Overview of the Shared Task on Fake News Detection in Dravidian Languages-DravidianLangTech@NAACL 2025

Malliga Subramanian[1], Premjith B[2], Kogilavani Shanmugavadivel[1], Santhiya Pandiyan[1], Balasubramanian Palani[3], Bharathi Raja Chakravarthi[4]
[1]Kongu Engineering College, Tamil Nadu, India,
[2]Amrita School of Artificial Intelligence, Coimbatore, Amrita Vishwa Vidyapeetham, India,
[3]Indian Institute of Information Technology Kottayam, Kerala, India,
[4]School of Computer Science, University of Galway, Ireland

Abstract

Detecting and mitigating fake news on social media is critical for preventing misinformation, protecting democratic processes, preventing public distress, mitigating hate speech, reducing financial fraud, maintaining information reliability, etc. This paper summarizes the findings of the shared task "Fake News Detection in Dravidian Languages—DravidianLangTech@NAACL 2025." The goal of this task is to detect fake content in social media posts in Malayalam. It consists of two subtasks: the first focuses on binary classification (Fake or Original), while the second categorizes the fake news into five types—False, Half True, Mostly False, Partly False, and Mostly True. In Task 1, 22 teams submitted machine learning techniques like SVM, Naïve Bayes, and SGD, as well as BERT-based architectures. Among these, XLM-RoBERTa had the highest macro F1 score of 89.8%. For Task 2, 11 teams submitted models using LSTM, GRU, XLM-RoBERTa, and SVM. XLM-RoBERTa once again outperformed other models, attaining the highest macro F1 score of 68.2%.

1 Introduction

In the modern age, information is spreading rapidly across the world, and the quality and truthfulness of the news affect society. This fast, rapid connectivity democratizes access to knowledge and information. However, it has also created a possibility for the proliferation of fake news and misinformation. Therefore, detecting and mitigating the fake and misinformation has become critical. Detecting fake news is a complex task due to the structure of the sentence. Unlike tasks like hate speech detection and sentiment analysis, where we have overt words/phrases that explain the meaning of the sentence, fake news doesn't contain such words, whereas it mimics the legitimate content.

Detecting fake news from low-resource languages like Malayalam is even more challenging due to linguistic diversity and resource limitations (Raja et al., 2024, 2023b). In addition, the presence of code-mixed text (Coelho et al., 2023) and the requirement of fine-tuning pre-trained models (Raja et al., 2023a) pose other challenges.

The goal of the shared task DravidianLangTech@NAACL 2025 is to address the difficulties in Malayalam fake news detection. This paper presents an overview of the submissions to this shared task. This task has two subtasks. The first task is to identify whether a given news item is fake or not, and the second task is about categorizing news into different fake categories (Subramanian et al., 2024). As a part of this task, we curated our own dataset. We gathered news from various fact-checking websites in Malayalam.

2 Related Works

Fake news detection and categorization are important tasks in languages like Malayalam due to the rapid spread of misinformation. Various approaches, including machine learning, deep learning, and transformer-based models for feature extraction as well as classification.

Machine learning algorithms such as random forest, support vector machine (SVM), logistic regression, and naive Bayes were widely used for detecting fake news and categories of fake news (Bade et al., 2024; Osama et al., 2024; Devika et al., 2024). Deep learning models also find success in this task. Long Short-Term Memory (LSTM) (Zamir et al., 2024) and Bidirectional LSTM (BiLSTM) models achieved a macro F1 score of 0.78. Convolutional Neural Networks (CNN) (Osama et al., 2024) were also employed for this task. Recently, researchers used XLM-RoBERTa (Malliga et al., 2023; Raja et al., 2023a), MuRIL (Farsi et al., 2024), m-BERT (Osama et al., 2024), and

Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 112–120
May 3, 2025

Malayalam BERT (Rahman et al., 2024; Tabassum et al., 2024). XLM-RoBERTa models achieved an F1 score of 0.87 and 0.90 in these tasks, whereas MuRIL-based models achieved an F1 score of 0.86 for the fake news detection task. Models built using m-BERT achieved similar performance with an F1 score of 0.85. Malayalam-BERT models achieved significant improvement in the categorization of fake news into different classes, with scores of 0.88 and 0.87.

3 Task description

3.1 Task 1

This task's objective is to determine if a particular social media text is original or fake; these data were sourced from numerous social media sites, including Facebook, Twitter, and others. The shared task's goal is to categorize a social media comment as either original or fake news. The classification of this task takes place at the comment/post level. The participant-submitted methods ought to classify a YouTube comment as either original or fake news.

