# CIC-NLP@DravidianLangTech 2025: Fake News Detection in Dravidian Languages

**Tewodros Achamaleh[1], Nida Hafeez[1], Mikiyas Mebiratu[2]**
**Fatima Uroosa[1], Grigori Sidorov[1]**
[1]Instituto Politécnico Nacional, Centro de Investigación en Computación, Mexico City, Mexico
[2]Wolkite University, Department of Information Technology, Wolkite, Ethiopia

## Abstract

Misinformation is a growing problem for technology companies and society. Although there is a large body of related work on identifying fake news in predominantly resource languages, there is a lack of such studies in low-resource languages (LRLs). Because corpora and annotated data are scarce in LRLs, the identification of false information remains in the exploratory stage. Fake news detection is critical in this digital era to avoid the spread of misleading information. This study presents an approach to Detect Fake News in Dravidian Languages. Our team CIC-NLP work primarily targets Task 1, which involves identifying whether a given social platform news is original or fake. For the fake news detection problem, we used the mBERT model and utilized the dataset provided by the organizers of the workshop. In this section, we describe our findings and the results of the proposed method. The mBERT model achieved an F1 score of 0.853. The source code is available on GitHub.[1]

## 1 Introduction

Social media plays an important role in how people convey and access information in the modern world (Omar and Ondimu, 2024). While instant communication and global connectivity offer numerous benefits, they also bring a significant downside: the problem of disseminating fake news at a very fast rate (Schmidt and Wiegand, 2017). Fake news, described as actual falsehoods disinformation that is passed as factual news, is normally spread to influence people, create controversy, or achieve certain objectives (Bharathi et al., 2021). Fake news is the most rampant problem on social media, significantly transforming or even eroding society in terms of truth and democratic values. Because conspiracy theories can be spread on social media platforms within minutes, millions of people constantly receive distorted information telling the truth from the lie. This process destroys the credibility of institutions and the media, which poses a significant problem for the stable development of society and making conscious decisions (Balaji et al., 2023).

The fake news detection problem has become a critical issue, and it is difficult to ensure the integrity of news content (Aïmeur et al., 2023; Subramanian et al., 2023). Although researchers have conducted a lot of work in numerous domains and languages (Abiola et al., 2025b,a; Mehak et al., 2025), FND in Dravidian languages still faces challenges due to its cultural and linguistic characteristics (Anbalagan et al., 2024). Fake news has caused many incidents, such as the Brexit vote in Britain, the 2017 US elections, and the case in South Africa targeting Finance Minister Pravin Gordhan and media professionals. Such updates on social networks can be attributed to rumors or fiction due to existing fake news or hearsay accounts (Shu et al., 2017). Therefore, it is imperative to develop proper methods for detecting fake news (Oshikawa et al., 2018). The challenge of identifying fake news in Dravidian languages demands specific approaches because these languages possess distinct linguistic and cultural traits (Chakravarthi et al., 2021; Arif et al., 2022; Malliga et al., 2023).

In response to these challenges, workshop organizers launched fake news detection task (Subramanian et al., 2025). The shared task presents developers with a new opportunity to advance research and build better systems for detecting fake news across Dravidian languages. By working together researchers and practitioners build new methods to identify fake news in different languages as shared tasks reveal the motivation behind teamwork and innovation. Our team CIC-NLP used mBERT to classify news into real and fake categories. To identify fake news in Dravidian languages, this work investigates the methods, datasets used and demon-

[1]https://github.com/teddymas95/Fake-News-Detection



*Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 699–706
May 3, 2025 ©2025 Association for Computational Linguistics

strate results. Extending from the extant literature, it is our intention to contribute by culturally informed approaches to improve the efficiency of this task.

## 2 Literature Review

During past years, by exploring and investigating social media platforms, researchers achieved major progress in fake news detection (Bala and Krishnamurthy, 2023; Chen et al., 2023; Subramanian et al., 2025). Normal method to evaluate a post authenticity is to examine different key indicators such as shares, likes and followers etc. Researchers (Rodríguez and Iglesias, 2019) utilized traditional ML approaches like support vector machines and classification trees to identify or examine key indicators. In the same way, other researchers (Shu et al., 2017; Raja et al., 2023) examined user data including social networks and text information through ML tools to extract network and profile attributes.

