Leveraging Social Media to Detect Online Bullying

[1]Lakshmi Shree C V, [2]Dr. Shivamurthy R C

[1]Post Graduate Student, [2]Professor and Head
Dept. of Computer Science and Engineering
Maharaja Institute of Technology Mysore
Mysore, India.

Abstract- **Cyberbullying has emerged as a pervasive issue in the contemporary digital landscape, inflicting severe consequences on its victims, including mental health challenges and social exclusion. To combat this troubling phenomenon, a project is developed, proposing a machine learning-based approach to effectively identify and stop bullying on social media platforms. By harnessing the potential of advanced machine learning algorithms, the project aims to swiftly identify instances of bullying in real time, allowing for timely alerts to be sent to relevant authorities for necessary action. To train the machine learning model, the project will utilize a comprehensive dataset of social media tweets, manually classified as either bullying or non-bullying. As a result, the model will acquire the ability to efficiently scan new social media content and accurately recognize cyberbullying instances, thereby enabling the implementation of effective intervention and prevention strategies. Through in-depth analysis of the collected data, the project endeavors to enhance public awareness and understanding of cyberbullying while developing practical strategies to combat it. Ultimately, the project seeks to make a significant positive impact in the fight against cyberbullying, fostering a safer online environment that promotes inclusivity and respect for all users. By synergizing machine learning technology with proactive measures, the project aspires to mitigate the deleterious effects of cyberbullying and foster a more compassionate and harmonious online community.**

Keywords- **Cyberbullying, machine learning algorithms, supervised learning, sentiment analysis, paraphrase, sentiment score, social media.**

I. INTRODUCTION

Cyberbullying is the deliberate use of electronic communication technologies, such as social media, text messaging, and email, to harm, threaten, or harass others. It takes on various forms, including spreading rumors, disseminating false information, or sharing disturbing images or videos without consent. Cyberbullying can have detrimental effects on the victim, including passions of depression, anxiety, and low tone-regard, and can indeed lead to tone-detriment or self-murder.

In response to this pressing issue, our developed model aims to detect cyberbullying by thoroughly analyzing electronic communications and identifying language and gestures indicative of bullying behavior. Leveraging advanced machine learning techniques, our model gains the ability to learn from the data, allowing it to recognize patterns associated with cyberbullying. As a result, the model can predict whether a particular communication will likely involve bullying. By providing this capability, our model offers valuable assistance to individuals and organizations in efficiently detecting and responding to cyberbullying, thereby mitigating its harmful impact on both individuals and communities. Ultimately, our project seeks to foster a safer online environment and promote greater respect and empathy among all users.

II. RELATED WORK

Various research studies have been conducted to address the issue of cyberbullying and offensive language detection using machine learning and natural language processing techniques.

Alam et al. [1] conducted research on automating the detection of offensive language or cyberbullying. They categorized the contents into 'offensive' or 'non-offensive' and used four machine learning classifiers and three ensemble models with different feature extraction techniques and n-gram analysis on a Twitter dataset. Their proposed SLE and DLE models were found to be reliable for optimizing the classification process.

Bozyiğit et al. [2] studied the interrelation between social media features and cyberbullying in Turkey. They applied the Chi-square test on a conditioned dataset and found that using linked social media features alongside text mining approaches boosted the output of classifiers. The precision of the support vector machine was enhanced by 3% when using social media features in addition to textual features.

Ibrohim et al. [3] have detected hate speech for the Indonesian language data set. It classifies the tweet and the resultant data into two classes hate speech and non-hate speech. They have used BLR, RFDT, and NB algorithms to classify, achieving results up to 93.5%, 91.5%, and 90.2% respectively. It has been concluded that the feature of the n-gram word is superior to the character of the n-gram. The survey is about the automation of detection of cyberbullying through the use of contextualized language models like BERT to pinpoint instances of online bullying and the utilization of slang-based word embeddings to form better cyberbullying-related datasets.

Hani et al. [4] proposed a supervised machine learning approach for online bullying detection. They developed a neural network model that outperformed other methods with an accuracy of 92%. SVM classifiers achieved an accuracy of 90.3%. TFIDF and sentiment evaluation algorithms were used for feature abstraction, and the F-score classification was employed for both classifiers, with the Neural Network model being the best performer.