3.2 Task 2

The primary objective of Task 2 is to classify fake news into different categories. In this task, we consider four classes of fake news, namely, false, mostly false, partly false, and half true. This classification helps people understand how much they have to rely on a specific news source to make their own decisions.

4 Dataset description

4.1 Task 1

The objective is to classify news items into 'Fake' and 'Original' categories. The dataset for this task comprises 1,599 training instances for 'Fake' and 1,658 for 'Original,' with respective testing sets of 507 and 512 instances nd development sets of 406 and 409 instances. Table 1 provides the number of data points in the training, development, and testing sets as well as the class-wise distribution

4.2 Task 2

In this task, the dataset was curated to contain different fake categories of Malayalam news, rather than classifying news into either fake or benign categories (Devika et al., 2024). We collected the news and their corresponding annotations from various fact-checking websites in Malayalam. We prepared a set of keywords to search for the news and identify their categories. To validate the authenticity of the annotations, we cross-checked them with multiple fact-checking tools. We provided train and test data for the participants of the shared task. Initially, we provided annotated training data for model building, and later, we provided test data without labels. Table 2 provides the number of data points in the training and testing sets as well as the class-wise distribution. The data is highly imbalanced, and the majority of the data in both the training and testing sets belong to the false category.

Data	Class	Count	Total
Train	Fake	1,599	3,257
	Original	1,658	
Development	Fake	406	815
	Original	409	
Test	Fake	507	1,019
	Original	512	

Table 1: Distribution of the data for Task 1

Data	Class	Count	Total
Train	False	1,386	1,900
	Mostly False	295	
	Partly False	57	
	Half True	162	
Test	False	100	200
	Mostly False	56	
	Partly False	7	
	Half True	37	

Table 2: A table explaining the distribution of the data in Train and Test datasets in Task 2

5 Methodology of participants

5.1 Task 1

Task 1 received 122 registrations. However, twenty-one teams actively participated and implemented their models. They tested the performance of the proposed models using the given fake news dataset, and the results are shown in Table 3.

5.1.1 Bytesizedllm

The team "bytesizedllm" (Manukonda and Kodali, 2025) developed an automatic fake news detection (FND) framework that uses a transformer-based fine-tuned XLM-RoBERTa model to leverage the strengths of both contextualized embeddings and sequential modeling. The transformer layer, integrated on top of the fine-tuned embeddings, further captures sequential dependencies, making the model highly effective for multilingual and transliteration-heavy tasks. The model has achieved the highest macro F1 score of 0.898 among all the other models proposed by other teams.

5.1.2 CUET_NLP_MP_MD

The team (Kabir et al., 2025) designed an FND system that combines multiple models, including Malayalam BERT, XLM-R, and Sarvamai/Sarvam-1 for contextual embedding and a majority voting classifier to detect fake news. This ensemble method leverages the strengths of each individual model to enhance performance and robustness. Hence, the model achieves a macro F1 score of 0.893 on the test data.

5.1.3 Awy

The team has employed a novel FND framework that consists of a mixture of multilingual models for contextual embedding and LLMs for emotion extraction to detect fake news effectively. The model has achieved a macro F1 score of 0.889.

5.1.4 Nayel

A machine learning-based system has been developed by the team 'Nayel' (Nayel et al., 2025), which integrates TF-IDF and n-grams as a feature extraction approach and sends the extracted features to ML-based classifiers such as SVM, SGD, Naive Bayes, and the Multi-Voting ensemble method to identify fake news. The highest macro F1 score of 0.875 is obtained by the ensemble model.

5.1.5 KCRL

The team (Haq et al., 2025) has implemented a text classification approach utilizing the XLM-RoBERTa base model augmented with a multi-pooling strategy. The methodology incorporates three distinct pooling mechanisms: CLS token extraction, mean pooling, and max pooling to capture comprehensive contextual representations from the input sequences. This unified pooling mechanism, enhanced by adaptive thresholding optimization, enables more robust classification by leveraging different semantic perspectives of the input text. Hence, their proposed model has achieved a macro F1 score of 0.874.

5.1.6 CUET-NLP_Big_O

The team (Sakib et al., 2025) has employed the XLM-RoBERTa (XLMR) large model, a multilingual transformer-based architecture, to classify social media text as either "Fake" or "Original". The model is tested on the dataset and achieves a macro F1 score of 0.874.

5.1.7 Celestia

The team (Noor et al., 2025) has designed an FND system that employs different embedding techniques and various ML and DL algorithms to detect fake news. The main advantage of this work includes the indic-transliteration library to create a consistent language format, English to Malayalam. The model achieves a macro F1 score of 0.859 on the test data.

