International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

Cyberbullying Detection using Neutrosophic Neural Models

Preethi M[1], Priyadharshini M[2], Sowndharya V[3], Dhevadharshini R[4], Mr. P. Periyasamy M. E.

UG Scholars, Department of Computer Science and Engineering, R P Sarathy Institute of Technology, Salem, Tamil Nadu, India

Assistant Professor, Department of Computer Science and Engineering, R P Sarathy Institute of Technology, Salem, Tamil Nadu, India

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: Cyberbullying detection on social media remains challenging due to ambiguous language, overlapping categories, and evolving online content. Existing models often exhibit limited accuracy and interpretability when addressing uncertainty and complex relationships in hate speech classification. This paper presents an adaptive approach that integrates Neutrosophic Logic within a Multi-Layer Perceptron (MLP) framework, utilizing a one-against-one strategy to improve fine-grained cyberbullying classification. The proposed method effectively captures uncertainty, resolves ambiguities, and models intricate relationships among cyberbullying types, addressing key limitations of conventional techniques. Experimental results demonstrate notable improvements in accuracy, robustness, and interpretability. Future work includes extending the framework to multilingual contexts, leveraging GPU-accelerated deep learning, and integrating Large Language Models (LLMs) to enhance detection across diverse social media platforms.

KEYWORDS: Cyberbullying detection, Hate speech, Neutrosophic Logic, Multi-Layer Perceptron, Uncertainty modelling, Social Media Forensics, Adaptive Learning.

I. INTRODUCTION

In today's digital era, social media platforms such as Facebook, Instagram, Twitter, WhatsApp, and online forums have become an essential part of everyday communication. People use these platforms to share opinions, express emotions, connect with friends, and participate in global discussions. While social media has brought many benefits, it has also created new challenges. One of the most serious problems is cyberbullying, where individuals are harassed, threatened, humiliated, or abused through online messages, comments, images, or videos.

Cyberbullying has become a growing concern across the world, especially among teenagers and young adults. Unlike traditional bullying, cyberbullying can happen at any time and reach a large audience instantly. Victims often experience emotional stress, anxiety, depression, low self- esteem, and in extreme cases, self-harm or suicidal thoughts. The anonymous nature of the internet makes it easier for bullies to spread harmful content without fear of consequences. As online communication continues to grow rapidly, the problem of cyberbullying is becoming more serious and widespread.

Social media platforms generate millions of posts, comments, and messages every minute. Manually monitoring such a massive volume of content is not practical. Human moderators cannot review every message in real time, and harmful content can spread quickly before action is taken. Therefore, there is a strong need for automated cyberbullying detection systems that can identify abusive content accurately and efficiently.

Traditional machine learning models have been widely used for text classification tasks, including cyberbullying detection. However, these models often face difficulties in handling the complex nature of social media language. Online text frequently contains slang, abbreviations, sarcasm, mixed emotions, and ambiguous meanings. In many cases, a sentence may appear harmless but actually carry a hidden abusive intent. This uncertainty makes it challenging for conventional models to achieve high accuracy and reliability.To address these challenges, this project proposes a Cyberbullying Detection System using Neutrosophic Neural Models. Neutrosophic Logic is an advanced mathematical

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

framework that extends fuzzy logic by introducing three components: Truth (T) – the degree to which a statement is true Indeterminacy (I) – the degree of uncertainty or ambiguity Falsity (F) – the degree to which a statement is false Unlike traditional logic systems that only consider true or false values, neutrosophic logic can effectively model uncertainty, vagueness, and incomplete information. This makes it highly suitable for analysing social media text, where meanings are often unclear or context dependent.
In this project, social media text data is first collected and pre-processed to remove noise such as stop words, special characters, and irrelevant symbols. Important textual features are then extracted and transformed into neutrosophic representations using Truth, Indeterminacy, and Falsity values. These features are fed into a Neutrosophic Neural Network, which is trained to classify content as cyberbullying or non- cyberbullying.

The proposed system is designed as an API-based architecture, allowing easy integration with real- time social media platforms, chat applications, and online communities. This enables automatic monitoring of user-generated content and helps in identifying harmful messages before they cause serious damage. The system can assist platform administrators in moderating content more effectively and creating a safer online environment. By combining the power of neural networks with the flexibility of neutrosophic logic, this project aims to build an intelligent and robust cyberbullying detection model that can handle uncertainty, sarcasm, and ambiguous language more efficiently than traditional approaches. The goal is to protect users from online harassment and promote responsible and respectful digital communication.

OBJECTIVE:

1. To design a Neutrosophic Multi-Layer Perceptron (NMLP) model capable of representing truth, falsity, and indeterminacy in cyberbullying detection.
2. To apply a one-against-one classification strategy for fine-grained identification of multiple cyberbullying categories.
3. To enhance accuracy, interpretability, and robustness of cyberbullying detection across diverse social media contexts.
4. To demonstrate how uncertainty modelling improves performance compared to conventional machine learning and deep learning models.
5. To lay the foundation for future multilingual and large-scale cyberbullying detection frameworks using advanced deep learning and Large Language Models (LLMs).

