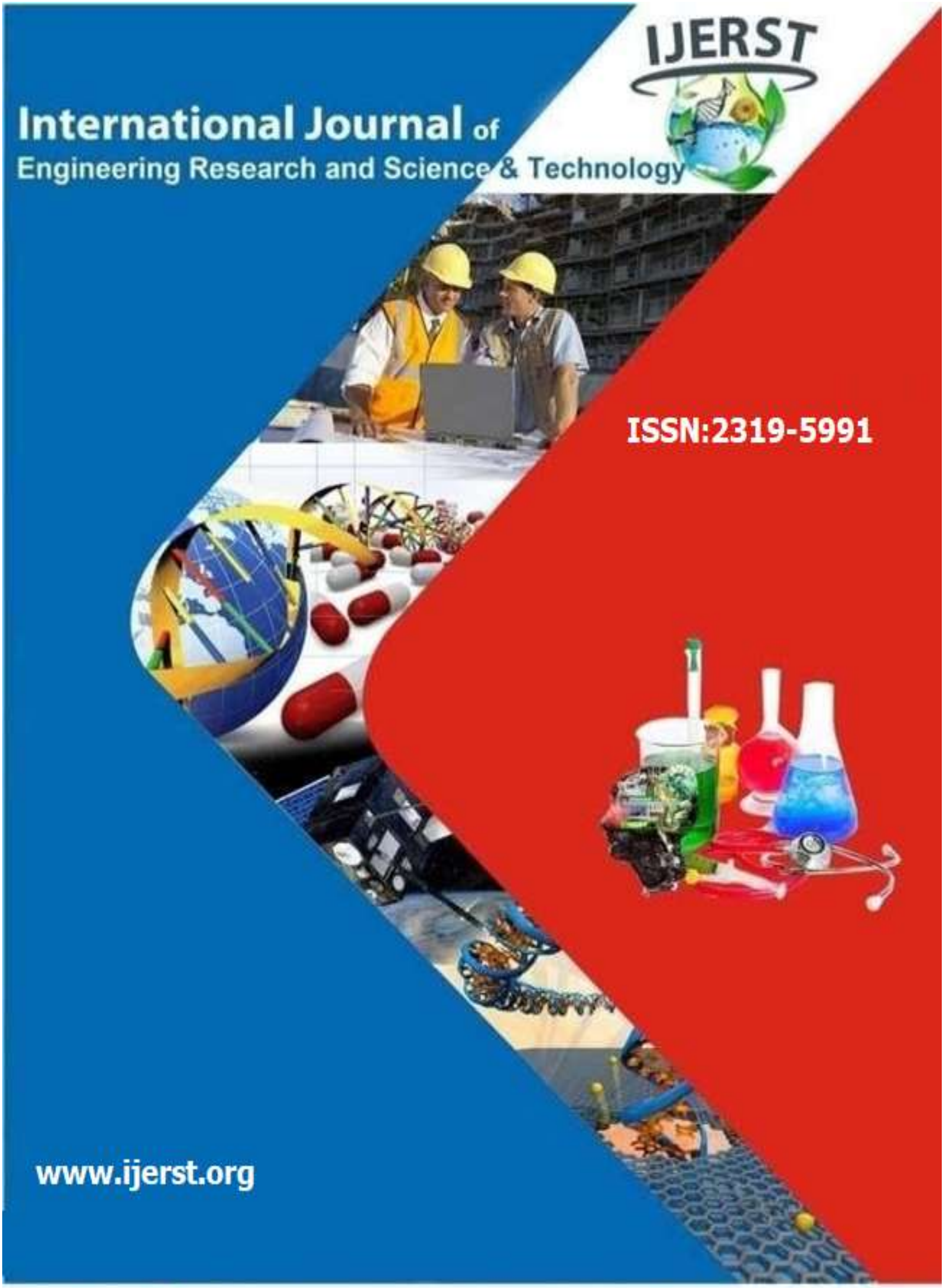
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025



