



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 4, April 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)



# Detection of Cyberbullying on Social Media Platforms

Dr. Gururaj T <sup>1</sup>, Prof. Vaishnavi Ajay Inamdar <sup>2</sup>, Puneeth GL <sup>3</sup>, Shashank G J <sup>4</sup>, Sinchana I S <sup>5</sup>,  
Yashaswini N S <sup>6</sup>

Associate Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India<sup>1</sup>

Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India<sup>2</sup>

U.G Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India<sup>3</sup>

U.G Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India<sup>4</sup>

U.G Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India<sup>5</sup>

U.G Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India<sup>6</sup>

**ABSTRACT:** Growing internet usage and easier access to online communities like social media have led to the emergence of cybercrime. Cyberbullying is a new form of bullying that has emerged recently with the development of social networks. It is a significant issue that both adults and teenagers experience online. Cyberbullying numbers are rising annually, which has detrimental implications. It has led to the miscues like anxiety, aggression, depression, and suicide. The need for content regulation on social media platforms is growing. This project mainly focuses on detection of hate speech and personal attacks on social media platforms. The main goal is to build a model which is used for detection of Cyberbullying in text data using Deep Learning technique. The method named Semantic-Enhanced Marginalized Denoising Auto-Encoder (smSDA) is developed by semantic extension of the deep learning model stacked denoising autoencoder. The semantic extension comprises of sparsity constraints and semantic dropout noise, the latter of which is created using word embedding technology and domain knowledge. Our suggested approach can learn a strong and discriminative representation of text while taking advantage of the hidden feature structure of bullying information.

**KEYWORDS:**Cyberbullying, Stacked Denoising Autoencoders, Word Embedding, Deep Learning.

## I. INTRODUCTION

Social media is described as "a group of Internet-based applications that expand on the philosophical and technological underpinnings of Web 2.0 and that enable the creation and exchange of user-generated content". Through social media, people may easily communicate with others, access a multitude of information, and more. Social media may have certain unintended consequences, such as cyberbullying, which may have a harmful impact on people's lives, particularly those of children and teenagers.

The deployment of hostile, intentional behaviours by an individual or group of individuals against a victim through digital communication channels like messages and comments is known as cyberbullying. Furthermore, it disrupts a person's mental stability. Depression is a common side effect of cyberbullying for many people. Additionally, they engage in self-harm. They feel inferior because of all the criticism directed at them. Contrary to, traditional bullying,



which often takes place at school during face-to-face interactions, cyberbullying on social media can occur anytime, anywhere.

One method to address the problem of cyberbullying is automated detection and quick reporting of bullying communications, enabling the appropriate action to be taken to avoid potential disasters. Natural language processing and machine learning are effective research methods for bullying, as demonstrated by earlier computer studies of bullying. As a supervised learning problem, the detection of cyberbullying can be formulated.

## II. LITERATURE SURVEY

1. Devi Venkata Sai Chandra Prasad, Sujith Mudraboyina, Sri Hari Babu Gole, "Detecting Cyber-Bullying with Denoising Auto Encoder," International Journal of Advanced Research in Computer and Communication Engineering, vol. 11, Issue 1, January 2022.

The three types of information that are frequently employed in the detection of cyberbullying are text, user demography, and social network traits. Due to the reliability of text material, the above paper focuses on text-based cyberbullying detection. Study shows that the stacked denoising autoencoder (SDA), a deep learning technique, is the major topic of discussion. Multiple denoising autoencoders are stacked by SDA. Furthermore, concatenates each layer's output into a learned representation. Each denoising autoencoder in SDA is taught to recover the input data from a corrupted version of it. Dropout noise is the result of corrupting the input by arbitrarily turning some of it to zero. The autoencoders can learn robust representation using this denoising technique.

2. Dr. K.N.S Lakshmi, MaduturiNikitha, "Recognize and Prevent the Cyberbullying Conversation on Social Networks Using Machine Learning Techniques", IJCRT , vol 10, no 9, September 2022.

In this study, a noisily auto encoder is used for the first time. By developing the recommended technique, one can locate the users engaging in bad and abusive communication inside a network and stop the transmission of the abusive messages. It is possible to further improve the robustness of the learned representation by accounting for word order in messages and using natural language processing techniques to predict any abusive words that are not in the dataset and include the same by way of feedback into the BoW database.

3. Nguyen, D. T., Kim, J. M., & Jeong, D. G., "Cyberbullying detection using marginalized denoising autoencoder with dropout and convolutional neural network," Future Generation Computer Systems, 117, 378-388, 2021.