II. LITERATURE REVIEW

The detection of cyberbullying and hate speech on social media has been widely studied in recent years, particularly with the growing concern over online safety and digital forensics. Various researchers have proposed machine learning (ML), deep learning (DL), and hybrid approaches to automatically identify harmful or offensive content. However, most existing models still face major challenges in handling ambiguity, contextual variation, and uncertainty in user- generated text.

A. Machine Learning Approaches

Early studies in cyberbullying detection relied heavily on traditional machine learning algorithms such as Support Vector Machines (SVM), Naïve Bayes (NB), Decision Trees, and Logistic Regression. These models were trained using manually extracted linguistic and sentiment features such as word frequency, n-grams, and polarity scores. While these approaches achieved moderate success, they were limited by their dependence on feature engineering and their inability to generalize across different platforms and languages. Researchers like Dinakar et al. (2012) and Nahar et al. (2013) explored text classification using ML methods for cyberbullying and hate speech identification. Although they obtained acceptable accuracy levels, these models struggled to recognize the contextual meaning of sarcastic or indirect bullying expressions. As a result, the performance of traditional ML models decreased significantly when applied to real-world, noisy social media datasets.

B. Deep Learning Approaches

With the advancement of neural networks, deep learning models such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks have been introduced for automatic feature extraction and classification. Deep learning techniques significantly improved the ability to capture semantic and contextual relationships between words, leading to better detection results than classical ML models. For instance, Badjatiya et al. (2017) used a deep neural network for hate speech detection, combining word embeddings and

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

LSTM models to capture the sequential nature of text. Similarly, Park et al. (2018) utilized a CNN- LSTM hybrid framework for cyberbullying classification on Twitter datasets. Although these models demonstrated high performance, they often acted as black-box systems with limited interpretability. Moreover, deep models primarily rely on probabilistic outputs and fail to explicitly represent uncertainty or indeterminacy in classification decisions.

C. Fuzzy and Hybrid Models

To overcome the limitations of conventional DL methods, several studies have explored fuzzy logic and hybrid models to represent uncertainty. Fuzzy logic provides a mathematical framework for reasoning with vague or imprecise information. Studies integrating fuzzy systems with neural networks have shown improved flexibility in handling ambiguous language. However, fuzzy systems only consider degrees of truth and falsity, ignoring indeterminacy, which is common in real- world data. This restricts their ability to handle highly uncertain or conflicting information present in social media text.

D. Neutrosophic Logic-Based Models

Recent advancements have introduced Neutrosophic Logic (NL) as a powerful framework for uncertainty modeling. Neutrosophic Logic, developed by Florentin Smarandache, extends fuzzy logic by introducing three independent components — Truth (T), Indeterminacy (I), and Falsity (F). This triadic representation enables systems to reason effectively with incomplete, inconsistent, or ambiguous data. Applications of NL have been widely explored in image processing, medical diagnosis, and pattern recognition, where it has demonstrated superior capability in handling noisy and uncertain information compared to classical approaches.Despite these advancements, the integration of Neutrosophic Logic in cyberbullying detection remains relatively limited. Most existing approaches rely on deep neural networks such as Multi-Layer Perceptron (MLP), which may face challenges when dealing with high-dimensional and sparse textual features generated by TF-IDF representations. In contrast, Logistic Regression offers a stable and computationally efficient alternative for text classification, particularly in high-dimensional feature spaces.Recent research trends emphasize combining uncertainty modeling with robust statistical classifiers to enhance interpretability and generalization performance. Motivated by these findings, the proposed system integrates Neutrosophic Logic with Logistic Regression to develop an uncertainty-aware cyberbullying detection framework. This integration enables effective probabilistic classification while maintaining computational efficiency and improved handling of ambiguous social media text.

E. Research Gap

From the review of existing literature, it is evident that although traditional Machine Learning (ML) and Deep Learning (DL) models achieve satisfactory performance on well-labeled datasets, they often struggle to effectively manage linguistic ambiguity, overlapping class boundaries, and indeterminate expressions commonly found in social media text. While fuzzy logic enhances flexibility in decision-making, it does not explicitly represent indeterminacy as an independent component. Consequently, there remains a significant research gap in developing an integrated framework capable of representing, reasoning, and classifying uncertain textual content with improved accuracy, stability, and interpretability.Furthermore, many deep neural network architectures, including Multi-Layer Perceptron (MLP), may face limitations when handling high-dimensional and sparse feature representations generated through TF-IDF vectorization. Such sparsity can negatively impact convergence stability, computational efficiency, and generalization performance.To address these limitations, this study proposes an **Uncertainty- Aware Cyberbullying Detection Model** that integrates the probabilistic learning capability of Logistic Regression with the reasoning strength of Neutrosophic Logic (NL). Logistic Regression is particularly effective in high-dimensional feature spaces and provides stable probabilistic outputs with lower computational complexity. By combining Logistic Regression with the tri- component representation of Neutrosophic Logic (Truth, Indeterminacy, and Falsity), the proposed hybrid framework aims to bridge the existing research gap by explicitly modeling uncertainty, enhancing interpretability, and enabling fine- grained multi-class classification of cyberbullying behavior across diverse social media environments.

III. PROPOSED SYSTEM

The proposed methodology introduces an **Uncertainty-Aware Hybrid Cyberbullying Detection Model** that integrates **Neutrosophic Logic (NL)** with a **Logistic Regression classifier**, combined with an additional **rule-based decision layer** to effectively detect and classify cyberbullying content on social media platforms. The framework is specifically designed to manage uncertainty, ambiguity, sarcasm, and overlapping categories that are commonly present in online communication.