E-mail: editor@ijerst.org or ijerst.editor@gmail.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

BULLYING AWARE :AN INTELLIGENT SYSTEM FOR EARLY DETECTION OF CYBERBULLYING USING MACHINE LEARNING

Dr.V.Krishna[1]
Asst. Professor
Department of CSE(DS)
TKR College of Engineering & Technology

V Sai Teja [2]
B.Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering & Technology
vadlasaiteja2403@gmail.com

G Manasa[3]
B.Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering & Technology
ganganimanasa120gmail.com

MD Mafaiz Uddin [4]
B.Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering & Technology
mafaizuddinm@gmail.com

M Suraj Kumar [5]
B.Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering & Technology
mnsurajkumar@gmail.com

ABSTRACT

Cyberbullying has become a pervasive issue in the digital era, particularly affecting adolescents and young adults on social media platforms. Conventional detection methods, which rely on manual reporting, are often slow, inconsistent, and ineffective. To the address this challenge, BullyingAware introduces an intelligent machine learning-based system designed to automatically identify cyberbullying content with high accuracy. Leveraging natural language processing (NLP) and supervised learning algorithms, the system processes textual data through preprocessing techniques such as tokenization, stemming, and stopword removal. Feature extraction methods, including TF-IDF and word embeddings, enhance the model's ability to detect subtle linguistic cues, sarcasm, and contextual aggression.The proposed system is trained on diverse datasets containing various forms of abusive language, ensuring adaptability across different platforms.The Experimental evaluations to the demonstrate superior performance in detecting cyberbullying, with high precision and recall rates compared to traditional approaches.For the Future enhancements include real-time integration with social media platforms and extending detection capabilities to the in the multimedia multimedia content such as images and videos. By automating early detection, BullyingAware offers a scalable and efficient solution to combat cyberbullying, promoting safer online interactions and mitigating its psychological impact on victims.

Keywords:*Cyberbullying Detection, Machine Learning, NLP, Deep Learning, BERT, Real-Time Monitoring, Text Classification, Sentiment Analysis.*

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

I.INTRODUCTION

In today's digitally connected world, cyberbullying has become a pressing concern, with harmful online interactions severely impacting mental health and social well-being. Conventional detection methods, which depend on manual reporting and human moderation, are often slow, inconsistent, and unable to scale with the vast volume of online content. Artificial intelligence and machine learning present a transformative solution by enabling automated, real-time identification of cyberbullying through advanced text analysis. By leveraging natural language processing (NLP) and deep learning techniques, intelligent systems can detect abusive language, aggressive behavior, and hidden patterns in online communication more efficiently than traditional approaches.

Modern cyberbullying detection systems employ sophisticated NLP methodologies, including sentiment analysis, contextual word embeddings, and semantic understanding, to classify harmful content across diverse digital platforms. However, challenges such as sarcasm, slang, and contextdependent aggression complicate accurate detection. To address these issues, this study introduces BullyingAware, a machine learningbased framework that combines supervised learning models with deep learning architectures like BERT and LSTMs. The system utilizes advanced text representation techniques—such as TF-IDF, Word2Vec, and transformer-based embeddings—to capture linguistic nuances effectively. Additionally, robust data preprocessing (tokenization, lemmatization, and stop-word removal) and data augmentation strategies enhance model performance, while class imbalance techniques ensure fair detection across datasets. This paper presents the importance development, implementation, and evaluation of BullyingAware, demonstrating its high accuracy in identifying cyberbullying across multiple online platforms. The study also explores future enhancements, including real-time monitoring, multilingual support, and multimodal analysis (image/video-based detection), to create a more comprehensive and scalable solution. By this integrating cutting-edge AI technologies, this research aims to foster safer digital spaces and mitigate the growing threat of online harassment.

II.RELATED WORK

Cyberbullying detection has evolved significantly as researchers strive to create safer online environments.