5.1.8 MNLP

The team has developed an FND model that explores different deep learning-based models to identify fake news. The model has achieved a macro F1 score of 0.858.

5.1.9 CIC_NLP

The team (Achamaleh et al., 2025) has developed a novel FND framework that utilizes multilingual BERT (mBERT) for contextual word embedding for Tamil and Malayalam languages, and then the extracted features are sent to classifiers to detect fake news. The model achieves a macro F1 score of 0.853 on the unseen test dataset.

5.1.10 NLP_goats

The team (V K et al., 2025) implemented an automatic FND system that uses a multilingual BERT (mBERT) model for efficient fake news detection in Malayalam. These features make the model versatile and a very efficient solution for fake news detection in the Malayalam language. The model is tested on the dataset and achieves a macro F1 score of 0.839.

5.1.11 Necto

The team has utilized Sentence BERT, a fine-tuned model on the given data with a binary classification head for the classification downstream task. The sentence-level embeddings of the given text and the size of the model are small so that the training time of the model is faster in any system. Hence, the model achieves a macro F1 score of 0.832 on the test data.

5.1.12 Lowes

The team has developed an FND system that utilizes mBERT and various LLM-based approaches to detect fake news. The model has achieved a macro F1 score of 0.826.

5.1.13 Lemlem

The team 'Lemlem' has employed the pre-trained multilingual transformer model named mBERT for the word embedding that captures contextual relationships between words to detect fake news. This model is particularly suitable for multilingual text processing as it can handle diverse scripts and linguistic features effectively. A classification head was added to the BERT base model, which outputs probabilities for the predefined classes of fake news. Hence, the model achieves a macro F1 score of 0.823 on the test data.

5.1.14 Data_drifters

For this task, the team (Shanmugavadivel et al., 2025a) has employed a comprehensive methodology combining traditional machine learning models, embedding techniques, and advanced transfer learning models to achieve robust text classification. The team utilized four base models: Random Forest, Support Vector Machine (SVM), Logistic Regression, and Multinomial Naive Bayes, leveraging their diverse strengths in classification tasks. Two classical count-based techniques, such as TF-IDF and Count Vectorizer, are used to convert words into vectors. In addition, mBERT and XLNet are utilized for their exceptional ability to understand contextual semantics and multilingual text. The model is tested on the dataset and achieves a macro F1 score of 0.814.

5.1.15 ST_1 CIOL

The team (Anik et al., 2025) has designed a novel FND model that employs a multilingual encoder to effectively encode the text into meaningful embeddings. These embeddings were then fed into a Multi-Layer Perceptron (MLP) model for training, enabling the prediction of sentiment classes. The team has adopted a balance-aware modeling approach, actively tracking the best-performing model throughout the training process to ensure optimal performance. Finally, the best model is utilized for generating predictions on the test set and achieves a macro F1 score of 0.814.

5.1.16 Fact_fusion

The team has implemented a multilingual pipeline for detecting fake news for the Dravidian languages based on a systematic methodology. TF-IDF (Term Frequency-Inverse Document Frequency) was used for extracting features from the textual data, considering the term's significance but minimizing noise. The logistic Regression model has achieved a better macro F1 score of 0.803 when compared to other models.

5.1.17 YenCs

The team (Gowda and Hegde, 2025) has employed a novel automatic FND system that uses four different deep learning models (BiRNN, DNN, GRU, LSTM + RNN) with pre-trained word embeddings and combines their predictions using a weighted average based on validation accuracies. These models and the ensemble model were evaluated using metrics like accuracy and F1-score and achieved a macro F1-score of 0.792 on the unseen test data.

5.1.18 Blue_ray

The team (Shanmugavadivel et al., 2025b) has developed an FND system to classify Malayalam news into two categories: Original and Fake. After text pre-processing, features are extracted using TF-IDF to capture significant patterns in the text. Various machine learning models, such as Logistic Regression, Random Forest, and SVM, are trained on these features to predict the labels. The model achieves a macro F1-score of 0.790 on the unseen test dataset.

5.1.19 CIC

The team has developed a novel model to detect fake news. First, the proposed model performed tokenization and other pre-processing with indic_nlp method and then applied feature extraction using a fine-tuned mBERT model for training and prediction. The model is tested on the dataset and achieved the macro F1 score of 0.659.

5.1.20 DLRG

For the fake news classification in Malayalam, the team implemented an FND system that employs TF-IDF to convert text data into numerical features, highlighting important words. Then, Passive passive-aggressive classifier (PAC) is used to classify the TF-IDF transformed data into fake or real news. In addition, a Voting Classifier is utilized to combine predictions from multiple classifiers to enhance accuracy. Hence, the model achieves a macro F1-score of 0.473 on the test data.