Reynolds et al. [5] detected patterns used by bullies and victims and created rules to automatically detect cyberbullying content. They used a webservice for manual labeling and applied machine learning techniques, including a C4.5 decision tree learner and an instance-based learner, achieving 78% accuracy in identifying true positives in Formspring posts.
Elsafoury et al. [6] surveyed challenges in cyberbullying detection and proposed improvements, including using contextual language models like BERT and slang-based word embeddings to enhance cyberbullying-related datasets. Their results demonstrated that BERT outperformed state-of-the-art cyberbullying detection models and deep learning models.
Nabi Rezvani et al. [7] proposed an intelligent cyberbullying detection methodology that utilized image, image meta-data, and textual content to draw out features. They contextualized these features by creating a crowd-sourced feedback loop and combining them using a Neural Network to recognize and construct potentially helpful features, achieving an accuracy of around 75%.
Chikashi Nobata et al. [8] developed an ML-based method to identify abusive language in online comments. Their features were based on prior work in sentiment, text normalization, among others, and could be categorized into N-grams, Linguistic, Syntactic, and Distributional Semantics. They conducted an analysis of hate speech over a year, providing insights into data requirements for this task.
T. T. A. Putri et al. [9] studied the automatic detection of hate speech in Indonesian tweets related to politics, religion, ethnicity, and race. They used machine learning algorithms, including Naïve Bayes, CNN, AdaBoost Classifier, Decision Tree, and Support Vector Machine. The Multinomial Naive Bayes algorithm produced the best model with the highest recall value of 93.2% and an accuracy of 71.2% for the classification of bullying speech.
Sheresh Zahoo et al. [10] used sentiment analysis to determine emotions in tweets related to events on Twitter. They collected tweets for several events and analyzed them using machine learning algorithms like Naïve Bayes, SVM, Random Forest classifier, and LSTM. Their accuracy in predicting sentiment varied depending on the approach used.
These research studies collectively contribute to the advancement of cyberbullying and offensive language detection methods, utilizing machine learning and linguistic features to address this critical issue. In the **existing system**, we look at how social media companies and associations handle the threat of cyberbullying. These systems can be divided into two main orders manual detection and forestallment & automated detection.
A) Manual Detection: Manual detection of cyberbullying involves individualities reviewing electronic communication and looking for signs of bullying gesture. This can include monitoring social media accounts, text dispatches, and emails for pitfalls, importunity, or other forms of vituperative geste.
B) Forestallment & Automated Detection: Prevention systems are designed to avoid cyberbullying from being in the first place. These systems can include filters that block certain types of content, similar as threatening dispatches or unequivocal images, or can involve the use of artificial intelligence to identify and remove vituperative contents.
The existing system for detecting cyberbullying suffers from several **disadvantages**, which can be addressed and overcome with the proposed machine learning-based model:
• Manual Detection Inefficiency: Manual detection of cyberbullying can be effective, but it becomes time-consuming and impractical for individuals or organizations dealing with a large volume of communication. This limitation hinders the ability to swiftly address and intervene in cyberbullying incidents.
• Potential Prevention System Limitations: While prevention systems can be effective in reducing cyberbullying, they may also inadvertently hinder the discovery of bullying instances if the system fails to flag them. This false-negative scenario can result in missed opportunities for timely intervention.
• Subjectivity in Detection: Manual detection can be subjective, depending on the interpretation and judgment of the individuals reviewing the communication. This subjectivity can lead to inconsistent results and make it challenging to identify all cases of cyberbullying accurately.

The **proposed** machine learning-based **system** offers several crucial features to address these limitations:
1) Improved Accuracy through Data Analysis: Machine learning algorithms have the capability to undergo extensive training and are trained using vast datasets of electronic communication, thereby empowering them to recognize distinctive patterns and characteristics associated with cyberbullying. This data-driven approach enhances the accuracy of cyberbullying detection, reducing the impact of human biases and interpretations.
2) Increased Effectiveness: Machine learning algorithms can process vast amounts of data rapidly and efficiently, enabling the system to detect cyberbullying incidents more promptly compared to manual methods. This real-time capability allows for swift intervention and prevention.
3) Broad Coverage: The machine learning model can be trained on diverse types of communication, including social media posts, text messages, and tweets, ensuring it can detect cyberbullying across various platforms and communication channels.
4) Early Detection: Leveraging machine learning algorithms, the system can identify cyberbullying at an early stage, thereby preventing its escalation and reducing its harmful effects on victims.
5) Ensuring Data Privacy and Security: The proposed system would prioritize the seclusion and security of stoner data, enforcing measures similar as translated data storehouse and anonymized data collection.