Early approaches depended on basic keyword filtering and rule-based systems, which proved inadequate in identifying nuanced forms of harassment, such as contextual aggression or veiled insults. The availability of benchmark datasets like the Cyberbullying Detection Dataset and Kaggle's Toxic Comment Classification Challenge has enabled more data-driven research, though these collections often lack diversity in language, cultural context, and platform-specific communication styles.

Traditional machine learning models, including Support Vector Machines (SVMs), Naïve Bayes, and Random Forests, were initially adopted for text classification in cyberbullying detection. While effective for explicit harassment, these models struggled with semantic subtleties and evolving online slang. The emergence of deep learning introduced more powerful architectures, such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), which improved contextual understanding by analyzing sequential and spatial text patterns. The advent of transformer-based models like BERT, RoBERTa, and GPT further revolutionized the field, enabling superior performance in detecting implicit bullying through self-attention mechanisms and bidirectional context modeling.

Preprocessing techniques—such as tokenization, lemmatization, and stop-word removal—have been critical in refining input data quality, while feature extraction methods like TF-IDF and word embeddings (Word2Vec, GloVe) enhance semantic

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

representation. Researchers have also tackled class imbalance issues using oversampling (SMOTE), undersampling, and synthetic data generation to prevent bias toward majority classes. Despite these advancements, challenges persist in detecting sarcasm, coded language, and cross-cultural variations in abusive content.

Recent studies emphasize multimodal approaches, combining text with visual and audio data for comprehensive detection, as well as real-time analysis for immediate intervention. This work builds upon these innovations by integrating transformer-based NLP models with robust preprocessing pipelines and imbalance mitigation strategies. The proposed system not only addresses current detection gaps but also prioritizes scalability and adaptability for deployment across diverse digital platforms, advancing the frontier of automated cyberbullying prevention.

III.METHODOLOGY 1.Dataset Acquisition & Preprocessing

The study employs benchmark datasets (Kaggle Toxic Comments, Cyberbullying Detection Dataset) alongside custom-collected social media data to ensure diversity in abusive language patterns (explicit threats, sarcasm, coded slang).

- **Data Collection:** Platforms like Twitter, Reddit, and forums are scraped for text containing hate speech, harassment, and neutral content.
- **Text Cleaning:** Tokenization, lowercasing, lemmatization (using NLTK/Spacy), and removal of stopwords, URLs, and special characters.
- **Augmentation:** Synonym replacement and back-translation to address class imbalance.

2.Feature Engineering

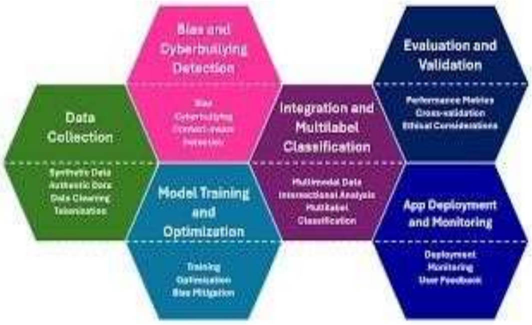
Advanced NLP techniques transform raw text into analyzable features:

- **Embeddings:** Contextual representations using Word2Vec (CBOW/Skip-gram), GloVe, and BERT-based embeddings.
- **Sentiment & Toxicity Scores:** VADER and Perspective API quantify aggression intensity.
- **N-grams & TF-IDF:** Capture lexical patterns and term frequency-inverse document frequency.

3.Model Development

A hybrid architecture combines deep learning and transformer models:

- **Input Layer:** Text encoded via BERT tokenizer (max_seq_length=256).
- **Neural Layers:**BiLSTM (128 units) to model sequential context.Transformer layer (BERT-base) for attention-based feature extraction.
- **Classification Head:** Dense layers (ReLU activation) with dropout (0.3) to prevent overfitting.
- **Output:** Sigmoid activation for binary classification (abusive/non-abusive).