5.1.21 CUET_ChiSquare

The team has designed a novel automatic FND system that utilizes a transformer-based approach leveraging XLM-RoBERTa, a multilingual pre-trained transformer model, fine-tuned for binary classification. The system's capability to generalize across varied linguistic structures and its efficient handling of imbalanced data make it particularly noteworthy for tasks involving low-resource and diverse language datasets. The model is achieved a macro F1-score of 0.334.

5.2 Task 2

Similar to Task 1, 122 teams registered for Task 2. However, only 11 teams submitted the predictions for the test data shared with the participants. The rank list for this task is shown in table 4. The following are the descriptions of the systems submitted by the participants.

5.2.1 byteSizedLLM

This team (Kodali and Manukonda, 2025) employed an advanced hybrid methodology, combining a customized BiLSTM network with a fine-tuned XLM-RoBERT base model to leverage the strengths of both contextualized embeddings and sequential modeling. The XLM-RoBERTa base model was fine-tuned using masked language modeling (MLM) on a carefully curated subset of the AI4Bharath dataset designed to enhance its multilingual contextual understanding. The dataset had original data, fully transliterated text, and partially transliterated data, with 20% to 70% of words randomly transliterated. This was done to add transliteration-based diversity. This method lets the model learn strong cross-lingual representations and adjust to different transliteration patterns that are common in collections of texts written in more than one language. The BiLSTM layer, which is added on top of the fine-tuned embeddings, captures even more sequential dependencies. This makes the model very good at tasks that require a lot of transliteration and more than one language.

5.2.2 YenCS

In this submission, the team (Gowda and Hegde, 2025) pre-processed the input text. The preprocessing step includes cleaning and tokenisation using Keras's Tokeniser. A pre-trained fastText model transformed the cleaned words into embeddings. The team trained three different models: a Convolutional Neural Network (CNN) with a GRU layer, an LSTM model, and an LSTM-GRU hybrid model. They subsequently use the predictions from these models as features in a stacking ensemble. A random forest classifier serves as the meta-learner, trained on the stacked predictions to produce the final classification. The team evaluated the effectiveness of both individual models and groups of models through accuracy and classification reports.

5.2.3 Fact Fusion

This team used a machine learning pipeline for the fake news classification task. This pipeline begins with text preprocessing, which removes noise like punctuation and extra white spaces. They used the term frequency-inverse document frequency (TF-IDF) for transforming the input text into embeddings. Unigrams and bigrams were considered for defining the features, and they restricted the vocabulary size to 5000 words. They performed the classification using a logistic regression classifier. They optimised the model training by fine-tuning the hyperparameter "max_iter."

5.2.4 Lowes

This team finetuned Malayalam BERT and also tried other LLM-based approaches. They also tried LLM-based synthetic data generation for this task.

5.2.5 KCRL

The team (Haq et al., 2025) developed a text classification system using the XLM-RoBERTa transformer model and improved it with a full pooling strategy and a data-balancing method. This method uses three different pooling methods—CLS token, mean pooling, and max pooling—to capture different parts of textual representations. To address class imbalance issues, they implemented an oversampling approach for minority classes, targeting a balanced distribution across all classes. Before the final classification, the concatenated pooled features go through a dense layer transformation. This lets the model use both global and local semantic features while keeping training levels even across all classes.

5.2.6 NLP_goats

The team (V K et al., 2025) developed a model that begins with preprocessing of the Malayalam text. They then encode the dataset labels and address imbalances through oversampling techniques.

The team trains a multilingual BERT model for multiclass classification.

5.2.7 MNLP

This team used deep learning-based models for classification.

5.2.8 Akatsuki-CIOL

This team (Anik et al., 2025) used a variety of encoders, such as Indic-specific and language-specific models, to get meaningful text embeddings that fit the multilingual nature of the data. They then processed these embeddings using the multilayer perceptron (MLP). It was the classification layer that predicted the corresponding classes. To ensure robust performance, the team adopted a balance-aware modeling approach, actively tracked and selected the best-performing model throughout the training process. Then, we used the chosen model to make predictions on the test set. This gave us a complete and flexible way to solve the multilingual sentiment classification problem.