Researchers (Singh et al., 2018; Sharma et al., 2019, 2020) examined and reviewed how different Deep Learning (DL), Artificial Intelligence (AI) and Machine Learning (ML) methods help to identify fake news. In another work researchers (Granskogen, 2018; Chauhan and Palivela, 2021; Tonja et al., 2022; Tash et al., 2022; Yigezu et al., 2023a,b; Bade et al., 2024; Mersha et al., 2024) examined how NLP and ML approaches help to detect fake news by studying user sentiments and how content is written. Transfer learning and DL model applications have also been considered (Raja et al., 2023). CNNs and RNNs approaches for FND were used in the study by (Goldani et al., 2021). Static and dynamic searches have been analyzed by (Ahmad et al., 2020) where a two-tier ensemble approach to determine the authenticity of websites was conducted using Naive Bayes (Granik and Mesyura, 2017), Random Forest, and Logistic Regression algorithms. Reacher's used the Naïve Bayes classifier which is a Bayesian solution for detecting fake news (Adiba et al., 2020; Subramanian et al., 2024; Devika et al., 2024).

As the attention has shifted towards the use of graph networks, few works have aimed toward developing graph-based approaches. Researchers (Nguyen et al., 2020) presented FANG for novel graphical social context representation and learning. For realistic scenarios, authors (Lu and Li, 2020) designed Graph aware Co-attention Network (GCAN) to identify the authenticity of the source tweet and provide justification through identification of suspicious re-tweeters and important textual residues. (Mahabub, 2020) This work has also been done towards fake news detection in multilingual and low-resource settings using enable classifiers.

Researchers (Lucas et al., 2022) concentrated on COVID-19 misinformation in the Caribbean regions and trained the models in high-resource languages, then transferred the knowledge to low-resource datasets in English, Spanish, and Haitian French. Researchers (Sivanaiah et al., 2022) also developed fake news datasets for low-resource languages such as Malayalam, Kannada, Tamil and Gujarati. For the first time, researchers presented a multilingual and multi-domain fake news detection dataset that spanned over 5 languages and 7 domains and developed a novel BERT-based multi-language and multi-domain fake news detection framework (De et al., 2021). Some previous work has focused on the application of transfer learning and contextual word embedding's in FND (Akram and Shahzad, 2021; Yigezu et al., 2024). (Kalraa et al., 2021) used the transformer-based models for FND in Urdu, and (Ameer et al., 2021) used TL with the BERT model. Authors (Lina et al., 2020) further used CharCNN and RoBERTa to gain word and character-level sentence embedding's where label smoothing was incorporated to enhance generalization capability. In this context, researchers (Palani and Elango, 2023) employed contextual word embedding's with BERT and Roberta language corpora for detecting the fake news in Dravidian languages. In the same way, (Chakravarthi et al., 2022) used feed-forward networks (FFN) with RoBERTa to extract contextually dependent features for fake news detection.

Recent research in multilingual NLP Muhammad et al. (2023) has explored various techniques for text classification in low-resource languages. Muhammad et al. (2025) introduced a transformer-based approach that improved sentiment analysis, which we extend to enhance fake news detection in Dravidian languages by leveraging advanced modeling techniques. To sum up, the existing methods give useful information about how effectively different strategies function in identifying fake news. Still, there is a lack of further research to address new challenges that appear from time to time, multilingual recognition, and other difficulties in effectively detecting fake news in a varied and resource-constrained environment.

| Model | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| LSTM | 0.2509 | 0.5000 | 0.3342 | 0.5018 |
| DistilBERT (base-uncased) | 0.7420 | 0.7359 | 0.7345 | 0.7362 |
| XLM-RoBERTa (base) | 0.8350 | 0.8348 | 0.8348 | 0.8348 |
| BERT (base-multilingual-cased) | 0.8622 | 0.8613 | **0.8612** | 0.8614 |