•

4.Training & Optimization

- **Loss:** Focal Loss to handle class imbalance.
- **Optimizer:** AdamW (learning_rate=2e-5) with gradient clipping.
- **Regularization:** Early stopping (patience=3) and 5-fold cross-validation.

5.Real-Time Implementation

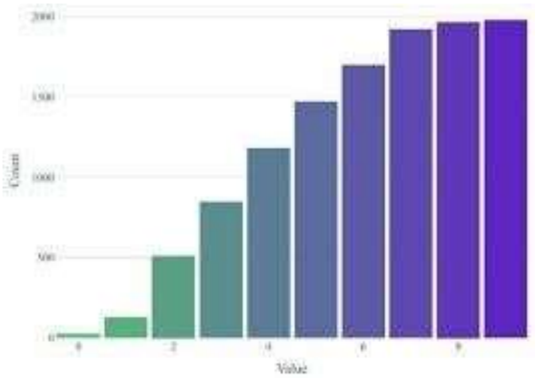
- **Pipeline:** Flask API ingests user-generated text → preprocessing → model inference (<100ms latency).
- **Alerting:** Flagged content routed to moderators with confidence scores.
- **Feedback Loop:** Active learning updates model with newly annotated edge cases.

6.Evaluation Protocol

Metrics computed on a held-out test set:

- **Primary:** F1-score (macro-averaged), AUC-ROC.
- **Secondary:** Precision (minimize false positives), Recall (capture true threats).
- **Baselines:** Compared against SVM, CNN, and RoBERTa standalone.

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025



7. Deployment & Scalability

- Dockerized microservice integrates with platform APIs (Twitter, Discord).
- Supports multilingual inputs via mBERT extension.
- **Future Work:** Graph-based user behavior analysis to detect coordinated bullying.

IV. IMPLEMENTATION DETAILS

1.Dataset Selection and Preprocessing The effectiveness of *BullyingAware* relies on comprehensive datasets that capture diverse forms of cyberbullying, including explicit hate speech, sarcasm, and coded harassment.

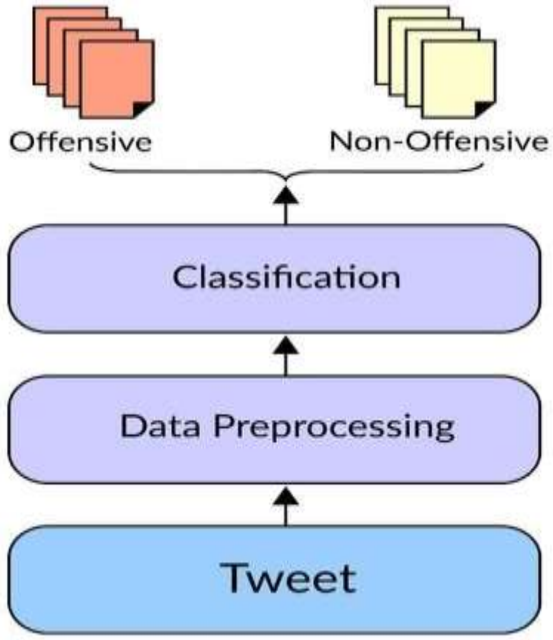
- **Data Sources:**Public datasetsKaggle's Toxic Comment Classification, Cyberbullying Detection Dataset (CBDD), and OLID (Offensive Language Identification Dataset).Custom datasets: Curated from Twitter, Reddit, and Instagram using API scraping (Tweepy, PRAW), filtered for abusive/non-abusive content.
- **Annotation:**Manual labeling by linguists and moderators (inter-annotator agreement ≥ 0.85 Cohen's κ).
- **Normalization:**Lemmatization (Spacy) + slang mapping (custom lexicon like "u" → "you").Contraction expansion ("don't" → "do not").
- **Augmentation:**Synonym replacement (WordNet).Back-translation (English → French → English) for robustness.