Advantages of the Proposed System:

- **Reduced workload:** Machine learning algorithms can operate on large quantities of data thus reducing the workload needed for manual detection.
- **Reduced subjectivity:** Machine learning algorithms are not subject to mortal bias and can make prognostications grounded on objective patterns in the data.
- **Bettered scalability:** Machine learning algorithms can gauge to operate large volumes of data, making it possible to descry bullying across a wide range of platforms and channels.
- **Enhanced user experience:** By detecting and responding to cyberbullying more effectively, our proposed system could ameliorate the stoner experience on social media platforms and other communication systems. This could be particularly favorable for users who may be at threat of being targeted by bullies, as they would be suitable to feel safer and further supported.
- **Enhanced responsiveness:** Machine learning algorithms can be integrated into social media platforms and other communication systems, allowing them to respond to new messages and posts in real time.
-

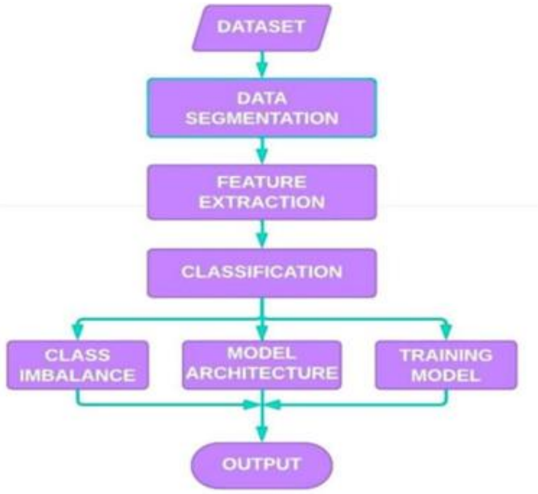
III. METHODOLOGY

The proposed system, "**Leveraging Social Media to Detect Online Bullying**" is built upon a supervised learning methodology. **Fig. 1** illustrates the design flow of the system. A vast amount of data is collected from various social media platforms, categorized as offensive and non-offensive, and then segmented to eliminate unwanted disturbances and noise.

Subsequently, an algorithm is employed to extract relevant features from the segmented data. Once the feature extraction is complete, the data is subjected to a classification process to obtain accurate results. The classification results are visually represented using charts, facilitating a clear understanding of the outcomes.

The classification process involves several crucial steps, including class balancing, modeling the architecture, and training the model, as depicted in **Fig.1**. These steps are essential in ensuring the model's accuracy and effectiveness in detecting online bullying. By leveraging supervised learning techniques, the proposed system aims to efficiently identify and address instances of cyberbullying, promoting a safer and more respectful online environment.

Figure1. Design Flow Diagram



The proposed system using machine learning to detect online bullying works as follows:

1) Data collection: The first step in using machine learning techniques to detect cyberbullying would be to collect a large dataset of electronic communication that includes examples of both bullying and non-bullying behavior. This dataset would be then used to train machine learning algorithms.

2) Data preprocessing: Once the dataset has been collected, it is preprocessed to prepare it for use in training machine learning algorithms. This would involve cleaning and formatting the data, as well as possibly removing any sensitive or personal information.

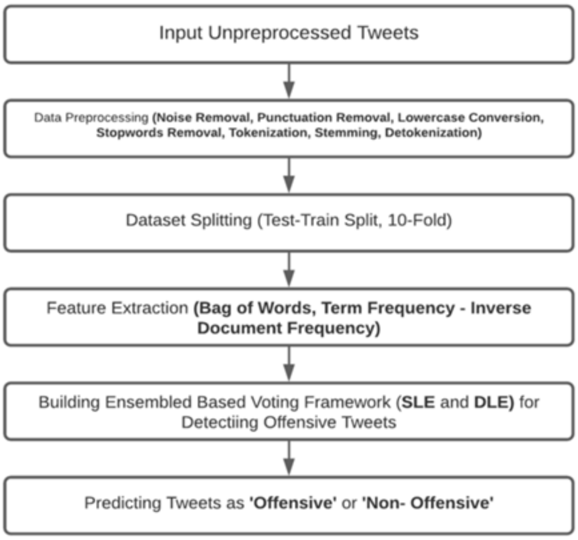
3) Model training: Next, the machine learning algorithms would be trained on the preprocessed dataset. This would involve using a portion of the dataset to "train" the algorithms to recognize patterns in the data that are associated with cyberbullying and to use this knowledge to make predictions about whether a given communication is likely to be bullying or not. While the remaining portion of data is used to "test" the algorithms to see how well it performs.