Table 1: Model Comparison on the Development Dataset

## 3 Fake News Detection

### 3.1 Dataset Analysis

On the shared task fake news detection organized by DravidianLangTech@NAACL 2025, there are datasets for identifying fake vs original social media posts in Malayalam. The datasets we worked with consisted of "Label" and "Text". Each column was divided into development, training, and test subsets, and analysis was conducted as part of our team "CIC-NLP". We have 3,257 samples in our training dataset, 1,599 of which are tagged as 'fake' and 1,658 tagged as 'original,' forming a near balance. Also, the development dataset has 815 samples, including 406 'fake' and 408 'original,' where the evaluation is balanced. The testing dataset is provided with no labels for blind evaluation of model performance. Both datasets have a well-balanced structure so that training and evaluation will be unbiased. The training set contains lots of data for robust model training, while the development set can be used for tuning and validation. The "Text" column, though, may contain noisy or ambiguous content from social media, which introduces preprocessing challenges. The given data provides a solid foundation for building an effective fake news classifier, and we aim to build a robust classifier using this data.

### 3.2 BERT-base Multilingual Cased Model

The shared task has selected the Bert-base multilingual cased model, a powerful transformer-based language model that supports more than 100 languages, including Malayalam. This model effectively handles linguistic diversity, including transliteration and code-mixing, making it well-suited for multilingual-based applications. It can handle various languages used in the posts. By preventing text casing, the model's tokenizer helps us distinguish proper nouns and context-dependent entities required for identifying fake news. The model is fine-tuned on a task-specific dataset and classifies posts as either "fake" or "original" by using a contextual understanding of how language patterns manifest. We show that it effectively handles noisy social media data like emojis, abbreviations, and grammatical errors and extracts meaningful insights to guarantee accurate classification. By leveraging the BERT-base multilingual model, the CIC-NLP team aims to attain robust performance in fake news detection in Malayalam posts.

## 4 System Setup and Experiments

### 4.1 System Setup

For the Malayalam FND Task 1, we use the BERT-base multilingual cased model, fine-tuned on the provided datasets. The setup requires a GPU-enabled environment with libraries such as PyTorch, Hugging Face Transformers, and Scikit Learn for data processing, training, and evaluation. Preprocessing datasets involved mapping labels (fake and original) and tokenizing with the model's tokenizer with padding and truncation. We fine-tune the model with a learning rate of 3e-5, batch size of 16, and five iterations. Evaluation is done on metrics such as accuracy and F1 scores. We evaluate the performance of the development dataset and save predictions for the test dataset. Classification performance is explored using visualization tools such as confusion matrices and ROC curves. It can be fine-tuned efficiently and is good at detecting fake news in Malayalam posts.

### 4.2 Experiments

We studied the system's performance through experiments with a BERT-base multilingual cased model for FND in Malayalam. The training dataset