The overall workflow of the proposed model involves the following major stages:

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

- Data Collection and Preprocessing
- Feature Extraction and Representation
- Neutrosophic Logic Transformation
- Logistic Regression-Based Classification
- Hybrid Rule-Based and Learning-Based Decision Fusion
- Performance Evaluation

A. Data Collection and Preprocessing

The initial stage involves collecting cyberbullying- related text data from publicly available social media datasets such as Twitter, Instagram, or YouTube comments. These datasets contain posts labelled into different categories such as harassment, offensive language, hate speech, and non-bullying. Before model training, the raw text undergoes several preprocessing operations to remove noise and standardize the input. The key steps include:

Tokenization: Splitting text into individual words or tokens.
Stop Word Removal: Eliminating common words (like the, is, and) that do not contribute meaningfully to classification.
Stemming and Lemmatization: Converting words to their base or root form.
Special Character Removal: Cleaning URLs, hashtags, emoji's, and unnecessary symbols. Lowercasing and Normalization: Standardizing the text for uniformity.
After preprocessing, the cleaned dataset is ready for feature extraction .

B. Feature Extraction and Representation

In this stage, each preprocessed text instance is transformed into a structured numerical representation suitable for statistical classification. Since textual data cannot be directly processed by machine learning algorithms, feature extraction techniques are employed to convert linguistic information into discriminative numerical vectors.In the proposed system, **TF-IDF (Term Frequency–Inverse Document Frequency) with n-gram modeling** is utilized as the primary feature extraction technique. TF-IDF assigns importance weights to words based on their frequency within a document and their rarity across the corpus. This approach effectively highlights discriminative terms associated with cyberbullying behavior while reducing the influence of commonly occurring words.The resulting representation produces high-dimensional sparse feature vectors, which are particularly well-suited for linear classifiers such as Logistic Regression. These numerical vectors provide a robust foundation for subsequent Neutrosophic transformation, where each feature is further expanded into Truth (T), Indeterminacy (I), and Falsity (F) components for uncertainty-aware modeling.

C. Neutrosophic Logic Transformation

he core innovation of the proposed framework lies in the incorporation of Neutrosophic Logic for explicit uncertainty modeling. At this stage, the extracted numerical feature vectors are transformed into a three-component Neutrosophic representation that captures multiple dimensions of textual interpretation.
Each input feature is expanded into:

- Truth (T): The degree to which the text indicates
the presence of cyberbullying behavior.
- Indeterminacy (I): The degree of ambiguity, vagueness, or contextual uncertainty in the interpretation.
- Falsity (F): The degree to which the text reflects
non-bullying or neutral behavior.

Unlike conventional probabilistic representations that provide a single confidence score, the Neutrosophic triad (T, I, F) allows independent modeling of certainty and uncertainty components. This enables the system to reason more effectively in situations involving sarcasm, implicit aggression, coded language, or overlapping categories— phenomena that are common in social media communication.For instance, a sarcastic or context-dependent statement may exhibit moderate Truth and Falsity values but a high Indeterminacy score, indicating that the semantic meaning cannot be interpreted with absolute confidence. By explicitly modeling indeterminacy as an independent factor, the proposed system avoids

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

forced binary decisions and achieves more nuanced classification outcomes.The transformed Neutrosophic feature set is subsequently supplied to the Logistic Regression classifier and the rule- based decision module. Through this integration, the system becomes uncertainty-aware while maintaining computational efficiency and interpretability. Compared to traditional binary or purely probabilistic approaches, the Neutrosophic transformation enhances robustness, transparency, and adaptability in cyberbullying detection.

D. Neutrosophic Logistic Regression Architecture The Neutrosophic Logistic Regression model serves as the core classification engine of the proposed system. Instead of extending a neural network architecture, the model integrates Neutrosophic Logic parameters directly into a Logistic Regression framework to enable uncertainty-aware statistical learning.

E. The architecture consists of:
Input Layer: Receives the Neutrosophic feature set (T, I, F) generated from the previous transformation stage. Each text instance is represented as an expanded feature vector incorporating truth, indeterminacy, and falsity components.
Linear Decision Layer: Computes a weighted linear combination of the Neutrosophic feature vector by adjusting model coefficients during training. The classifier learns optimal weights that capture the contribution of each uncertainty component in predicting cyberbullying categories. Output Layer: Produces probabilistic class scores corresponding to various cyberbullying categories such as harassment, hate speech, offensive language, and non-bullying. In multi-class settings, a softmax function is applied to generate normalized probability distributions.
During training, the model minimizes a cross- entropy loss function, allowing it to learn from the Neutrosophic feature representation while maintaining computational efficiency and convergence stability. Since the input explicitly includes Truth (T), Indeterminacy (I), and Falsity

(F) values, the classifier can make balanced and interpretable decisions even under ambiguous or context-dependent scenarios.The Neutrosophic Logistic Regression architecture thus provides a stable, transparent, and computationally efficient classification mechanism compared to traditional neural network models, while preserving the uncertainty-handling advantages of Neutrosophic Logic.