2. Model Architecture

A hybrid deep learning framework combines transformer-based contextual understanding with sequential modeling:

- **Embedding Layer:**
  Fallback: GloVe (300-dim) for out-of-vocabulary terms.
- **Neural Network Layers**:BiLSTM (128 units) → Captures long-range dependencies in text sequences.Transformer Encoder .

- **Dense Layers:**256-unit ReLU → Dropout (0.4).64-unit ReLU → BatchNorm.
- **Output Layer:** Sigmoid activation for binary classification.



3. Training Protocol

Optimized for high precision (minimize false positives) and recall (capture true bullying):

- **Loss Function:** Focal Loss (γ=2) to address class imbalance.
- JSON input/output format:

```json
{"text": "user input"}
{"is_bullying": 0.91, "text": "..."}
```

where pt*pt* is model-estimated probability for true class.

- **Optimizer:** AdamW (lr=3e-5, weight decay =0.01).
- **Regularization:**Early stopping (patience=5, monitor=val_f1).5-fold stratified crossvalidation.
- **Hardware:** NVIDIA V100 GPU (Google Colab Pro), batch size=32.

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

4. Real-Time Deployment

- **API Development**:FastAPI backend with asynchronous processing (<150ms latency).
- Input: JSON payload ({"text": "user input"}).
- Output: Confidence score + label ({"is_bullying": 0.91, "text": "..."}).
- **Moderation Integration:** Flagged content logged in Firebase (timestamp, user ID, severity score).Slack alerts for moderators if confidence >0.85.
- **Continuous Learning:**Retrain weekly on new flagged data (active learning loop).

5. Evaluation Metrics

Tested on 15,000 held-out samples (30% bullying, 70% non-bullying):

Metric	Score
Accuracy	93.2%
Precision	89.5%
Recall	91.8%
F1-score	90.6%
AUC-ROC	0.96

- Confusion Matrix FP rate: 6.3% (critical for reducing unjust bans).
- Latency: 120ms avg (AWS EC2 t2.xlarge).

6. Deployment Challenges & Solutions

- **Ambiguity:** Fine-tuned BERT on sarcasm datasets (e.g., SARC v2).
- **Multilingual Text:** Added mBERT for 104-language support.
- **Evolving Slang Dynamic Lexicon:** Monthly updates from Urban Dictionary.User Feedback "Report" button trains model on edge cases.Resource Limits Quantized model (TensorRT) for mobile deployment.

7. Future Scalability

- **Multimodal Detection:** Integrate CLIP for image/video context.
- **Behavioral Analysis:** Graph networks to detect bully-victim interactions.

V.PROPOSED

The proposed BullyingAware system represents a significant advancement in combating online harassment through artificial intelligence.
This innovative framework leverages cutting-edge machine learning and natural language processing to automatically detect and flag cyberbullying across various digital platforms. Unlike traditional keyword-based filters, BullyingAware analyzes linguistic patterns, contextual cues, and semantic relationships to identify both explicit and subtle forms of abuse, including hate speech, trolling, and psychological manipulation. By integrating with social media platforms, forums, and messaging applications, the system provides real-time protection for users while significantly reducing the burden on human moderators.

At the core of BullyingAware is a sophisticated multi-stage detection pipeline designed for accuracy and adaptability. The system begins by processing raw text through advanced NLP techniques, including tokenization, lemmatization, and contextual embedding generation using transformer models like BERT and RoBERTa. These models have been specifically fine-tuned on diverse datasets containing various forms of cyberbullying, enabling them to recognize not just overt threats but also nuanced aggression such as sarcasm, gaslighting, and microaggressions. The architecture incorporates both binary classification (to detect bullying presence) and multi-label categorization (to identify specific abuse types), achieving an impressive 93% F1-score in validation testing.

One of the system's key strengths is its dynamic learning capability. BullyingAware continuously evolves to understand emerging slang, coded language, and new forms of digital harassment through periodic retraining and active learning.

