produce at the end of the process. WordNet ability to recognize words that may be interpreted in a number of different ways was one of the most crucial roles that it played. The findings would then be applied to the production of additional data sets.

After the annotator had finished applying labels to the data that would later be used for training, they would store the entire thing as a corpus within the data set. This would be done after the labels had been applied to the data. In the course of our inquiry, the research took use of real-world data sets, which are collections of information that have been assembled from the perspectives of actual customers regarding a variety of products.

The third step involved compiling an enlarged term list for WordNet by categorizing words into different groups according to the likelihood that they have similar semantic associations. The research were able to generate hidden topics by using the expectation-maximization (EStep and M-Step) algorithm that was provided by the LSA. This approach allowed us to determine the relevance of the document sentences and the similarities that existed between the topics.

Because of its flexibility and the fact that it enables the integration of the document training corpus from both the E-step and the M-step algorithms, the latent semantic analysis (LSA) technique was selected as the method to apply. Because it addresses natural language processing from the angle of machine learning, the LSA is in a position to generate improved hidden themes.

As an example of a LSA algorithm, please take into consideration the following: The researcher is responsible for determining the initial values of the probability parameters for each subject area. The relevant values for these parameters are P(z), P (d|z), and P (w|z).

Both of these probabilities are expressed as percentages. The formula that is displayed illustrates the procedure that was followed in order to determine the total number of words:

$$P\left(d_i, w_j\right) = \sum_{k=1}^{K} a \cdot a \cdot c \cdot d \qquad (1)$$

$a=P(Z_k)$ (2)

$c=P(d_i|Z_k)$ (3)

$d=P(w_j|Z_k)$ (4)

P(d|z) - probability that a document will contain the topic, and

P(w|z) - probability that words will randomly appear in the topic.

The study employ the methodologies known as the expectation step and the maximization step, respectively, in order to ascertain the probability value of the words contained inside each parameter. The extent to which the information included in the text is likely to conform to the formula that is shown:

$$P\left(Z_k \middle| d_i, w_j\right) = \frac{P\left(w_j \mid Z_k\right) P\left(z_k \mid d_i\right)}{\sum_{l=1}^{K} P\left(w_j \mid Z_l\right) P\left(z_l \mid d_i\right)} \qquad (5)$$

The final step that needs to be figured out is the one that maximizes the benefit (M-step). In order to discover the new value, the research will utilize M-step, as stated in below, to iteratively solve for the parameters of the document:

$$P\left(w_j \middle| Z_k\right) = \frac{\sum_{l=1}^{K} n\left(d_i \mid w_l\right) P\left(z_k \mid d_i, w_j\right)}{\sum_{m=1}^{M} \sum_{i=1}^{N} n\left(d_i \mid w_m\right) P\left(z_k \mid d_i, w_m\right)} \qquad (6)$$

$$P\left(Z_k \middle| d_i\right) = \frac{\sum_{j=1}^{K} n\left(d_i \mid w_j\right) P\left(z_k \mid d_i, w_j\right)}{n\left(d_i\right)} \qquad (7)$$

The result is the hidden-topic-generating word probability, which can be calculated by combining the above equations.

4. Results and Discussions

In addition to this, the research also carried out an investigation on Twitter. In order to evaluate how well the proposed model worked when it came to detecting data extracted from an actual stream of tweets on Twitter, the research made use of a GHSOM that had been trained on it in the past. The research gathered a total of one thousand tweets during the summer of 2015 and did not put them through any form of filter other than the one that tested for language (English).

Table 1: Dataset

Parameter	Value
Character limit	140
Unique hashtags	250
Retweets	210
Unique users	20
Tweets	1232

As a result of the platform character limit of 140, users of Twitter are obligated to submit only the information that is most pertinent to their posts. This restriction presents a significant challenge because the techniques that are most frequently used for text analysis are most efficient when applied to lengthier pieces of writing. This makes the restriction a significant obstacle.