This paper proposes a cyberbullying detection system using a marginalized denoising autoencoder with dropout and convolutional neural network. The model is trained on a large dataset of tweets and is shown to achieve high accuracy in detecting cyberbullying.

4. S Manikanta, Abdul Ahad, DDD Suribabu, "A survey on Semantic-Enhanced Marginalized Denoising Auto-Encoder," JETIR, vol.5, no 6, June 2018.

The psychological science approach to cyberbullying based on individual surveys is time-consuming and might not be appropriate for automatically detecting cyberbullying.

The computational study of cyberbullying has piqued scholars' interest as machine learning has grown in popularity in recent years. Cyberbullying detection is strongly related to a number of study fields, including topic detection and affective analysis. Their efforts have made it possible to detect cyberbullying automatically.

5. Monirah A. Al-Ajlan; Mourad Ykhlef, "Optimized Twitter Cyberbullying Detection based on Deep Learning," 21st Saudi Computer Society National Computer Conference (NCC), 2018.

This paper suggests that an innovative solution to the aforementioned problems: optimised Twitter cyberbullying detection based on deep learning. Unlike earlier research in this area, the proposed method encodes a tweet as a series of word vectors rather than extracting features from tweets and feeding them to a classifier. By doing this, the feature extraction and selection stages can be skipped while still maintaining the meaning of the words. Deep learning will be employed in the classification phase, together with a metaheuristic optimisation technique for parameter adjustment.



6. D. Swathi, S. Babu, “Cyberbullying Detection based on Semantic Enhanced Marginalized Denoising Auto-Encoder,” International Journal of Science, Engineering and Management (IJSEM), vol. 3, Issue 4, April 2018.

The primary focus of this paper is on the text-based cyberbullying detection issue, where reliable and discriminative representations of messages are essential for a successful detection system. Studies on the effects of bullying have been done, but nothing has been done to monitor social media for cyberbullying. Therefore, the suggested system focuses on identifying the existence of cyberbullying behaviour in social networks and preventing the posting of cyberbullying terms in these networks. The suggested solution focuses on both identifying the predator and preventing their access to social networks. This might contribute to the creation of a positive and secure online community.

### III. METHODOLOGY

#### Modules

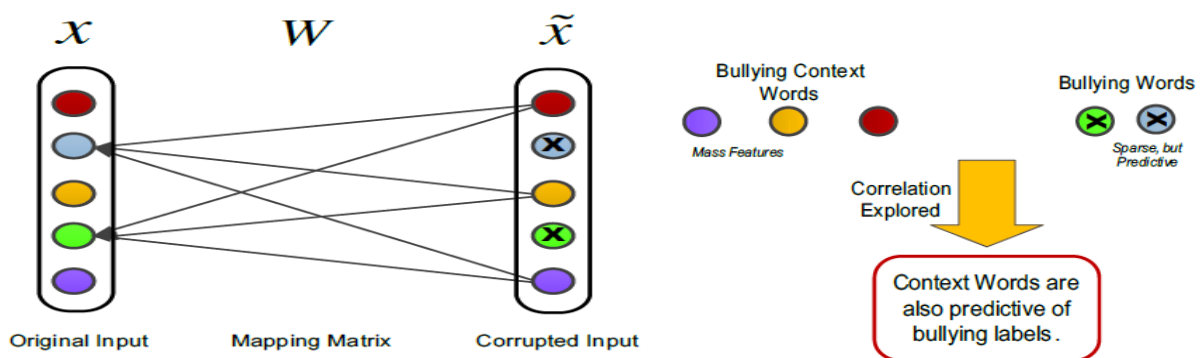
1. Marginalized Stacked Denoising Auto-encoder
2. Semantic Enhancement for mSDA
3. Construction of Bullying Feature Set
4. smSDA for Cyberbullying Detection

1. Marginalized Stacked Denoising Auto-encoder

This method proposes a revised version of stacked denoising auto encoder that uses a linear projection rather than non-linear projection to obtain a closed-form solution. The fundamental principle behind denoising auto-encoder is to reconstruct the original input from a corrupted one to obtain a robust representation. The main goal behind denoising encoder is to reconstruct original data using linear projection.

2. Semantic Enhancement for mSDA

A denoising auto-encoder is trained to reconstruct the removed feature values from rest of the uncorrupted data. Thus, the learned mapping matrix  $W$  is able to capture correlation between the removed features and that of the other features. The major modifications include semantic dropout noise and sparse mapping constraints.