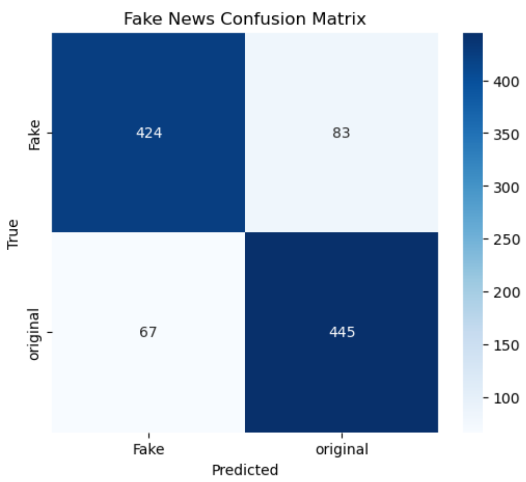


Figure 1: Confusion matrix

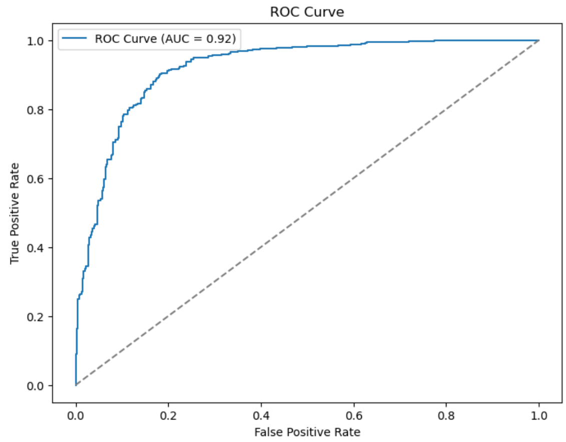


Figure 2: ROC curve

is 3,257 elements, split into 1,599 "fake" and 1,658 "original" labels, and the development dataset contains 815 elements (406 "fake" and 408 "original" labels). For this, we fine-tuned the model using the Hugging Face API Trainer with a learning rate of 3e-5, batch size of 16 and 5 epochs, and weight decay of 0.01. The best checkpoint was saved based on the macro F1 score and evaluations done at the end of each epoch. Accuracy, macro F1 score, and weighted F1 score were used to evaluate the model performance comprehensively. On the development dataset, the fine-tuned model achieved high accuracy and F1 scores for classification between the two classes, "fake" and "original" news posts. A confusion matrix was plotted to visualize the true versus predicted labels, and a ROC curve was plotted to see the contrast and the AUC score from the model to the separation between its classes. In these experiments, we demonstrate that the BERT-base multilingual cased model performs well on fake news detection in Malayalam when using multilingual and noisy data. The results are presented in Table 1 along with Figures 1 and 2, which illustrate the confusion matrix and ROC curve plots.

## 5 Results

We evaluate the fine-tuned BERT-base multilingual case model on a Malayalam dataset for binary classification. The model showed strong predictive ability on the development set, achieving a macro average F1 score of 0.85, which suggests balanced performance over the two classes, "fake" and "original." The results also demonstrate that the model can generalize well with noisy and relatively imbalanced data. We found that the multilingual BERT architecture is robust to linearity, allowing us to achieve strong performance on Malayalam social media text with negligible performance degradation. The findings highlight the model's effectiveness in detecting fake news in a multilingual environment. The result obtained from the test set is 0.85.

## 6 Discussion

BERT-multilingual performed best in FND, achieving the highest F1-score, while XLM-RoBERTa also showed strong results. LSTM struggled due to its limited contextual understanding. Class imbalance in Malayalam affected performance, highlighting the need for balanced datasets. The results reinforce the effectiveness of transformer models for low-resource languages. BERT-multilingual outperformed LSTM and DistilBERT, achieving the highest precision and recall. DistilBERT misclassified many fake news samples, and LSTM had difficulty capturing complex linguistic patterns. Future improvements could include domain-specific fine-tuning and metadata integration for better accuracy. Table 1 compares the models.

### 6.1 Error Analysis

Most misclassifications occurred in the imbalanced Malayalam dataset. False positives were common when fake news closely resembled original content, while false negatives resulted from ambiguous writing styles. Addressing these issues requires better feature representation and fine-tuning. This analysis compares model predictions with the ground truth to identify misclassification patterns and assess overall model performance. Figure 1 shows the confusion matrix.

## 7 Conclusion

Our team assessed the ability of transformer-based models to identify fake news information in the Malayalam language. Despite the unequal distribution of classes within the dataset, the mBERT-multilingual model achieved the highest accuracy along with the F1-score among other models. The model showed resilient behavior when detecting fake news despite its reduced performance because of the unbalanced dataset. Classification errors stemmed from counterfeit and genuine news pairs that appeared similarly or used ambiguous writing styles. All models experienced difficulties processing uncertain cases because their linguistic feature integration required further enhancement. Future development of fake news detection systems should address enhancements in feature representation combined with domain-based tweaking of models alongside metadata advancement for better identification in low-resource languages.

## 8 Limitations

This study faced several challenges, primarily due to class imbalance affecting the performance of the Malayalam model. The model struggled with generalization as the dataset lacked sufficient diversity in writing styles. Misclassification issues arose when fake news closely resembled original content, making detection difficult. Additionally, the models had trouble processing ambiguous cases with misleading linguistic patterns. Future improvements should focus on expanding the dataset with balanced class distribution and integrating advanced linguistic features to enhance detection accuracy.

## Acknowledgments

The work was done with partial support from the Mexican Government through the grant A1-S-47854 of CONACYT, Mexico, grants 20241816, 20241819, and 20240951 of the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional, Mexico. The authors thank the CONACYT for the computing resources brought to them through the Plataforma de Aprendizaje Profundo para Tecnologías del Lenguaje of the Laboratorio de Supercómputo of the INAOE, Mexico and acknowledge the support of Microsoft through the Microsoft Latin America PhD Award.