Twitter users are limited to publishing a maximum of 140 characters at once, users commonly resort to utilizing informal language, acronyms and slang sentences. This is because Twitter only enables users to publish a maximum of 140 characters at once. The research tested the capabilities of our unsupervised model by applying it to the Twitter dataset which is shown in Table 1.

The research wanted to explore how well our model would work on this dataset. It is possible that the poor recall and F1 score results can be explained by the fact that the dataset was originally used for sentiment analysis rather than for cyberbullying.

The percentage of tweets in the real-world test sample that contain unevenly distributed cyberbullying ranges between 4 and 10% of the whole dataset. Figure 4 shows the results of accuracy between the LSA and existing state-of-art cyberbullying texts. The simulation shows that the LSA has a higher rate of accuracy in classifying the bullying tweets than the other methods.

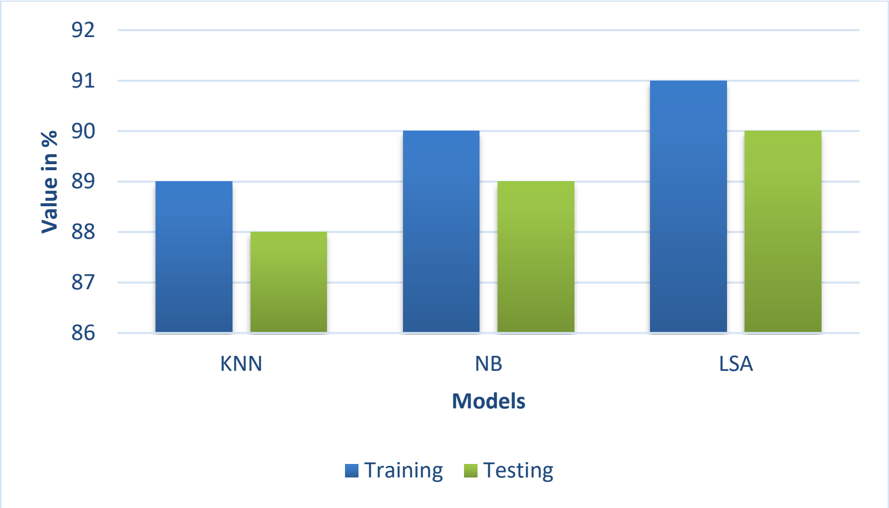

Fig 4: Accuracy

Figure 5 shows the results of precision between the LSA and existing state-of-art cyberbullying texts. The simulation shows that the LSA has a higher rate of classification precision in classifying the bullying tweets than the other methods.

Fig 5: Precision

Figure 6 shows the results of recall between the LSA and existing state-of-art cyberbullying texts. The simulation shows that the LSA has a higher rate of classification recall in classifying the bullying tweets than the other methods.