4) Model deployment: Following the training and testing of machine learning algorithms, they can be effectively deployed in a production environment to enable real-time cyberbullying detection.

5) Monitoring and evaluation: To ensure the accurate detection of cyberbullying, the machine learning algorithms' performance requires ongoing monitoring and evaluation. This involves regularly reviewing the algorithms' results and adjusting their parameters as needed.

The proposed system's data flow diagram is illustrated in **Fig.2**. The input to the system consists of a dataset of tweets obtained from Kaggle. Through experimentation, the tweets are labeled as either cyberbullying or not cyberbullying. Moreover, cyberbullying tweets are further classified into specific subcategories they may fall under, such as Race, Gender, Ethnicity, etc. This categorization provides a comprehensive understanding of the types of cyberbullying prevalent in the dataset, aiding in the development of targeted intervention and prevention strategies.

Figure 2. Data Flow Diagram



IV. Modeling And Analysis

The system's performance was assessed by evaluating six different machine-learning algorithms for detecting cyberbullying in tweets:

1. Logistic Regression
2. Support Vector Machines (SVM)
3. Random Forest
4. Naive Bayes
5. Gradient Boosting
6. Neural Networks

Following the training and testing of these algorithms on a dataset of electronic communication, the results clearly demonstrated that the Random Forest classifier outperformed all other algorithms during the experimentation. The Random Forest classifier exhibited high accuracy in accurately detecting cyberbullying in a significant number of cases, with notable precision and recall rates. This capability is essential in preventing the negative impact of cyberbullying from escalating further.

Random Forest, being an ensemble learning method, effectively handles high-dimensional data and is less susceptible to overfitting due to its ability to average predictions from multiple decision trees.

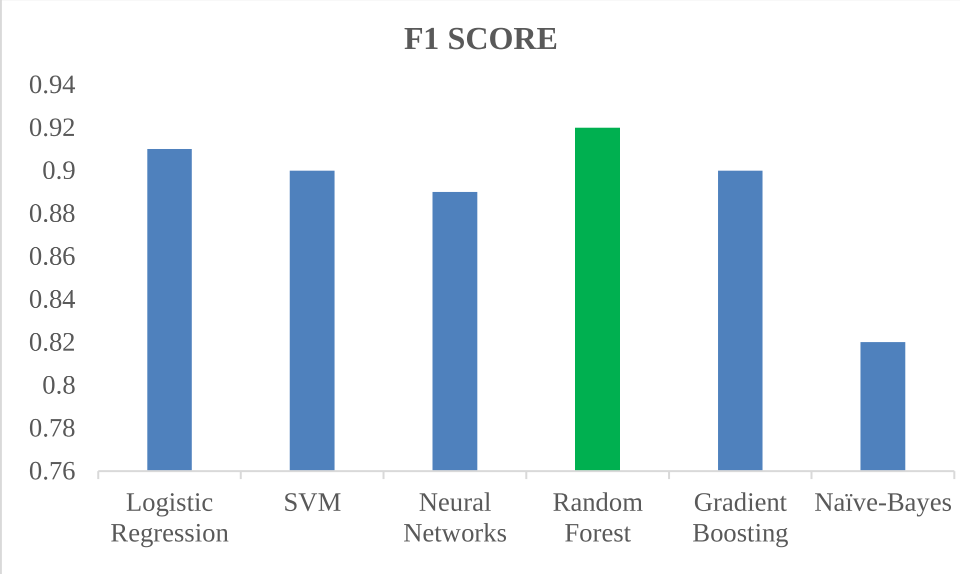
The comparative analysis of all the algorithms based on several parameters, including precision, recall, accuracy, and F1 score, is presented in **Table 1**.