3. Construction of Bullying Feature Set

Expert knowledge and word embeddings are used to build a list of words with negative affective features and compare it with BoW features. Discriminative feature selection is conducted to expand the list of insulting words based on word embeddings. In a vector space well-trained word embeddings with similar words lie close to each other. The semantic similarity between words can also be measured using the cosine similarity between word embeddings. Due to different word usages and insulting expressions, insulting words are located at different regions.

4. smSDA for Cyberbullying Detection

smSDA is used to leverage semantic information and feature correlation in order to construct a pre-defined bullying wordlist. This bullying wordlist is then used to automatically expand and refine insulting words, which are then used as bullying features to layer one of the model using word embeddings.

IV. EXPERIMENTAL RESULT

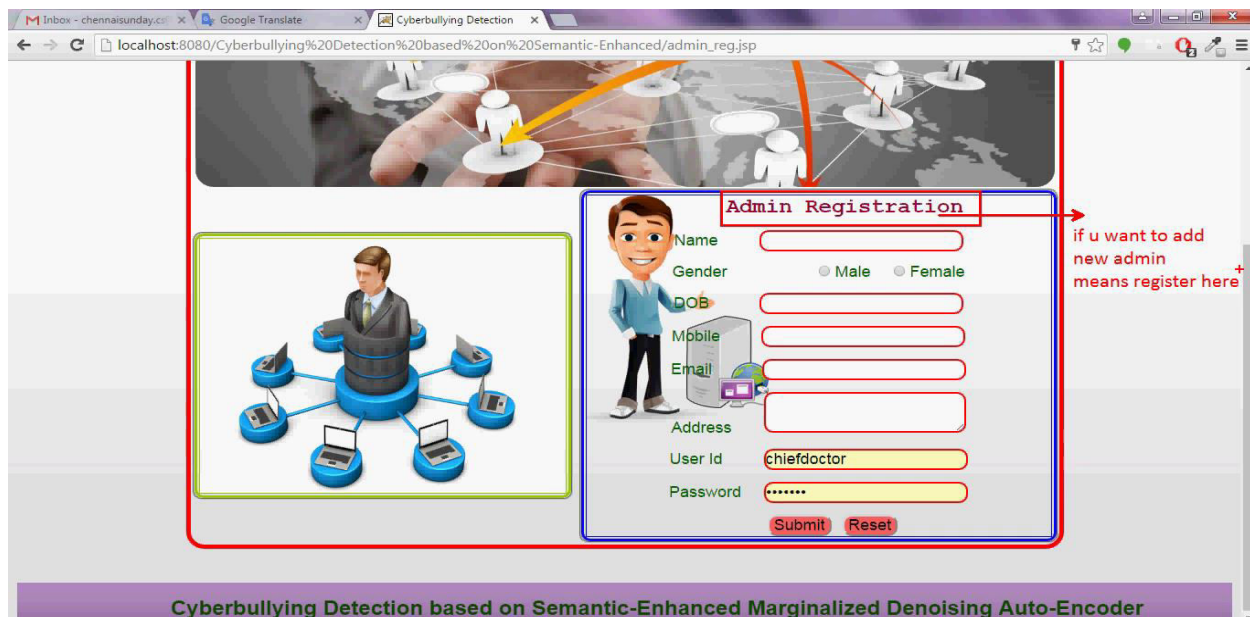


Figure 1: Admin registration page



Figure 2: Admin login page



|| Volume 12, Issue 4, April 2023 ||

| DOI:10.15680/IJIRSET.2023.1204277 |