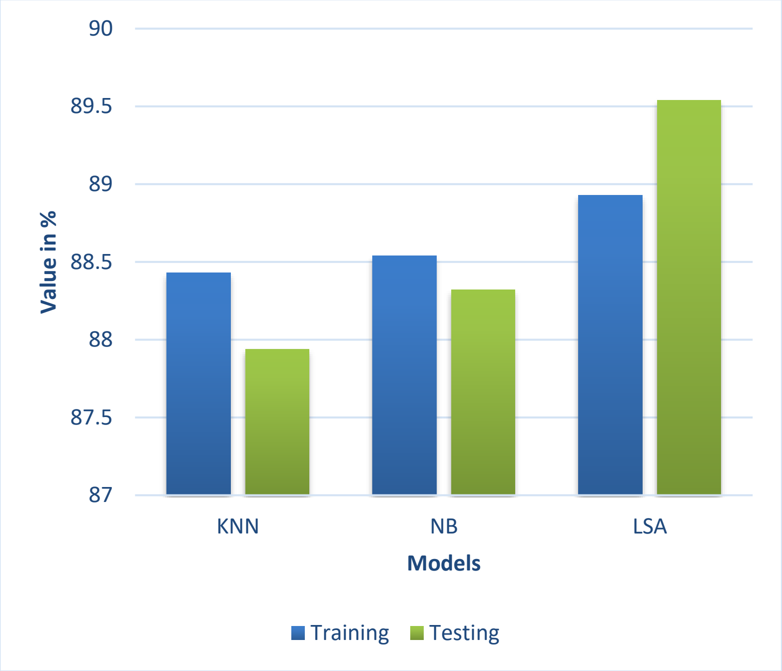

Fig 6: Recall

Figure 7 shows the results of F-Measure between the LSA and existing state-of-art cyberbullying texts. The simulation shows that the LSA has a higher rate of classification F-Measure in classifying the bullying tweets than the other methods.

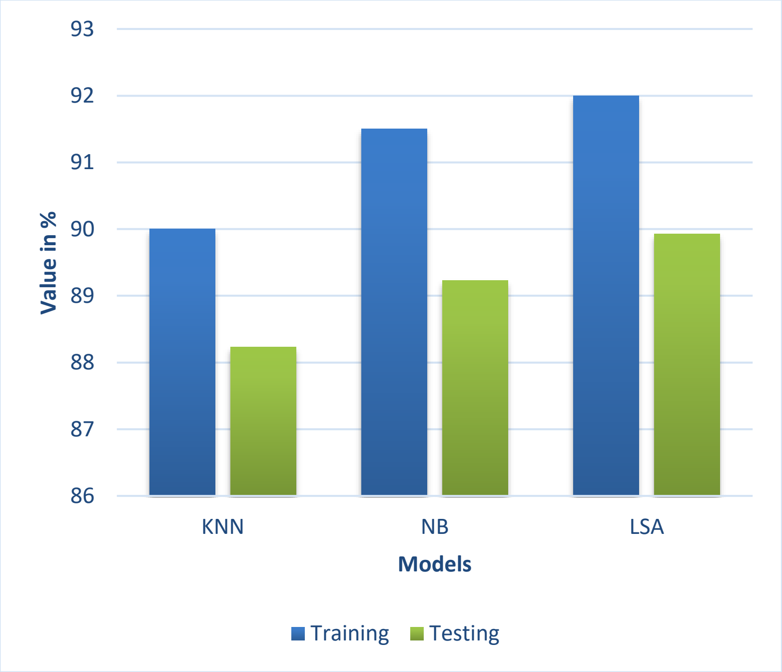

Fig 7: F-Measure

Figure 8 and 9 shows the results of MAPE and MAE between the LSA and existing state-of-art cyberbullying texts. The simulation shows that the LSA has a reduced rate of MAPE and MAE in classifying the bullying tweets than the other methods.