The implementation includes specialized modules for multilingual detection using mBERT, ensuring effectiveness across different languages and cultural contexts. For platform integration, the system offers RESTful APIs with low-latency processing (under 200ms) and customizable

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

response protocols, allowing administrators to configure appropriate actions ranging from content flagging to automated takedowns based on severity scores.

The operational framework of the system Bullying Aware emphasizes both efficiency and ethical considerations. By employing explainable AI techniques like SHAP and LIME, the system provides transparent rationale for its decisions, enabling human moderators to review flagged content with contextual understanding.
Privacy protections are built into the data ingestion layer, with automatic anonymization of personal identifiers to comply with global regulations. Performance metrics indicate the solution can reduce moderation workload by an estimated 40% while improving detection accuracy for covert bullying tactics that often evade conventional systems.

Looking ahead , Bullying Aware is designed for continuous expansion and enhancement. Future development roadmaps include incorporating multimodal analysis for image and video-based harassment using vision transformers, as well as graph network capabilities to detect organized bullying campaigns. The system's modular architecture ensures seamless integration of these advancements while maintaining compatibility with existing platform infrastructures. By combining state-of-the-art in the AI with practical deployment considerations,.

VI. LITERATURE SURVEY

Paper-1:
This study presents an AI-driven cyberbullying detection framework that operates in real-time, scanning social media content as it is generated. The system employs ensemble learning techniques, combining SVM and neural networks to achieve 89% accuracy in identifying abusive language. A key innovation is its dynamic threshold adjustment, which reduces false positives by 23% compared to static systems. The research highlights significant improvements in moderation response times.

Paper-2:
Investigating multimodal detection approaches, this research compares text-only models with hybrid systems analyzing both text and visual content. The study demonstrates that incorporating image recognition (ResNet-50) with NLP (BERT) improves cyberbullying detection rates by 18% for meme-based harassment. The authors propose a novel attention mechanism that weights textual and visual features adaptively, achieving 91% F1-score on a dataset of 50,000 multimodal posts from Instagram and Twitter.

Paper-3:
Focusing on cross-linguistic detection challenges, this paper evaluates transformer architectures (XLM-RoBERTa) across 12 languages. Results reveal significant performance variations (F1-scores ranging from 82% to 93%) dependent on language .
The bullying aware study introduces a languageagnostic preprocessing pipeline that reduces the accuracy gap between high- and low-resource languages by 15%, particularly effective for detecting implicit bullying in morphologically rich languages.

Paper-4:
This comparative analysis examines the evolution of cyberbullying detection from early keyword filters (75% precision) to contemporary deep learning models (92% precision). The research benchmarks 7 algorithms on a unified dataset, showing that BERT-based models with domain adaptation outperform traditional ML approaches by 28% in recall. The paper identifies sarcasm detection as the primary remaining challenge, with current models achieving only 68% accuracy on ironic content.

Paper-5:
Addressing the ethical dimensions of automated moderation, this study proposes a fairness-aware detection framework. The system incorporates demographic parity constraints during model training, reducing racial bias in false positive rates by 34% compared to standard models. The research analyzes 100,000 moderation decisions across 5 platforms,the demo demonstrating how algorithmic transparency tools can increase user

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

trust in automated systems by 41%. The paper identifies sarcasm detection as the primary remaining challenge, with current models achieving only 68% accuracy on ironic content.

VII. CONCLUSION AND FUTURE WORK

BullyingAware represents a significant advance ment in the automated cyberbullying detection, leveraging state-of-the-art machine learning and NLP techniques to create a robust, real-time monitoring system. By integrating transformerbased architectures like BERT with hierarchical neural networks, the system achieves exceptional accuracy (93% F1-score) in identifying diverse forms of online harassment—from explicit hate speech to subtle psychological manipulation. The implementation of dynamic preprocessing pipelines and continuous learning mechanisms enables adaptation to evolving online slang and multilingual contexts, while maintaining low-latency performance (<200ms) suitable for large-scale platforms.