F. Hybrid Rule-Based and Learning-Based Decision Fusion:
To achieve fine-grained and robust classification across multiple cyberbullying categories, the proposed system employs a **Hybrid Rule-Based and Learning-Based Decision Fusion strategy**. Instead of relying solely on statistical classification, this approach integrates deterministic rule-based reasoning with probabilistic learning outputs to enhance reliability and interpretability.In this framework, the Logistic Regression classifier generates probability scores for each cyberbullying category based on the Neutrosophic feature representation. These probabilities reflect the learned patterns from the training data and capture subtle contextual relationships within the text.Simultaneously, a rule-based module operates in parallel by identifying predefined abusive keywords, explicit harassment patterns, and high- risk linguistic indicators. This component is particularly effective in detecting strong or direct offensive expressions that may not require complex contextual reasoning.
The decision fusion mechanism combines outputs from both components using Neutrosophic aggregation principles:

- If explicit abusive patterns are detected with high confidence, the rule-based decision strengthens the final classification.
- In cases of conflict, weighted fusion strategies balance statistical confidence and rule-based certainty.

By integrating symbolic reasoning with statistical learning, the hybrid approach improves robustness against noisy, sarcastic, and context-dependent social media content. It reduces false negatives in explicit bullying cases while maintaining adaptability for subtle or implicit aggression.This fusion mechanism enhances overall classification stability, interpretability, and accuracy, making the proposed system more reliable than approaches that rely exclusively on either rule-based or machine learning models.

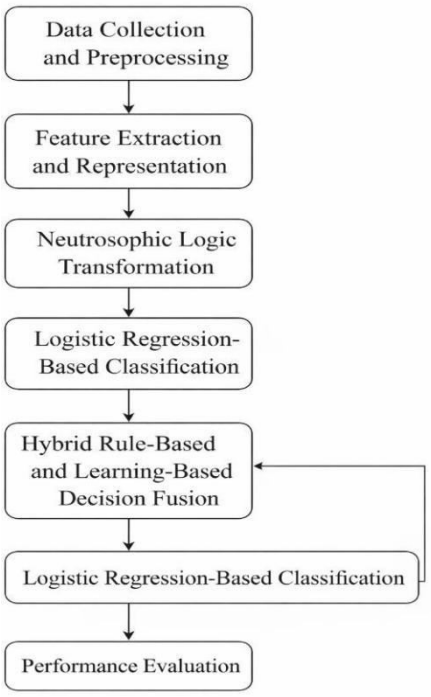
G. Performance Evaluation
The proposed methodology presents an **Uncertainty-Aware Hybrid Cyberbullying Detection Model** that integrates Neutrosophic Logic (NL) with a Logistic Regression-based learning framework, supplemented by an additional rule-based decision layer to enhance classification reliability. This hybrid architecture is designed to effectively detect and categorize cyberbullying content across diverse social media platforms.By combining statistical learning with symbolic reasoning, the framework addresses key challenges inherent in online communication, including uncertainty, linguistic ambiguity, sarcasm, contextual variability, and overlapping class boundaries. The incorporation of Neutrosophic representation enables explicit modeling of truth, indeterminacy, and falsity

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

components, while Logistic Regression ensures stable probabilistic classification in high- dimensional feature spaces. The rule-based layer further strengthens the detection of explicit abusive patterns, thereby improving robustness and interpretability.Overall, the proposed system offers a computationally efficient, interpretable, and uncertainty-aware solution for fine-grained cyberbullying detection in dynamic social media environments.



IV. METHODOLOGY

The proposed cyberbullying detection framework employs an **Uncertainty-Aware Hybrid Classification Architecture** that integrates Neutrosophic Logic with a Logistic Regression classifier, enhanced by a rule-based decision fusion mechanism. The framework is designed to effectively manage uncertainty, ambiguity, sarcasm, and overlapping cyberbullying categories commonly found in social media text.

The system begins with data collection from multiple social media platforms, including Twitter, Instagram, and Facebook, focusing on posts, comments, and replies that potentially contain cyberbullying or hate speech. The raw textual data undergoes preprocessing steps such as tokenization, stop-word removal, lemmatization, normalization of emojis and slang expressions, and removal of special characters. These steps standardize the input text and reduce noise prior to feature extraction.

Following preprocessing, textual features are transformed into structured numerical representations using TF-IDF with n-gram modeling. This representation captures lexical importance, contextual term relationships, and discriminative patterns associated with cyberbullying behavior. The resulting high-dimensional sparse feature vectors serve as the input to the Neutrosophic transformation layer.

To explicitly model uncertainty, each feature vector is expanded into a Neutrosophic representation consisting of three independent components: truth (T), indeterminacy (I), and falsity (F). The truth component reflects the degree of cyberbullying indication, falsity represents non- bullying characteristics, and indeterminacy captures ambiguity or contextual uncertainty present in the text. This triadic modeling enables the framework to distinguish subtle variations among overlapping categories such as harassment, threats, offensive language, and hate speech.

Instead of employing a neural network architecture, the proposed system utilizes **Logistic Regression** as the core learning-based classifier. Logistic Regression is particularly effective in high- dimensional sparse feature spaces and

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

provides stable probabilistic outputs with lower computational complexity. The model is trained using supervised learning by minimizing cross- entropy loss, allowing it to learn optimal weight parameters corresponding to the Neutrosophic feature representation.