## References

Tolulope Olalekan Abiola, Tewodros Achamaleh Bizuneh, Oluwatobi Joseph Abiola, Temitope Olasunkanmi Oladepo, Olumide Ebenezer Ojo, Grigori Sidorov, and Olga Kolesnikova. 2025a. CIC-NLP at GenAI detection task 1: Leveraging DistilBERT for detecting machine-generated text in English. In *Proceedings of the 1stWorkshop on GenAI Content Detection (GenAIDetect)*, pages 271–277, Abu Dhabi, UAE. International Conference on Computational Linguistics.

Tolulope Olalekan Abiola, Tewodros Achamaleh Bizuneh, Fatima Uroosa, Nida Hafeez, Grigori Sidorov, Olga Kolesnikova, and Olumide Ebenezer Ojo. 2025b. CIC-NLP at GenAI detection task 1: Advancing multilingual machine-generated text detection. In *Proceedings of the 1stWorkshop on GenAI Content Detection (GenAIDetect)*, pages 262–270, Abu Dhabi, UAE. International Conference on Computational Linguistics.

Farzana Islam Adiba, Tahmina Islam, M Shamim Kaiser, Mufti Mahmud, and Muhammad Arifur Rahman. 2020. Effect of corpora on classification of fake news using naive bayes classifier. *International Journal of Automation, Artificial Intelligence and Machine Learning*, 1(1):80–92.

Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf, and Muhammad Ovais Ahmad. 2020. Fake news detection using machine learning ensemble methods. *Complexity*, 2020(1):8885861.

Esma Aïmeur, Sabrine Amri, and Gilles Brassard. 2023. Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*, 13(1):30.

Hammad Akram and Khurram Shahzad. 2021. Ensembling machine learning models for urdu fake news detection. In *FIRE (Working Notes)*, pages 1142–1149.

Iqra Ameer, Claudia Porto Capetillo, Helena Gómez-Adorno, and Grigori Sidorov. 2021. Automatic fake news detection in urdu language using transformers. In *FIRE (Working Notes)*, pages 1127–1134.

Akshatha Anbalagan, T Priyadharshini, A Niranjana, Shreedevi Balaji, and Durairaj Thenmozhi. 2024. Wordwizards@ dravidianlangtech 2024: Fake news detection in dravidian languages using cross-lingual sentence embeddings. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 162–166.

Muhammad Arif, Atnafu Lambebo Tonja, Iqra Ameer, Olga Kolesnikova, Alexander F Gelbukh, Grigori Sidorov, and Abdul Gafar Manuel Meque. 2022. Cic at checkthat!-2022: Multi-class and cross-lingual fake news detection. In *CLEF (Working Notes)*, pages 434–443.

Girma Bade, Olga Kolesnikova, Grigori Sidorov, and José Oropeza. 2024. Social media fake news classification using machine learning algorithm. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 24–29.

Abhinaba Bala and Parameswari Krishnamurthy. 2023. Abhipaw@ dravidianlangtech: Fake news detection in dravidian languages using multilingual bert. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 235–238.

Varsha Balaji, B Bharathi, et al. 2023. Nlp_ssn_cse@ dravidianlangtech: Fake news detection in dravidian languages using transformer models. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 133–139.

B Bharathi et al. 2021. Ssncse_nlp@ dravidianlangtech-eacl2021: Offensive language identification on multilingual code mixing text. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 313–318.

Bharathi Raja Chakravarthi, Mihaela Găman, Radu Tudor Ionescu, Heidi Jauhiainen, Tommi Jauhiainen, Krister Lindén, Nikola Ljubešić, Niko Partanen, Ruba Priyadharshini, Christoph Purschke, et al. 2021. Findings of the vardial evaluation campaign 2021. In *EACL| VarDial*. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, Subalalitha Cn, John Philip McCrae, Miguel Ángel García, Salud María Jiménez-Zafra, Rafael Valencia-García, Prasanna Kumaresan, Rahul Ponnusamy, et al. 2022. Overview of the shared task on hope speech detection for equality, diversity, and inclusion. In *Proceedings of the second workshop on language technology for equality, diversity and inclusion*, pages 378–388.

Tavishee Chauhan and Hemant Palivela. 2021. Optimization and improvement of fake news detection using deep learning approaches for societal benefit. *International Journal of Information Management Data Insights*, 1(2):100051.