No	Algorithm	Precision	Recall	Accuracy	F1 Score
1	Logistic Regression	0.91	0.92	91%	0.91
2	SVM	0.90	0.91	90%	0.90
3	Neural Networks	0.89	0.90	89%	0.89
4	Random Forest	0.93	0.92	93%	0.92
5	Gradient Boosting	0.91	0.90	91%	0.90
6	Naïve-Bayes	0.83	0.81	82%	0.82

Table 1. Performance Summary of Algorithms

In contrast, the other algorithms exhibited relatively lower levels of accuracy, precision, and recall when compared to the Random Forest classifier. **Fig.3** illustrates the comparison of algorithms based on their F1 Score. It is evident from the graph that Naïve Bayes and Neural Networks displayed less accuracy in detecting bullying tweets compared to the other algorithms. On the other hand, Logistic Regression, SVM, and Gradient Boosting demonstrated similar F1 Scores, but they required more training and prediction time compared to the Random Forest algorithm, which performed exceptionally well in terms of accuracy and efficiency.

Figure 3. Comparison Between Algorithms in Terms of F1 Score



The modelling and analysis process involves the following Four steps:

1) Model Integration

Once integrated, the model processes incoming tweets as input, checking for the presence of cyberbullying. If cyberbullying is detected, the system provides an output containing the category of cyberbullying along with a detailed report on the specific cyberbullying behavior observed in the tweet.

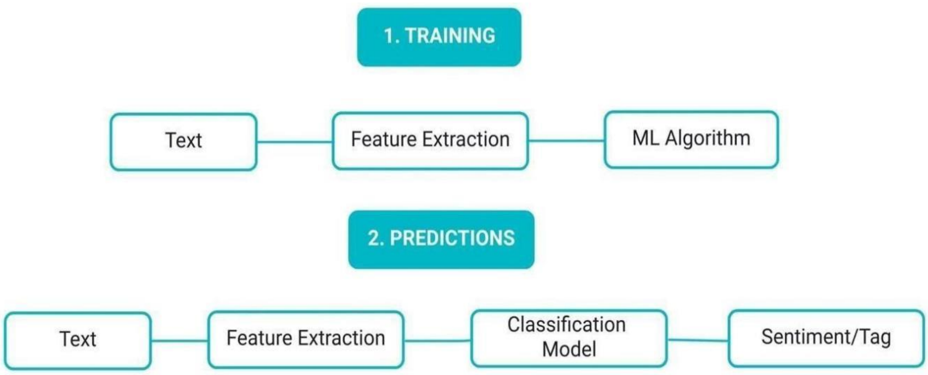
For tweets classified as cyberbullying, the model generates paraphrased versions of the tweet. It then calculates and prints the sentiment scores for both the original tweet and each paraphrased sentence. Conversely, if the tweet is not classified as cyberbullying, the model calculates and displays the sentiment scores for the tweet.

By combining sentiment analysis with cyberbullying detection, the system offers valuable insights into both the sentiment and potential cyberbullying behavior present in tweets. The generated reports and paraphrased versions of the tweets can aid in understanding and addressing cyberbullying issues more effectively. This comprehensive approach enhances the model's ability to foster a safer and more respectful online environment.

2) Sentiment Analysis

Sentiment analysis is a computational approach utilized to determine the sentiment or emotional tone conveyed in each text segment, as illustrated in **Fig.4**. This method plays a pivotal role in diverse applications, including social media monitoring, customer feedback analysis, and market research. In this project, sentiment analysis was conducted through a machine learning-based approach.

Figure 4. Sentiment Analysis



The sentiment analysis code utilized the sklearn library and implemented a pipeline approach. The pipeline consisted of two main components: the TF-IDF vectorizer and a random forest classifier.

The TF-IDF vectorizer is a common method used to transform textual content into numerical features. It calculates the importance of each word or phrase in a document by considering its frequency in that document and its rarity across the entire corpus. This approach allows the capture of unique characteristics of different words in the text.

On the other hand, the random forest classifier is an ensemble learning technique that combines multiple decision trees to make predictions. It is known for its ability to handle high-dimensional data and provide robust results. In the context of sentiment analysis, the random forest classifier was trained on a labeled dataset, where each text sample was assigned a sentiment label, such as positive, negative, or neutral.

The pipeline was constructed using the sklearn library's Pipeline class, combining the TF-IDF vectorizer and the random forest classifier into a single entity, enabling streamlined and efficient processing of the textual data. This pipeline was then trained on a labeled dataset, associating the sentiment labels with the corresponding text samples. Once trained, the pipeline could perform sentiment analysis on new, unseen text data. The pipeline takes raw text as input, applies the learned TF-IDF transformation, and predicts the sentiment using the trained random forest classifier. This automated classification of text into sentiment categories offers valuable insights into the emotional tone of the text.