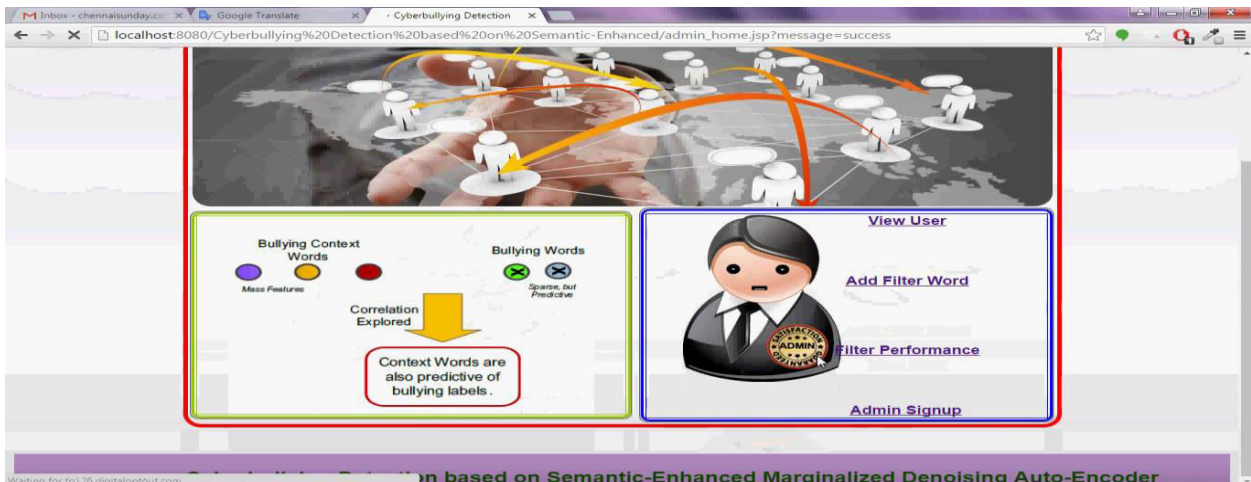


Figure 3: View of admin login page

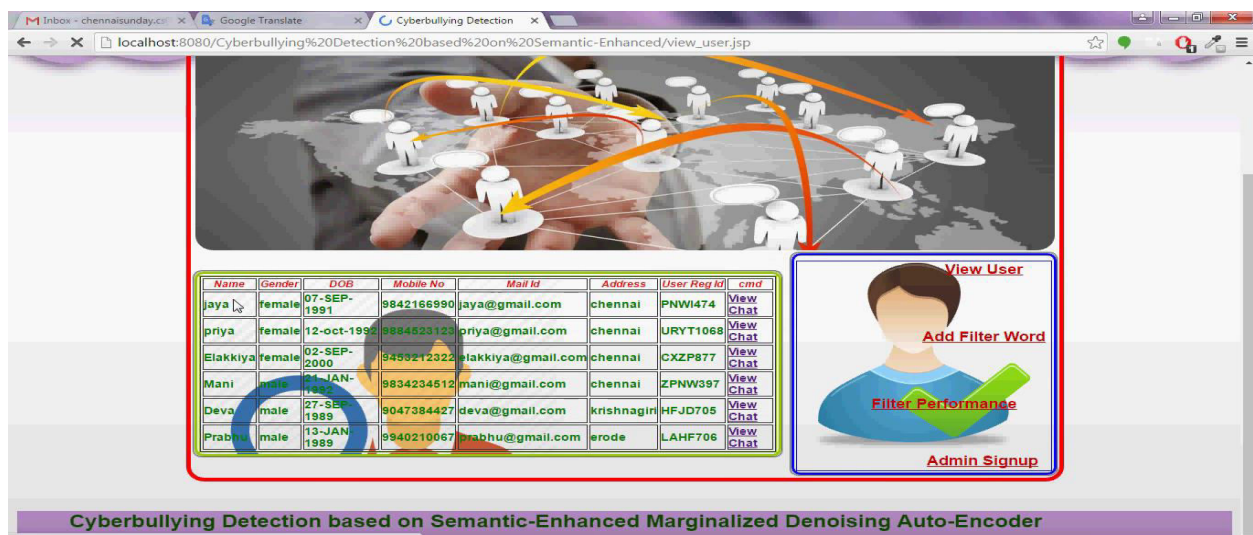


Figure 4: View of user details page



Figure 5: User registration page

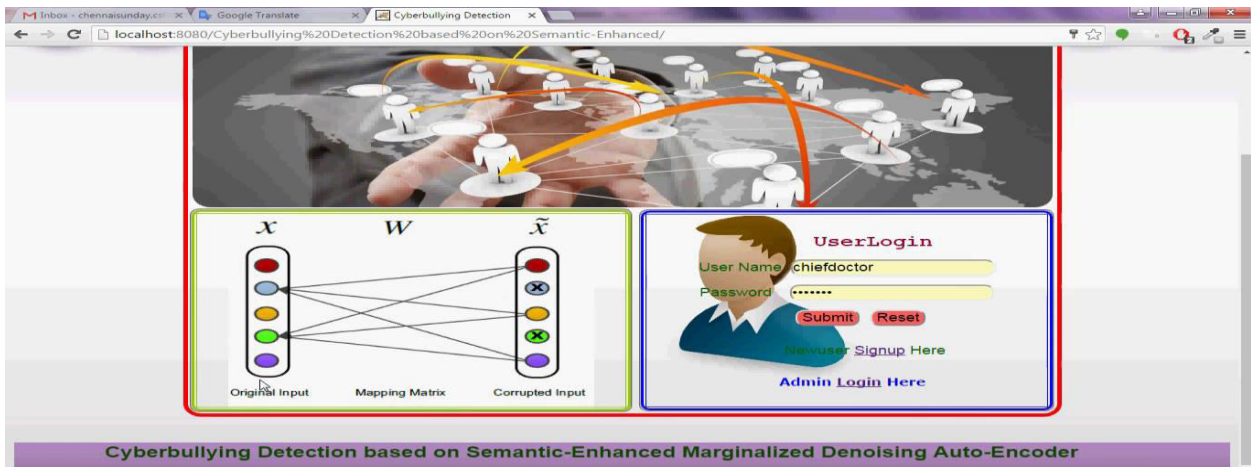


Figure 6: User login page

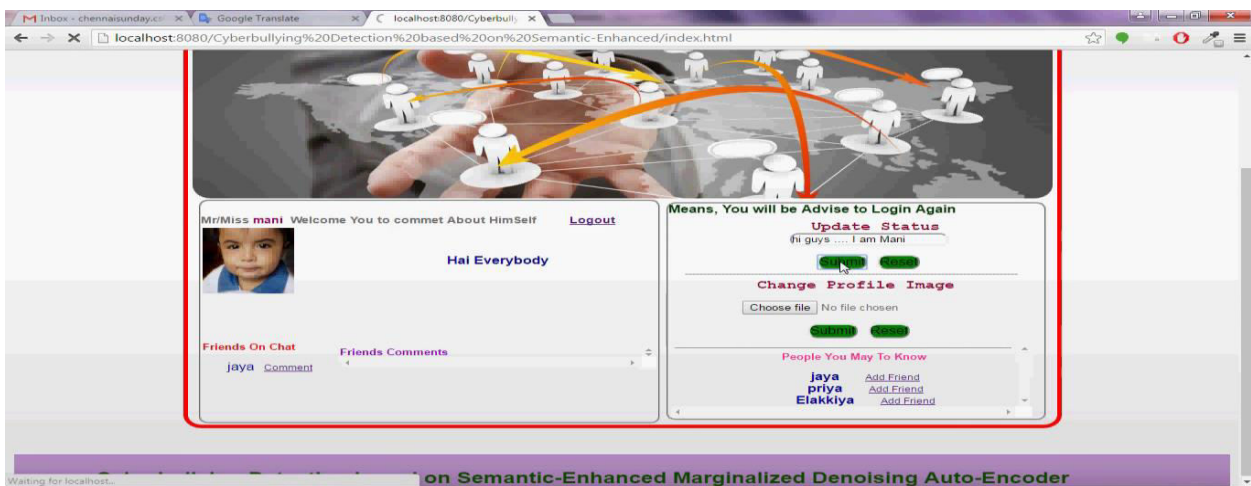


Figure 7: User status update page

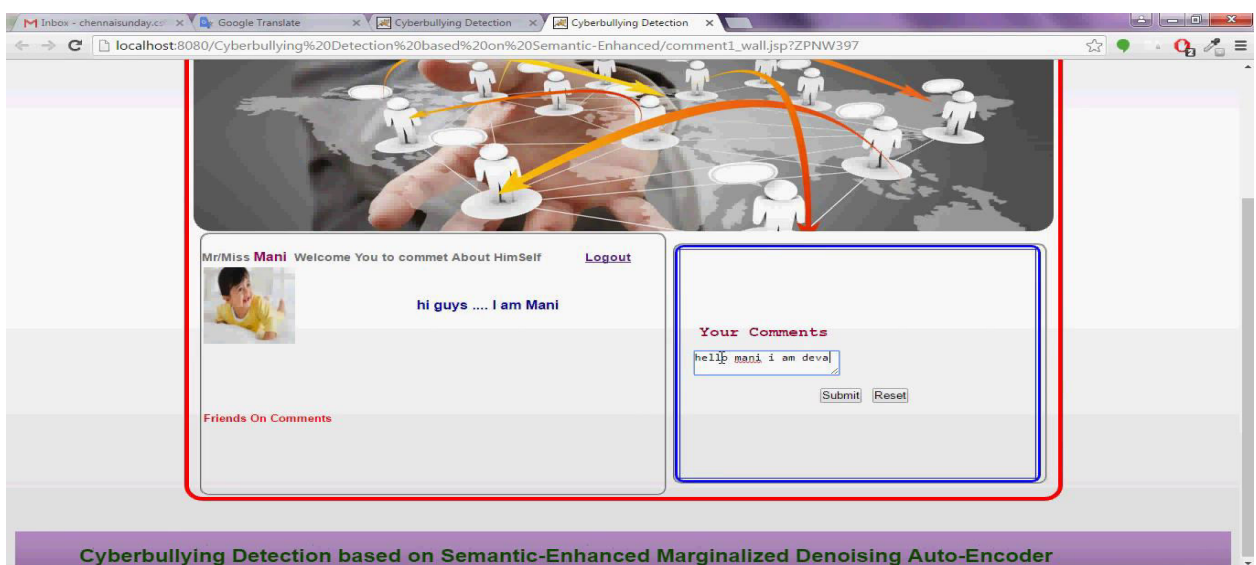


Figure 8: User comments update page



Figure 9: Page that shows bullying words cannot be commented or posted

## V. CONCLUSION

The problems of text-based cyberbullying detection, where a reliable and discriminative representation of communications is essential for a reliable detection system. The proposed system demonstrates a specialized representation learning model for the identification of cyberbullying by incorporating semantic dropout noise into the design process and enforcing sparsity. The automatic expansion and improvement of bullying word lists that are initialized by domain knowledge has also been done using word embeddings. The use of Semantic Enhanced Marginalized Denoising Auto Encoder (SEMDAE) in detecting cyberbullying is a promising approach. Overall, the use of SEMDAE in detecting cyberbullying is a step forward in addressing the growing problem of cyberbullying.