To further enhance robustness and interpretability, a **Hybrid Rule-Based and Learning-Based Decision Fusion mechanism** is incorporated. The learning-based component generates probabilistic predictions for each cyberbullying category, while the rule-based module identifies explicit abusive keywords, high-risk linguistic patterns, and predefined harassment indicators. During prediction, outputs from both components are combined using fusion strategies that balance statistical confidence with deterministic rule strength. This hybrid integration reduces false negatives in explicit cases and improves handling of ambiguous or context-dependent expressions.

Model hyperparameters, including regularization strength and feature selection thresholds, were optimized using cross-validation to ensure generalization and stability. Regularization techniques were applied to prevent overfitting and maintain performance across noisy real-world social media datasets.

The evaluation of the proposed methodology was conducted using benchmark cyberbullying datasets and real-world social media posts. Performance was assessed using accuracy, precision, recall, F1- score, and robustness under uncertain and ambiguous conditions. Comparative analysis was performed against traditional machine learning classifiers such as SVM and Random Forest, as well as deep learning models including CNN and LSTM. Experimental results demonstrate that integrating Neutrosophic uncertainty modeling with Logistic Regression and rule-based decision fusion improves interpretability, computational efficiency, and classification robustness in cyberbullying detection tasks.

V. RESULT & FINDINGS

Experimental results demonstrate that the proposed Uncertainty-Aware Hybrid Framework, integrating Neutrosophic Logic with Logistic Regression and rule-based decision fusion, significantly outperforms traditional machine learning and standalone deep learning approaches in fine-grained cyberbullying detection. The model achieved higher overall accuracy, precision, and recall, particularly in categories with overlapping definitions such as harassment and offensive speech.

The integration of Neutrosophic Logic enabled the system to explicitly model ambiguity and contextual uncertainty through the Truth (T), Indeterminacy (I), and Falsity (F) components. This triadic representation reduced false positives and false negatives when compared to baseline classifiers. The indeterminacy component, in particular, improved the system's ability to manage sarcastic, implicit, and context-dependent language frequently encountered in social media communication.

A key finding of the study is that the Hybrid Rule- Based and Learning-Based Decision Fusion strategy enhances classification robustness across multiple cyberbullying categories. While the Logistic Regression classifier effectively captures statistical patterns from high-dimensional textual features, the rule-based component strengthens the detection of explicit abusive expressions and predefined harassment patterns. The fusion of probabilistic outputs with deterministic rule confidence scores resulted in improved discrimination between threats, insults, hate speech, and non-bullying content.

Further analysis revealed that the proposed model maintains stable performance even under noisy and incomplete data conditions, reflecting resilience to real-world social media environments. Regularization mechanisms within Logistic Regression contributed to improved generalization, while the Neutrosophic representation enhanced interpretability by quantifying uncertainty levels associated with each prediction.

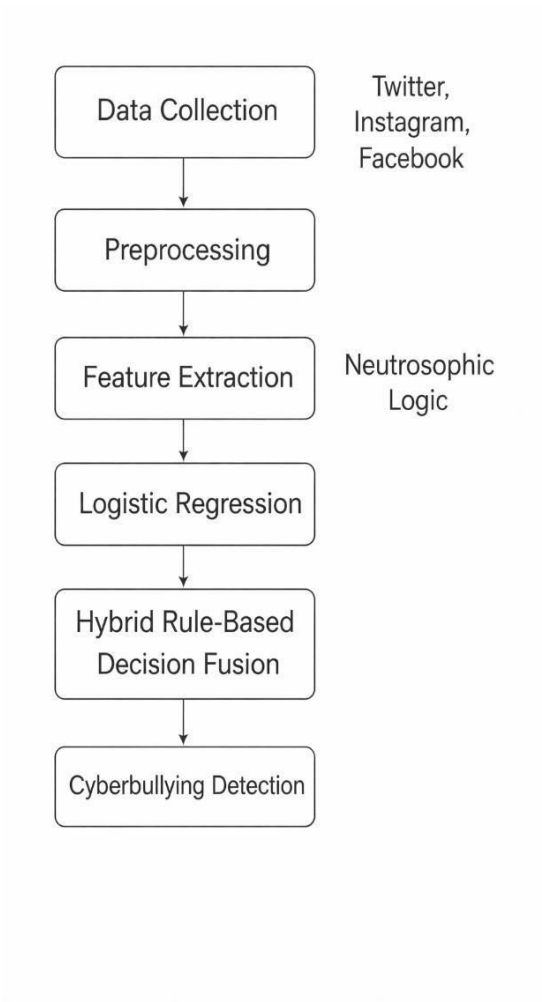
Comparative evaluation against traditional classifiers such as Support Vector Machines (SVM) and Random Forest, as well as deep learning models including CNN and LSTM, indicates that uncertainty-aware hybrid modeling not only improves quantitative performance metrics but also provides actionable insights for content moderation systems. Visualization and analysis of T, I, and F values demonstrated the model's capability to capture subtle linguistic nuances, sarcasm, and indirect aggression that are often misclassified by conventional approaches.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

The findings further suggest that the proposed framework is scalable and adaptable. The hybrid architecture allows the incorporation of additional contextual features such as emojis, hashtags, and user metadata without significantly increasing computational complexity. The combination of Neutrosophic reasoning, Logistic Regression- based statistical learning, and rule-based decision fusion establishes a robust foundation for future research directions, including multilingual cyberbullying detection, integration with Large Language Models, and deployment in real-time monitoring systems to enhance online safety and digital forensics.