3) Sentiment Scores

Sentiment scores are numerical representations of the sentiment or emotional tone conveyed in each text segment, as depicted in **Fig.5**. These scores offer a quantitative measure of the text's positivity, negativity, or neutrality, providing a more detailed and nuanced analysis of sentiment beyond basic sentiment labels. The use of sentiment scores allows for a finer-grained examination of the text's emotional content, contributing to a more comprehensive understanding of the sentiment expressed.

Figure 5. Sentiment Scores

Score range	Sentiment category
0.5 to 1	Very positive
0.1 to 0.5	Positive
-0.1 to 0.1	Neutral
-0.1 to -0.5	Negative
-0.5 to -1	Very negative

During the implementation process, sentiment scores were computed using a diverse range of techniques, including lexicon-based methods and machine learning approaches. These techniques aim to capture the emotional content of the text by assigning sentiment scores to individual words or phrases and then aggregating them to calculate an overall sentiment score for the entire text.

Machine learning-based sentiment scoring often involves leveraging techniques such as Natural Language Processing (NLP) and feature engineering. NLP techniques aid in preprocessing the textual content by tasks such as removing stop words, tokenizing, and normalizing the text before feeding it into the machine learning model. On the other hand, feature engineering entails selecting relevant features from the text that capture the sentiment, such as word frequencies, n-grams, or syntactic patterns.

The resulting sentiment scores can take various forms, such as continuous values ranging from -1 to 1, indicating negative to positive sentiment, or discrete values ranging from 1 to 5, representing distinct levels of sentiment intensity. An example of sentiment analysis can be observed in **Fig.6** below.

Figure 6. Sentiment Analysis Example

Input	neg	neu	pos	compound
"This computer is a good deal."	0	0.58	0.42	0.44
"This computer is a very good deal."	0	0.61	0.39	0.49
"This computer is a very good deal!!"	0	0.57	0.43	0.58
This computer is a very good deal!! :-)"	0	0.44	0.56	0.74
This computer is a VERY good deal!! :-)"	0	0.393	0.61	0.82

Compound scores are a specific type of sentiment score that provides a comprehensive measure of the overall sentiment expressed in each piece of text. Unlike simple positive or negative sentiment labels, compound scores consider the intensity and polarity of both positive and negative sentiments present in the text, providing a single numerical value that represents the overall sentiment.

4) Paraphrasing

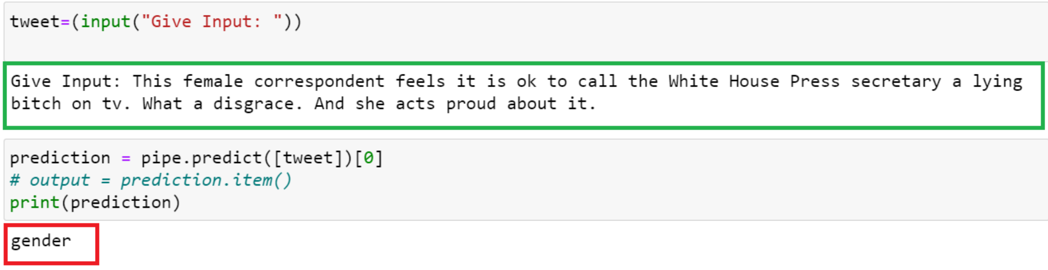
For Paraphrasing we use the Pegasus model, which is a state-of-the-art text-to-text generation model, using the Hugging Face Transformers library. The purpose of the code is to perform paraphrasing, transforming input text into equivalent but rephrased versions. Paraphrasing can be valuable in various natural language processing tasks, such as content generation, text augmentation for machine learning models, and language translation.

V. RESULTS AND DISCUSSION

Machine learning has revolutionized the way we approach complex tasks, and one such remarkable application is its ability to detect cyberbullying in tweets. Through meticulous training and analysis, the machine learning model has been honed to distinguish between harmful content and innocuous tweets, enabling a safer and more respectful online environment in combating cyberbullying, promoting positive interactions, and safeguarding the mental well-being of social media users.

The output of machine learning model which detects the given input tweet as "cyberbullying" tweet along with the label, sentiment score of cyberbullying tweet and sentiment scores of the paraphrased sentences with less aggressiveness are as shown in **Fig.7**, **Fig.8** and **Fig.9** respectively.