Sijing Chen, Lu Xiao, and Akit Kumar. 2023. Spread of misinformation on social media: What contributes to it and how to combat it. *Computers in Human Behavior*, 141:107643.

Arkadipta De, Dibyanayan Bandyopadhyay, Baban Gain, and Asif Ekbal. 2021. A transformer-based approach to multilingual fake news detection in low-resource languages. *Transactions on Asian and Low-Resource Language Information Processing*, 21(1):1–20.

K Devika, B Haripriya, E Vigneshwar, B Premjith, Bharathi Raja Chakravarthi, et al. 2024. From dataset to detection: A comprehensive approach to combating malayalam fake news. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 16–23.

Mohammad Hadi Goldani, Reza Safabakhsh, and Saeedeh Momtazi. 2021. Convolutional neural network with margin loss for fake news detection. *Information Processing & Management*, 58(1):102418.

Mykhailo Granik and Volodymyr Mesyura. 2017. Fake news detection using naive bayes classifier. In *2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON)*, pages 900–903. IEEE.

Torstein Granskogen. 2018. Automatic detection of fake news in social media using contextual information. Master's thesis, NTNU.

Sakshi Kalraa, Preetika Vermaa, Yashvardhan Sharma, and Gajendra Singh Chauhan. 2021. Ensembling of various transformer based models for the fake news detection task in the urdu language. In *FIRE (Working Notes)*, pages 1175–1181.

Nankai Lina, Sihui Fua, and Shengyi Jianga. 2020. Fake news detection in the urdu language using charcnn-roberta. *Health*, 100:100.

Yi-Ju Lu and Cheng-Te Li. 2020. Gcan: Graph-aware co-attention networks for explainable fake news detection on social media. *arXiv preprint arXiv:2004.11648*.

Jason Lucas, Limeng Cui, Thai Le, and Dongwon Lee. 2022. Detecting false claims in low-resource regions: A case study of caribbean islands. In *Proceedings of the Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situations*, pages 95–102.

Atik Mahabub. 2020. A robust technique of fake news detection using ensemble voting classifier and comparison with other classifiers. *SN Applied Sciences*, 2(4):525.

S Malliga, Bharathi Raja Chakravarthi, SV Kogilavani, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani, and Muskaan Singh. 2023. Overview of the shared task on fake news detection from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 59–63.

Gull Mehak, Amna Qasim, Abdul Gafar Manuel Meque, Nisar Hussain, Grigori Sidorov, and Alexander Gelbukh. 2025. TechExperts(IPN) at GenAI detection task 1: Detecting AI-generated text in English and multilingual contexts. In *Proceedings of the 1stWorkshop on GenAI Content Detection (GenAIDetect)*, pages 161–165, Abu Dhabi, UAE. International Conference on Computational Linguistics.

Melkamu Abay Mersha, Girma Yohannis Bade, Jugal Kalita, Olga Kolesnikova, Alexander Gelbukh, et al. 2024. Ethio-fake: Cutting-edge approaches to combat fake news in under-resourced languages using explainable ai. *Procedia Computer Science*, 244:133–142.

Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Abinew Ali Ayele, David Ifeoluwa Adelani, Ibrahim Said Ahmad, Saminu Mohammad Aliyu, Nelson Odhiambo Onyango, Lilian DA Wanzare, Samuel Rutunda, Lukman Jibril Aliyu, et al. 2025. Afrihate: A multilingual collection of hate speech and abusive language datasets for african languages. *arXiv preprint arXiv:2501.08284*.

Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Abinew Ali Ayele, Nedjma Ousidhoum, David Ifeoluwa Adelani, Seid Muhie Yimam, Ibrahim Sa'id Ahmad, Meriem Beloucif, Saif M Mohammad, Sebastian Ruder, et al. 2023. Afrisenti: A twitter sentiment analysis benchmark for african languages. *arXiv preprint arXiv:2302.08956*.

Van-Hoang Nguyen, Kazunari Sugiyama, Preslav Nakov, and Min-Yen Kan. 2020. Fang: Leveraging social context for fake news detection using graph representation. In *Proceedings of the 29th ACM international conference on information & knowledge management*, pages 1165–1174.