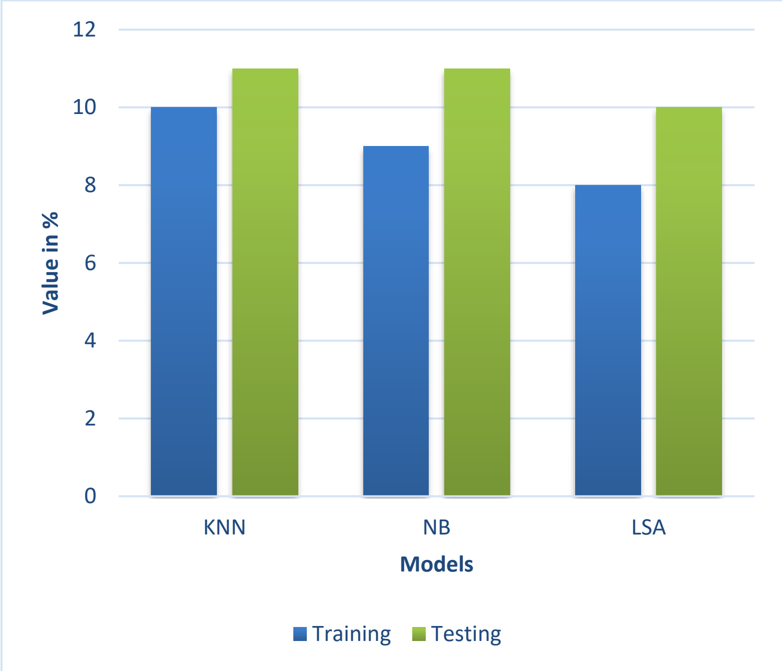

Fig 8: MAPE

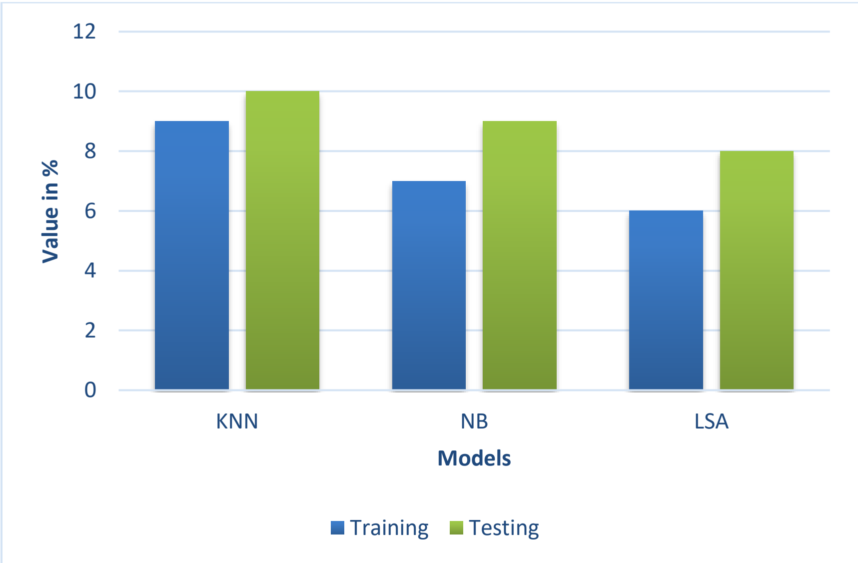

Fig 9: MAE

5. Conclusions

The research provides a LSA model for the detection of cyberbullying that is based on the method of sentiment analysis. This model was developed by us. LSA constitutes cyberbullying is assumed to be a highly negative message for the purposes of this investigation. The research made the decision to go with an unsupervised technique so that the research would not have to manually label the vast datasets and so that the research would not have to make any a priori assumptions about the classes that are most likely to be there. LSA allowed to save a significant amount of time.

Additionally, the research will go over the next steps that the research plan to take in order to further develop this model. According to the findings, our method performs rather well and has the potential to be used to actual monitoring systems in order to lessen the effect that cyberbullying has on society. This would help to reduce the negative effects that cyberbullying has on society. The ability to recognize sarcasm, for instance, is one strategy that clearly has room for development in the interest of producing more rich outcomes.

References

[1] Rosa, H., Pereira, N., Ribeiro, R., Ferreira, P. C., Carvalho, J. P., Oliveira, S., ... & Trancoso, I. (2019). Automatic cyberbullying detection: A systematic review. *Computers in Human Behavior*, *93*, 333-345.

[2] Kim, S., Razi, A., Stringhini, G., Wisniewski, P. J., & De Choudhury, M. (2021). A Human-Centered Systematic Literature Review of Cyberbullying Detection Algorithms. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1-34.

[3] Ali, W. N. H. W., Mohd, M., & Fauzi, F. (2018, November). Cyberbullying detection: an overview. In *2018 Cyber Resilience Conference (CRC)* (pp. 1-3). IEEE.