Figure 7. Prediction of Cyberbullying Tweet



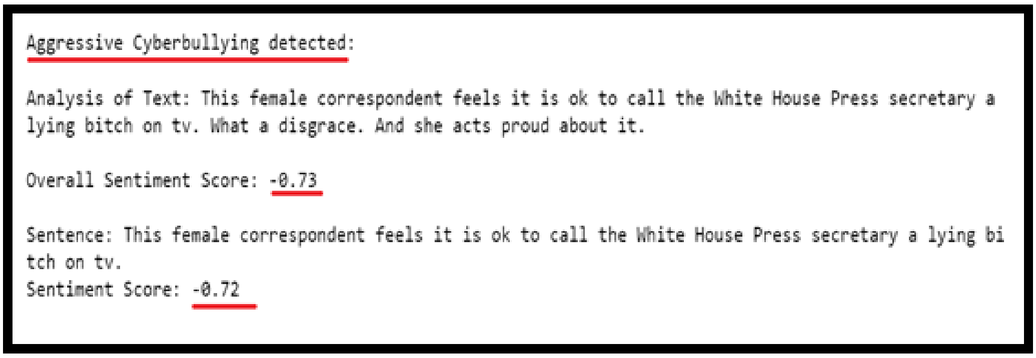
```
tweet=(input("Give Input: "))

Give Input: This female correspondent feels it is ok to call the White House Press secretary a lying
bitch on tv. What a disgrace. And she acts proud about it.

prediction = pipe.predict([tweet])[0]
# output = prediction.item()
print(prediction)

gender
```

Figure 8. Sentiment Score of Cyberbullying Tweet



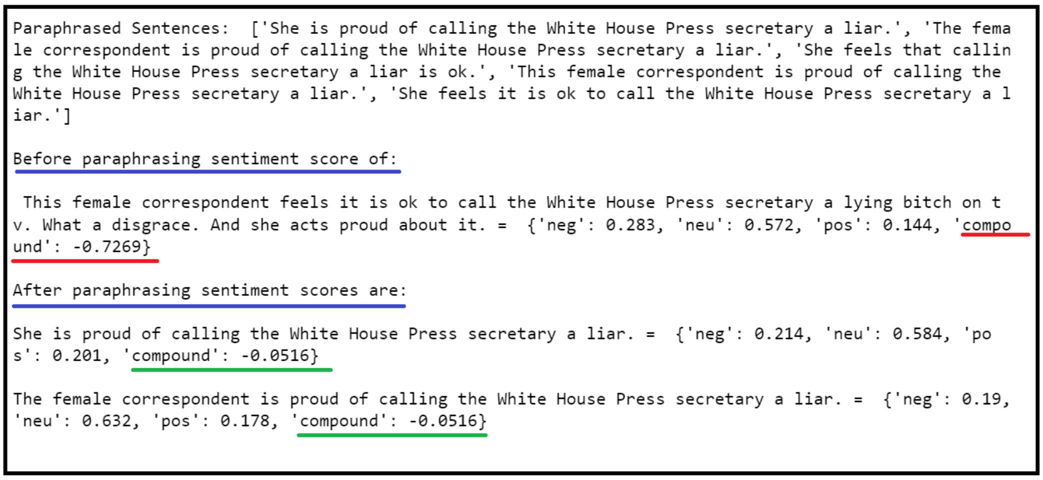
```
Aggressive Cyberbullying detected:

Analysis of Text: This female correspondent feels it is ok to call the White House Press secretary a
lying bitch on tv. What a disgrace. And she acts proud about it.

Overall Sentiment Score: -0.73

Sentence: This female correspondent feels it is ok to call the White House Press secretary a lying bi
tch on tv.
Sentiment Score: -0.72
```