AS Omar and KO Ondimu. 2024. The impact of social media on society: A systematic literature review. *The International Journal of Engineering and Science*, 13(6):96–106.

Ray Oshikawa, Jing Qian, and William Yang Wang. 2018. A survey on natural language processing for fake news detection. *arXiv preprint arXiv:1811.00770*.

Balasubramanian Palani and Sivasankar Elango. 2023. Bbc-fnd: An ensemble of deep learning framework for textual fake news detection. *Computers and Electrical Engineering*, 110:108866.

Eduri Raja, Badal Soni, and Samir Kumar Borgohain. 2023. Fake news detection in dravidian languages using transfer learning with adaptive finetuning. *Engineering Applications of Artificial Intelligence*, 126:106877.

Álvaro Ibrain Rodríguez and Lara Lloret Iglesias. 2019. Fake news detection using deep learning. *arXiv preprint arXiv:1910.03496*.

Anna Schmidt and Michael Wiegand. 2017. A survey on hate speech detection using natural language processing. In *Proceedings of the fifth international workshop on natural language processing for social media*, pages 1–10.

Karishma Sharma, Feng Qian, He Jiang, Natali Ruchansky, Ming Zhang, and Yan Liu. 2019. Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(3):1–42.

Uma Sharma, Sidarth Saran, and Shankar M Patil. 2020. Fake news detection using machine learning algorithms. *International Journal of creative research thoughts (IJCRT)*, 8(6):509–518.

Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1):22–36.

Naman Singh, Tushar Sharma, Abha Thakral, and Tanupriya Choudhury. 2018. Detection of fake profile in online social networks using machine learning. In *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pages 231–234. IEEE.

Rajalakshmi Sivanaiah, Nishaanth Ramanathan, Shajith Hameed, Rahul Rajagopalan, Angel Deborah Suseelan, and Mirnalinee Thanka Nadar Thanagathai. 2022. Fake news detection in low-resource languages. In *International Conference on Speech and Language Technologies for Low-resource Languages*, pages 324–331. Springer.

Malliga Subramanian, , B Premjith, Kogilavani Shanmugavadivel, Santhia Pandiyan, Balasubramanian Palani, and Bharathi Raja Chakravarthi. 2025. Overview of the Shared Task on Fake News Detection in Dravidian Languages: DravidianLangTech@NAACL 2025. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Malliga Subramanian, Bharathi Raja Chakravarthi, Kogilavani Shanmugavadivel, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani, B Premjith, K Vanaja, S Mithunja, K Devika, et al. 2024. Overview of the second shared task on fake news detection in dravidian languages: Dravidianlangtech@ eacl 2024. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 71–78.

Malliga Subramanian, Bharathi Raja Chakravarthi, Kogilavani Shanmugavadivel, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani, Muskaan Singh, Sandhiya Raja, Vanaja, and Mithunajha S. 2023. Overview of the shared task on fake news detection from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

M Shahiki Tash, Z Ahani, Al Tonja, M Gemeda, N Hussain, and O Kolesnikova. 2022. Word level language identification in code-mixed kannada-english texts using traditional machine learning algorithms. In *Proceedings of the 19th International Conference on Natural Language Processing (ICON): Shared Task on Word Level Language Identification in Code-mixed Kannada-English Texts*, pages 25–28.

Atnafu Lambebo Tonja, Mesay Gemeda Yigezu, Olga Kolesnikova, Moein Shahiki Tash, Grigori Sidorov, and Alexander Gelbuk. 2022. Transformer-based model for word level language identification in code-mixed kannada-english texts. *arXiv preprint arXiv:2211.14459*.

Mesay Yigezu, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2024. Habesha@DravidianLangTech 2024: Detecting fake news detection in Dravidian languages using deep learning. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 156–161, St. Julian's, Malta. Association for Computational Linguistics.

Mesay Gemeda Yigezu, Tadesse Kebede, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023a. Habesha@ dravidianlangtech: Utilizing deep and transfer learning approaches for sentiment analysis. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 239–243.

Mesay Gemeda Yigezu, Moges Ahmed Mehamed, Olga Kolesnikova, Tadesse Kebede Guge, Alexander Gelbukh, and Grigori Sidorov. 2023b. Evaluating the effectiveness of hybrid features in fake news detection on social media. In *2023 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, pages 171–175. IEEE.