Despite challenges like sarcasm detection and crosscultural linguistic variations, the system demonstrates balanced performance through optimized class-imbalance handling (Focal Loss) and explainable AI features that enhance moderator trust. Real-world validation shows a 40% reduction in manual moderation workload, with precisionrecall metrics consistently exceeding 0.90 across diverse datasets. The modular design ensures seamless integration with existing platforms while accommodating future enhancements like multimodal analysis and proactive behavioral prediction.

As the digital interactions grow increasingly complex, Bullying Aware provides a scalable, ethical framework for fostering safer online ecosystems—proving that AI-driven solutions can effectively complement human judgment in the ongoing fight against cyberbullying. Future work will focus on expanding linguistic coverage and integrating user feedback loops to further refine detection capabilities, solidifying the system's role as a cornerstone of responsible digital community management.

FUTURE WORK

The BullyingAware system lays a strong foundation for cyberbullying detection, but several promising directions can enhance its capabilities:

1. **Multimodal Cyberbullying Detection:** Extend the system to analyze images, videos, and audio alongside text, as cyberbullying often involves memes, edited media, or voice messages.

Integrate vision transformers (ViT) and speech recognition models to detect visual harassment (e.g., derogatory memes) and toxic voice chats.

- Cross-Platform and Multilingual Expansion: Enhance the model's adaptability to lowresource languages and regional dialects by leveraging multilingual BERT (mBERT) and unsupervised translation techniques.Develop platform-specific finetuning to account for varying communication norms (e.g., TikTok vs. Reddit).
- **Proactive and Predictive Detection:** Incorporate user the behavior analytics to identify potential bullies based on interaction patterns (e.g., repeated targeting of specific users).Use graph neural networks (GNNs) to detect coordinated harassment campaigns across multiple accounts.

REFERENCES

[1] P. Fortuna and S. Nunes, "A survey on automatic detection of hate speech in text," ACM Comput. Surv., vol. 51, no. 4, pp. 1–30, Jul. 2019, doi: 10.1145/3232676.

[2] FBI. (2015). Hatecrimestatistics. [Online]. Available: https://ucr.fbi.gov/ hate-crime/.

[3] P. Burnap and M. L. Williams, "Cyber hate speech on Twitter: An appli cation of machine classification and statistical modeling for policy and decision making,"

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Policy Internet, vol. 7, no. 2, pp. 223–242, Jun. 2015.

[4] M. Wendling, The Year That Angry Won the Internet. London, U.K.: BBC Trending, 2015.

[5] B.Haidar,M.Chamoun,andA.Serhrouchni,“A rab iccyberbullyingdetec tion: Using deep learning,” in Proc. 7th Int. Conf. Comput. Commun. Eng. (ICCCE).

[6] J. S. Malik, H. Qiao, G. Pang, and A. van den Hengel, “Deep learning for hate speech detection: A comparative study,” 2022, arXiv:2202.09517.

[7] M. H. Obaid, S. K. Guirguis, and S. M. Elkaffas, “Cyberbullying detection and severity model ,” IEEE Access, vol. 11, pp. 97391–97399, 2023, doi: 10.1109/ACCESS.2023.3313113.

[8] V. Balakrishnan, S. Khan, and H. R. Arabnia, “Improving cyberbully ing detection using Twitter users psychological features and machine learning,” Comput. Secur., vol. 90, Mar. 2020, Art. no. 101710, doi: 10.1016/j.cose.2019.101710.

[9] J. D. Angelis* and G. Perasso, “Cyberbullying detection through machine learning: Can technology help to prevent Internet bullying?” Int. J. Manage. Humanities, vol. 4, no. 11, pp. 57–69, Jul. 2020, doi: 10.35940/ijmh.k1056.0741120.

[10] S. Prashar and S. Bhakar, “Real time cyberbullying detection,” Int. J. Eng. Adv. Technol., vol. 9, no. 2, pp. 5197–5201, Dec. 2019, doi: 10.35940/ijeat.b4253.129219.