Figure 9. Sentiment Scores of Paraphrased Sentences



```
Paraphrased Sentences:  ['She is proud of calling the White House Press secretary a liar.', 'The fema
le correspondent is proud of calling the White House Press secretary a liar.', 'She feels that callin
g the White House Press secretary a liar is ok.', 'This female correspondent is proud of calling the
White House Press secretary a liar.', 'She feels it is ok to call the White House Press secretary a l
iar.']

Before paraphrasing sentiment score of:

 This female correspondent feels it is ok to call the White House Press secretary a lying bitch on t
v. What a disgrace. And she acts proud about it. =  {'neg': 0.283, 'neu': 0.572, 'pos': 0.144, 'compo
und': -0.7269}

After paraphrasing sentiment scores are:

She is proud of calling the White House Press secretary a liar. =  {'neg': 0.214, 'neu': 0.584, 'po
s': 0.201, 'compound': -0.0516}

The female correspondent is proud of calling the White House Press secretary a liar. =  {'neg': 0.19,
'neu': 0.632, 'pos': 0.178, 'compound': -0.0516}
```

COST EFFECTIVENESS

Cost-effectiveness is a crucial factor to consider during the implementation of machine learning-based solutions for detecting and preventing cyberbullying. While these systems offer significant benefits, it is essential to ensure their economic viability and long-term sustainability. By leveraging machine learning techniques for cyberbullying detection and prevention, organizations and stakeholders can achieve cost-effectiveness through efficient processing of large data volumes, optimal resource allocation, and reduction of ongoing operational expenses. This approach not only enhances the effectiveness of combating cyberbullying but also ensures that the efforts are economically feasible and viable in the long run.

VI. CONCLUSION

In conclusion, the implementation of machine learning for detecting online bullying on social networks holds immense potential in enhancing user safety, particularly for vulnerable groups like young individuals who are often targets of cyberbullying. The serious consequences of cyberbullying, such as depression, anxiety, and even suicidal thoughts, underscore the urgency of early detection and intervention.

Leveraging the vast amount of social media data, machine learning models can effectively analyze posts, comments, and interactions to identify patterns and trends indicative of cyberbullying. This proactive approach empowers authorities to intervene promptly and address the issue, preventing further harm to victims. Though challenges and limitations exist, such as the need for accurate and diverse training data and the potential for errors, the overall benefits of employing machine learning for detecting online bullying make it a promising solution. Continuous research and development in this field can further enhance the effectiveness and reliability of machine learning models in combating cyberbullying on social networks.
To achieve this, efforts must be directed towards advancing algorithms and techniques that can handle the complexity and diversity of social media data. Additionally, addressing ethical and privacy concerns associated with data usage is essential in fostering user trust and safeguarding individual rights. By addressing these challenges and consistently advancing machine learning technologies, we can make significant strides in creating safer online environments, promoting mental well-being, and mitigating the negative impact of cyberbullying on users.

REFERENCES:

[1] K. S. Alam, S. Bhowmik and P. R. K. Prosun, "Cyberbullying Detection: An Ensemble Based Machine Learning Approach”, 2021
[2] A. Bozyiğit, S. Utku, and E. Nasiboğlu, "Cyberbullying Detection by Using Artificial Neural Network Models", 2019
[3] Ibrohim, M. O., Budi, “A dataset and preliminaries study for abusive language detection in Indonesian social media”, 2018
[4] John Hani, Mohamed Nashaat, Mostafa Ahmed “Social Media Cyberbullying Detection using Machine Learning”, 2019
[5] Kelly Reynolds, April Kontostathis, and Lynne Edwards “Using machine learning to detect cyberbullying”, 2011
[6] Fatma Elsafoury, Stamos Katsigiannis and Zeeshan Pervez “When the Timeline Meets the Pipeline: A Survey on Automated Cyberbullying Detection”, 2021
[7] Nabi Rezvani and Alireza Tabebordbar “Linking Textual and Contextual Features for Intelligent Cyberbullying Detection in Social Media”, 2020
[8] Chikashi Nobata, Joel Tetreault, Achint Thomas, Yashar Mehdad, and Yi Chang, “Abusive language detection in online user content,” 2016
[9] T. T. A. Putri, S. Sriadhi, R. D. Sari, R. Rahmadani, and H. D. Hutahaean” A comparison of classification algorithms for hate speech detection”, 2020
[10] Zahoor, S., & Rohilla, R. “Twitter sentiment analysis using machine learning algorithms: a case study”, 2020

Web References:

1. https://www.kaggle.com/
2. https://www.researchgate.net/
3. https://towardsdatascience.com/
4. https://www.ncbi.nlm.nih.gov/
5. https://www.discoverdatascience.org/
6. https://www.semanticscholar.org/
7. https://www.igi-global.com/
8. https://www.mastersindatascience.org/
9. https://www.analyticsinsight.net/