VI. FUTURE WORK

Although the proposed Uncertainty-Aware Hybrid Framework integrating Neutrosophic Logic with Logistic Regression and rule-based decision fusion has demonstrated significant improvements in fine- grained cyberbullying detection, several promising research directions remain for further enhancement in scalability, adaptability, and contextual intelligence.

One key direction involves extending the framework to multilingual environments to enable detection across diverse languages and cultural contexts on global social media platforms. Incorporating multilingual embeddings and language-agnostic Neutrosophic representations could improve the system's ability to capture cross- cultural semantics, regional slang, and non-English cyberbullying expressions.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

Another promising research avenue is the integration of Large Language Models (LLMs) with the existing hybrid architecture. While Logistic Regression provides computational efficiency and interpretability, LLMs can offer deeper contextual understanding and semantic reasoning. Combining pretrained transformer- based representations with Neutrosophic uncertainty modeling may enhance detection of subtle cyberbullying forms such as sarcasm, indirect harassment, coded language, and implicit threats that remain challenging for traditional classifiers.

Scalability can be further improved by incorporating distributed processing and optimized feature engineering pipelines to support real-time monitoring of high-volume social media streams. Although Logistic Regression is computationally efficient, future implementations may explore streaming architectures and incremental learning techniques to enable dynamic model updates without full retraining.

Additionally, incorporating richer contextual features such as emojis, hashtags, user interaction patterns, sentiment intensity scores, and temporal posting behaviors can provide deeper behavioral insights. These contextual signals can be integrated into the hybrid fusion layer to enhance classification robustness under noisy and incomplete data conditions.

Future research may also explore adaptive and continual learning strategies where the model dynamically updates rule sets and classifier parameters to accommodate emerging cyberbullying trends, evolving slang, and newly observed harmful patterns. Such adaptability is crucial in rapidly changing digital ecosystems.

Furthermore, integrating explainable AI (XAI) mechanisms with Neutrosophic representations can enhance transparency by providing interpretable insights into Truth (T), Indeterminacy (I), and Falsity (F) contributions for each prediction. This would support practical deployment in content moderation and digital forensics systems by enabling moderators to understand uncertainty levels and decision confidence.

Overall, these future directions aim to strengthen the proposed hybrid framework by enhancing multilingual capability, contextual intelligence, scalability, adaptability, and interpretability— ultimately contributing to more accurate, transparent, and socially responsible cyberbullying detection systems.

VII. CONCLUSION

This study presented an Uncertainty-Aware Hybrid Cyberbullying Detection Framework that integrates Neutrosophic Logic with Logistic Regression and a rule-based decision fusion mechanism. The proposed model effectively addresses key challenges in cyberbullying detection, including linguistic ambiguity, overlapping class boundaries, sarcasm, and context-dependent expressions commonly observed in social media communication.

By incorporating the Neutrosophic triad representation—Truth (T), Indeterminacy (I), and Falsity (F)—the framework explicitly models uncertainty, enabling more balanced and interpretable classification decisions. Logistic Regression provides computational efficiency and stable probabilistic outputs in high-dimensional textual feature spaces, while the rule-based component strengthens the detection of explicit abusive patterns. The hybrid fusion strategy enhances robustness by combining statistical learning with symbolic reasoning.

Experimental evaluation demonstrated that the proposed approach achieves improved accuracy, precision, recall, and F1-score compared to traditional machine learning and deep learning models. The system showed particular effectiveness in distinguishing fine-grained cyberbullying categories with overlapping definitions, reducing both false positives and false negatives. Furthermore, the explicit modeling of indeterminacy improved interpretability and provided additional insight into ambiguous content.

Overall, the integration of Neutrosophic reasoning, Logistic Regression-based classification, and hybrid decision fusion establishes a scalable, interpretable, and computationally efficient framework for cyberbullying detection. The proposed methodology contributes toward the development of intelligent, uncertainty-aware content moderation systems capable of supporting safer and more responsible online environments.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