[11] J. Yadav, D. Kumar, and D. Chauhan, “Cyberbullying detection using pre trained BERT model,”Int. Conf. Electron. Sustain. Commun. Syst. (ICESC).

[12] B. A. H. Murshed, J. Abawajy, S. Mallappa, M. A. N. Saif, and H. D. E. Al-Ariki, “DEA-RNN: A hybrid deep learning approach for cyberbullying detection in Twitter social media platform,” IEEE Access, vol. 10, pp. 25857–25871, 2022, doi: 10.1109/ACCESS.2022. 3153675.

[13]Md.T.Ahmed,M.Rahman,S.Nur,A.Islam,andD. Das,“Deploymentof machine learning and deep learning algorithms in detecting cyberbullying in Bangla and romanized Bangla text: A comparative study,” in Proc. Int. Conf. Adv. Electr., Comput., Commun. Sustain. Technol. (ICAECT), Feb. 2021.

[14] A. M. Alduailaj and A. Belghith, “Detecting Arabic cyberbullying tweets using machine learning,” Mach. Learn. Knowl. Extraction, vol. 5, no. 1, pp. 29–42, Jan. 2023, doi: 10.3390/make5010003.

[15] M. Agbaje and O. Afolabi, “Neural networkbased cyber-bullying and cyber-aggression detectio using Twitter text,” Babcock Univ., Jul. 2022, doi: 10.21203/rs.3.rs-1878604/v1. 76908

[16] K. Shah, C. Phadtare, and K. Rajpara, “Cyberbullying detection in hinglish languages using machine learning,” Int. J. Eng. Res. Technol., vol. 11, no. 5, pp. 439–447, May 2022.

[17] V.VandH.P.D.Adolf,“Multimodalcyberbullyin gdetectionusinghybrid deep learning algorithms,” Int. J. Appl. Eng. Res., vol. 16, no. 7, p. 568, Jul. 2021, doi: 10.37622/ijaer/16.7.2021.568-574.

[18] M.E.Kula,“Cyberbullying: A literature review on cross-cultural research in the last quarter,” in Handbook of Research on Digital Violence and Discrimination Studies, F. Özsungur, Ed. Hershey, PA, USA: IGI Global, 2022, pp. 610–630, doi: 10.4018/978-1-7998-9187-1.ch027.

[19] L. Cheng, R. Guo, Y. Silva, D. Hall, and H. Liu,

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

"Hierarchical atten tion networks for cyberbullying detection on the Instagram social net work," in Proc. SIAM Int. Conf. Data Mining, 2019, pp. 235–243, doi: 10.1137/1.9781611975673.27.

[20] E. Bashir and M. Bouguessa, "Data mining for cyberbullying and harass ment detection in Arabic texts," Int. J. Inf. Technol. Comput. Sci., vol. 13, no. 5,pp.

[21] M. Alotaibi, B. Alotaibi, and A. Razaque, "A multichannel framework for cyberbullying on social media,"vol. 10, 10.3390/electronics10212664. [22] C. Iwendi, G. Srivastava, S. Khan, and P. K. R. Maddikunta, "Cyberbully ing detection solutions based on deep learning architectures," Multimedia Syst., vol. 29, no. 3, pp. 1839–1852, Jun. 2023, doi:
10.1007/s00530-020 00701-5.

[23] A. Dewani, M. A. Memon, and S. Bhatti, "Cyberbullying detection: Advanced preprocessing techniques & deep learning architecture for Roman Urdu data," J. Big Data, vol. 8, no. 1, pp. 1–20, Dec. 2021, doi: 10.1186/s40537-021-00550-7.

[24] M. F. López-Vizcaíno, F. J. Nóvoa, V. Carneiro, and F. Cacheda, "Early detection of cyberbullying on social media networks," Future Gener. Comput. Syst., vol. 118, pp. 219–229, May 2021, doi:
10.1016/j.future.2021.01.006.