## REFERENCES

1. Devi Venkata Sai Chandra Prasad, Sujith Mudraboyina, Sri Hari Babu Gole, "Detecting Cyber-Bullying with Denoising Auto Encoder," International Journal of Advanced Research in Computer and Communication Engineering, vol. 11, Issue 1, January 2022.
2. Dr. K.N.S Lakshmi, MaduturiNikitha, "Recognize and Prevent the Cyberbullying Conversation on Social Networks Using Machine Learning Techniques", IJCRT , vol 10, no 9, September 2022.
3. Nguyen, D. T., Kim, J. M., & Jeong, D. G., " Cyberbullying detection using marginalized denoising autoencoder with dropout and convolutional neural network," Future Generation Computer Systems, 117, 378-388, 2021.
4. S Manikanta, Abdul Ahad, DDD Suribabu, " A survey on Semantic-Enhanced Marginalized Denoising Auto-Encoder," JETIR, vol.5, no 6, June 2018.
5. Monirah A. Al-Ajlan; Mourad Ykhlef, "Optimized Twitter Cyberbullying Detection based on Deep Learning," 21st Saudi Computer Society National Computer Conference (NCC), 2018.
6. D. Swathi, S. Babu, "Cyberbullying Detection based on Semantic Enhanced Marginalized Denoising Auto-Encoder," International Journal of Science, Engineering and Management (IJSEM), vol. 3, Issue 4, April 2018.
7. Dabaghi-Zarandi, F., &Kuchaki Rafsanjani, "Community Detection in Social Networks. In Studies in Systems, Decision and Control", 2019.
8. Available[online]: [https://doi.org/10.1007/978-3030-00084-4\\_15](https://doi.org/10.1007/978-3030-00084-4_15)





**|| Volume 12, Issue 4, April 2023 ||**

**| DOI:10.15680/IJIRSET.2023.1204277 |**

9. J. Sui, “Understanding and fighting bullying with machine learning,” Ph.D. dissertation, THE UNIVERSITY OF WISCONSIN-MADISON, 2015.
10. Divyashree, Vinutha H, Deepashree N S, “An Effective Approach for Cyberbullying Detection and Avoidance”, International Journal of Innovative Research in Computer and Communication Engineering, pp. 8005-8010, 2016.
11. Q. Huang, V. K. Singh, and P. K. Atrey, “Cyber bullying detection using social and textual analysis,” in Proceedings of the 3rd International Workshop on Socially-Aware Multimedia. ACM, 2014, pp.3–6.



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.423**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details