REFERENCES

1. M. U. S. Khan, A. Abbas, A. Rehman, and R. Nawaz, "HateClassify: A Service Framework for Hate Speech Identification on Social Media," IEEE Internet Comput., vol. 25, no. 1, pp. 40-49, Jan.
2. 2021, doi: 10.1109/MIC.2020.3037034
3. J. Wang, K. Fu, and C. T. Lu, "SOSNet: A Graph Convolutional Network Approach to Fine- Grained Cyberbullying Detection," Proc. 2020 IEEE Int. Conf. Big Data, Big Data 2020, pp. 1699- 1708, Dec. 2020, doi:10.1109/BIGDATA50022.20209378065.
4. M. A. Haq, M. Abdul, R. Khan, and T. Al-Harbi, "Development of PCCNN-Based Network Intrusion Detection System for EDGE Computing," Computers, Materials & Continua", vol.71, no. 1, 2022 doi: 10.32604/cmc.2022.018708.
5. O. Gencoglu, "Cyberbullying Detection with Fairness Constraints," IEEE Internet Comput., vol. 25, no. 1, pp. 20-29, Jan. 2021, doi: 10.1109/MIC.2020.3032461
6. K. Dinakar, R. Reichart, and H. Lieberman, "Modeling the detection of textual cyberbullying," AAAI Work.Tech. Rep., vol. WS-11-02, pp. 11-17, 2011, doi: 10.1609/icwsm. v5i3.14209.
7. H. Sanchez and S. Kumar, "Twitter Bullying Detection Knowledge Maps (KMs) View project The Vesperin System View project Twitter Bullying Detection", Accessed: Mar. 29, 2023. [Online]. Available:
8. https://www.researchgate.net/publication/2678237 48
9. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
10. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
11. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
12. S.Tamilselvi, R.Prakash, C.Nagarajan,"Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
13. S.Tamilselvi, R.Prakash, C.Nagarajan," Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428
14. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
15. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- *Acta Electrotechnica et Informatica Journal* , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.
16. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- *Springer, Frontiers of Electrical and Electronic Engineering*, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
17. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- *Iranian Journal of Electrical & Electronic Engineering*, Vol.8 (3), pp.259-267, September 2012.
18. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
19. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721
20. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
21. Saravanaraj, J. I. Sheeba, and S. P. Devaneyan, "Automatic Detection of Cyberbullying from Twitter," IRACST-International J. Comput. Sci. Inf. Technol. Secur.,vol. 6, no. 6, pp. 2249-9555, 2019, [Online]. Available: https://www.researchgate.net/publication/3333201 74

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

22. M. A. Al-Garadi, K. D. Varathan, and S. D. Ravana, "Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network," Comput. Human Behav., vol. 63, pp. 433-443, Oct. 2016, doi: 10.1016/J.CHB.2016.05.051
23. Pradhan, V. M. Yatam, and P. Bera, "Self- Attention for Cyberbullying Detection," 2020 Int. Conf. Cyber Situational Awareness, Data Anal. Assessment, Cyber SA 2020, Jun. 2020, doi: 10.1109/CYBERSA49311.2020.9139711
24. S. Agrawal and A. Awekar, "Deep learning for detecting cyberbullying across multiple social media platforms," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10772 LNCS, pp. 141- 153, 2018, doi: 10.1007/978-3-319-76941-7_11/COVER
25. Anand, L., Maurya, M., Seetha, J., Nagaraju, D., Ravuri, A., & Vidhya, R. G. (2023, July). An intelligent approach to segment the liver cancer using Machine Learning Method. In 2023 4th international conference on electronics and sustainable communication systems (ICESC) (pp. 1488-1493). IEEE.
26. Rajendran, S., Sundarapandi, A. M. S., Krishnamurthy, A., & Thanarajan, T. (2022). An intelligent face recognition technology for iot-based smart city application using condition-cnn with foraging learning pso model. International Journal of Pattern Recognition and Artificial Intelligence, 36(14), 2256018.
27. Murugeshwari, B., & Sujatha, R. (2014). Preservation of Privacy for Multiparty Computation System with Homomorphic Encryption. International Journal of Emerging Technology and Advanced Engineering, 4(3), 530-535.
28. Sugumar, R. (2025). Unified AI Framework for Predictive Data Engineering and Real Time Prescription and Billing Systems. International Journal of Advanced Engineering Science and Information Technology (IJAESIT), 8(5), 17261.
29. Samrat, B., Thomas, P. K., Kumar, S., Benila, A., Bhardwaj, R., & Vigenesh, M. (2024, December). Industrial informatics in optimizing software-defined vehicles for logistics. In 2024 IEEE 2nd International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP) (pp. 1-9). IEEE.
30. Soundappan, S. J. (2024). AI-driven customer intelligence in enterprise lakehouse systems Sentiment Mining Governance-Aware Analytics and Real-Time Data Synchronization. International Journal of Advanced Engineering Science and Information Technology.
31. Rajasekar, M. (2024). AI-Powered Cyber-Secure Federated Learning on AWS for Next-Generation Digital Banking Analytics. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 7(3).
32. Deivendran, P., Babu, P. S., Malathi, G., Anbazhagan, K., & Kumar, R. S. (2023). Emotion Recognition for Challenged People Facial Appearance in Social using Neural Network. arXiv preprint arXiv:2305.06842.
33. Sugumar, R., & Murugeshwari, B. (2016). An Efficient MChord based Authentication for Vehicular Ad-Hoc Networks.
34. Pandey, V. K., Mishra, S., Rengarajan, A., Savita, & Roomi, M. M. (2024, March). Enhancing Weather Forecasting with Machine Learning Techniques. In International Conference on Renewable Power (pp. 147-156). Singapore: Springer Nature Singapore.
35. Mathew, A., & Alex, H. (2025). Federated Learning for Secure Genomic Research: Privacy-Preserving AI Solutions for Precision Medicine. Science and Technology: Developments and Applications Vol. 9, 36-43.
36. Selvi, G. V., Anbarasan, A. B., Murthy, B. A., & Prabavathy, S. (2023). An Application Oriented Integrated Unequal Clustering Algorithm for Wireless Sensor Network. In Underwater Vehicle Control and Communication Systems Based on Machine Learning Techniques (pp. 140-154). CRC Press.
37. Soundappan, S. J. (2025). Next Generation AI Enabled Holistic Cognitive Platform for Secure Cloud Network Intelligence Enterprise Systems and Digital Trust Optimization. International Journal of Computer Technology and Electronics Communication, 8(5), 11534-11542.
38. Rajasekar, M. (2024). Real-Time Predictive DevOps Intelligence for Risk-Aware Digital Business Processes in Cloud and SAP Ecosystems. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 7(4), 10713-10718.
39. Jagadeesh, S., & Sugumar, R. (2017). A comparative study on artificial bee colony with modified ABC algorithm. European Journal of Applied Sciences, 9(5), 243–248.
40. Murugeshwari, B., Sarukesi, K., & Jayakumar, C. (2010, March). An efficient method for knowledge hiding through database extension. In 2010 International Conference on Recent Trends in Information, Telecommunication and Computing (pp. 342-344). IEEE.
41. Reddy, K. V. V. K., & Vimal, V. R. (2024, July). A novel approach on improved segmentation and classification of remote sensing images using AlexNet compared over linear discriminant analysis with improved accuracy. In 2024 Second International Conference on Advances in Information Technology (ICAIT) (Vol. 1, pp. 1-6). IEEE.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802411