[4] Balakrishnan, V., Khan, S., & Arabnia, H. R. (2020). Improving cyberbullying detection using Twitter users' psychological features and machine learning. *Computers & Security*, *90*, 101710.

[5] Muneer, A., & Fati, S. M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on Twitter. *Future Internet*, *12*(11), 187.

[6] Iwendi, C., Srivastava, G., Khan, S., & Maddikunta, P. K. R. (2020). Cyberbullying detection solutions based on deep learning architectures. *Multimedia Systems*, 1-14.

[7] Gencoglu, O. (2020). Cyberbullying detection with fairness constraints. *IEEE Internet Computing*, *25*(1), 20-29.

[8] Cheng, L., Li, J., Silva, Y. N., Hall, D. L., & Liu, H. (2019, January). Xbully: Cyberbullying detection within a multi-modal context. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (pp. 339-347).

[9] Cheng, L., Guo, R., Silva, Y., Hall, D., & Liu, H. (2019, May). Hierarchical attention networks for cyberbullying detection on the instagram social network. In *Proceedings of the 2019 SIAM international conference on data mining* (pp. 235-243). Society for Industrial and Applied Mathematics.

[10] Balakrishnan, V., Khan, S., Fernandez, T., & Arabnia, H. R. (2019). Cyberbullying detection on twitter using Big Five and Dark Triad features. *Personality and individual differences*, *141*, 252-257.

[11] Hani, J., Mohamed, N., Ahmed, M., Emad, Z., Amer, E., & Ammar, M. (2019). Social media cyberbullying detection using machine learning. *International Journal of Advanced Computer Science and Applications*, *10*(5).

[12] Soni, D., & Singh, V. K. (2018). See no evil, hear no evil: Audio-visual-textual cyberbullying detection. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1-26.

[13] Pawar, R., & Raje, R. R. (2019, May). Multilingual cyberbullying detection system. In *2019 IEEE international conference on electro information technology (EIT)* (pp. 040-044). IEEE.

[14] Kumar, A., & Sachdeva, N. (2019). Cyberbullying detection on social multimedia using soft computing techniques: a meta-analysis. *Multimedia Tools and Applications*, *78*(17), 23973-24010.

[15] Bozyiğit, A., Utku, S., & Nasibov, E. (2021). Cyberbullying detection: Utilizing social media features. *Expert Systems with Applications*, *179*, 115001.

[16] Can, U., & Alatas, B. (2019). A new direction in social network analysis: Online social network analysis problems and applications. *Physica A: Statistical Mechanics and its Applications*, *535*, 122372.

[17] Hernandez-Suarez, A., Sanchez-Perez, G., Toscano-Medina, K., Martinez-Hernandez, V., Perez-Meana, H., Olivares-Mercado, J., & Sanchez, V. (2018). Social sentiment sensor in twitter for predicting cyber-attacks using ℓ 1 regularization. *Sensors*, *18*(5), 1380.

[18] Sharma, K., Bhasin, S., & Bharadwaj, P. (2019). A worldwide analysis of cyber security and cyber crime using Twitter. *Int. J. Eng. Adv. Technol*, *8*, 1051-1056.

[19] Founta, A. M., Chatzakou, D., Kourtellis, N., Blackburn, J., Vakali, A., & Leontiadis, I. (2019, June). A unified deep learning architecture for abuse detection. In *Proceedings of the 10th ACM conference on web science* (pp. 105-114).

[20] Al-Smadi, M., Qawasmeh, O., Al-Ayyoub, M., Jararweh, Y., & Gupta, B. (2018). Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. *Journal of computational science*, *27*, 386-393.

[21] Drishya, S. V., Saranya, S., Sheeba, J. I., & Devaneyan, S. P. (2019). Cyberbully image and text detection using convolutional neural networks. *CiiT Int. J. Fuzzy Syst*, *11*(2), 25-30.