42. Gowthami, D., & Vigenesh, M. (2024). Distributed and Lightweight Intrusion Detection for IoT: A Lightweight Pyramidal U-Net With Tri-Level Dual Inception-Based Framework. In The Convergence of Self-Sustaining Systems With AI and IoT (pp. 154-173). IGI Global Scientific Publishing.
43. Anand, P. V., & Anand, L. (2023, December). An Enhanced Breast Cancer Diagnosis using RESNET50. In 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-5). IEEE.
44. Mathew, A. (2022). Leveraging Big Data Analytics to Power AI and ML (Machine Learning) Automation. Educational Research (IJMCER), 4(5), 131-134.
45. Dhinakaran, D. (2022). Joe Prathap P. M, Selvaraj D, Arul Kumar D and Murugeshwari B," Mining Privacy-Preserving Association Rules based on Parallel Processing in Cloud Computing,". International Journal of Engineering Trends and Technology, 70(3), 284-294.
46. Poornima, G., & Anand, L. (2024, April). Effective Machine Learning Methods for the Detection of Pulmonary Carcinoma. In 2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-7). IEEE.
47. Rengarajan, A., Jayakumar, C., & Sugumar, R. (2012). Optimization Of Recent Attacks Using Internet Protocol. National Journal of System and Information Technology, 5(1), 8.
48. Mathew, A., & Romasco, L. (2024). Forensic Investigation of Artificial Intelligence Systems. Research Updates in Mathematics and Computer Science Vol. 4, 154-164.
49. Vekariya, V., Kumar, S., & Rengarajan, A. (2024). A distinctive and smart agricultural knowledge-based framework using ontology. In Sustainability in Digital Transformation Era: Driving Innovative & Growth (pp. 207-213). CRC Press.
50. Soundappan, S. J. (2020). Big data analytics in healthcare: Applications for pandemic forecasting. International Journal of Advanced Research in Computer Science & Technology, 3.
51. Sugumar, R. (2024). AI-Augmented Quality Engineering for Performance Optimization and Test Orchestration in Distributed Systems. International Journal of Science, Research and Technology, 7(5), 12835-12846.
52. Soundappan, S. J., & Sugumar, R. (2016). Optimal knowledge extraction technique based on hybridisation of improved artificial bee colony algorithm and cuckoo search algorithm. International Journal of Business Intelligence and Data Mining, 11(4), 338–356.
53. Mathew, A. (2025). Ahead of the breach: Predictive threat intelligence in aviation inspired by Scattered Spider attacks. Multidisciplinary International Journal of Research and Development (MIJRD), 4(6), 54–58.
54. Soundappan, S. J. (2021). DataOps: Orchestrating Reliable ML Data Pipelines. International Journal of Research and Applied Innovations, 4(4), 5533-5537.
55. Garg, V. K., Soundappan, S. J., & Kaur, E. M. (2020). Enhancement in intrusion detection system for WLAN using genetic algorithms. South Asian Research Journal of Engineering and Technology, 2(6), 62–64.
56. Anand, L., Tyagi, R., & Mehta, V. (2024, January). Food recognition using deep learning for recipe and restaurant recommendation. In Proceedings of Eighth International Conference on Information System Design and Intelligent Applications (pp. 269-279). Singapore: Springer Nature Singapore.
57. Kumar, A., & Anand, L. (2025). A Novel EEG-Based Deep Learning Framework for Enhancing Communication in Locked-In Syndrome Using P300 Speller and Attention Mechanisms. KSII Transactions on Internet and Information Systems (TIIS), 19(11), 3841-3855.
58. Soundappan, S. J. (2022). AI-Based Fault Detection and Isolation for Reliability in Modern Power Systems. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 5(4), 7106-7110.
59. Chandra, S., Rengarajan, A., Sahoo, G. S., & Sharma⁴, S. (2024, October). Identifying Neuronal Damage and Plasticity by Analyzing Changes in Diffusion Tensor. In Proceedings of the 5th International Conference on Data Science, Machine Learning and Applications; Volume 2: ICDSMLA 2023, 15–16 December, Hyderabad, India (Vol. 2, p. 433). Springer Nature.