5.2.9 Data_Drifters

For this task, the team (Shanmugavadivel et al., 2025a) employed a comprehensive methodology combining traditional machine learning models, embedding techniques, and advanced transfer learning models to achieve robust text classification. They used four base models: random forest, support vector machine (SVM), logistic regression, and multinomial logistic regression, leveraging their diverse strengths in classification tasks. To process text data, they implemented two embedding techniques: TF-IDF and Count Vectorizer, ensuring effective feature extraction and representation. Additionally, they incorporated two state-of-the-art transfer learning models, mBERT and XLNet.

5.2.10 Blue_Ray

The team implemented a multi-class classification system to categorize fake Malayalam news. The methodology involved preprocessing the text data by cleaning it and removing stopwords to ensure better feature representation. Feature extraction transformed the text data into the numerical format. They then split the processed data into training and testing subsets. Various machine learning and deep learning models were employed to train the data, and their performance was evaluated.

5.2.11 Cognitext

The team (Alladi and B, 2025) used a deep learning model to classify fake news articles. They first cleaned the text data by removing URLs, special characters, punctuation, and numbers, and then converted it to lowercase. They tokenized the cleaned text using Keras' Tokenizer and padded the sequences to ensure a uniform input length. The model architecture is made up of an embedding layer that stores word representations, an LSTM layer that tracks how events depend on each other, and a dense layer that uses softmax activation to sort words into multiple groups. They trained the model for five epochs using categorical cross-entropy loss and the Adam optimizer.

Team	Macro F1-score	Rank
bytesizedllm (Manukonda and Kodali, 2025)	0.898	1
CUET_NLP_MP_MD (Kabir et al., 2025)	0.893	2
One_by_zero (Chakraborty et al., 2025)	0.892	3
Awy	0.889	4
Nayel (Nayel et al., 2025)	0.875	5
KCRL (Haq et al., 2025)	0.874	6
CUET-NLP_Big_O (Sakib et al., 2025)	0.874	6
Celestia (Noor et al., 2025)	0.859	7
MNLP	0.858	8
CIC_NLP (Achamaleh et al., 2025)	0.853	9
NLP_goats (V K et al., 2025)	0.839	10
Necto	0.832	11
Lowes	0.826	12
Lemlem	0.823	13
Data_drifters (Shanmugavadivel et al., 2025a)	0.814	14
ST_1 CIOL	0.814	14
Fact_fusion	0.803	15
YenCs (Gowda and Hegde, 2025)	0.792	16
Blue_ray (Shanmugavadivel et al., 2025b)	0.790	17
CIC	0.659	18
DLRG	0.473	19
Technovators	0.387	20
CUET_ChiSquare	0.334	21

Table 3: Rank list of Task 1: Detecting fake news in Malayalam

Team	Macro F1-score	Rank
KCRL (Haq et al., 2025)	0.6283	1
byteSizedLLM (Kodali and Manukonda, 2025)	0.5775	2
NLP_goats (V K et al., 2025)	0.5417	4
Data_Drifters(Shanmugavadivel et al., 2025a)	0.5029	5
lowes	0.2902	6
YenCS (Gowda and Hegde, 2025)	0.2696	7
Blue_Ray (Shanmugavadivel et al., 2025b)	0.2631	8
Akatsuki-CIOL (Anik et al., 2025)	0.1978	11
Cognitext (Alladi and B, 2025)	0.1667	14
Fact-Fusion	0.1667	14
MNLP	0.1667	14

Table 4: Rank list of Task 2: Classification of fake news into various categories

This task saw a significant difference in the model performance. In Task 1, top teams achieved macro F1 scores greater than 0.89, whereas the lower-ranked teams attained scores around 0.33. The trend is similar in Task 2, too. The first-ranked team scored an F1 score of 0.6283, and the bottom-ranked teams scored only 0.1667. Most of these differences in performance can be traced back to the approaches that the teams devised with class imbalance and data augmentation, feature extraction, and model architecture.

Class imbalance is a primary challenge pertaining to this task, especially in Task 2. Some of the top-performing teams effectively addressed this issue by employing oversampling algorithms at the feature level. In addition, teams using the adaptive thresholding optimization approach ensured that the models did not overfit to the majority classes. The lower-ranked teams did not employ any mechanism to address the class imbalance issue and hence resulted in the poor generalization of the minority class data samples.

The selection of the model architectures played a pivotal role in the performance. The majority of the top-ranked teams leveraged the transformer-based architectures, which provided better context understanding of the data compared to the traditional machine learning classifiers and feature extraction approaches. The transformer-based models excelled because of their multilingual capabilities, making them effective for data in low-resource languages like Malayalam. In addition, these models provide better learning and contextual token embeddings due to their multilingual capability. Traditional feature extraction models such as TF-IDF, bag-of-words, and n-gram-based representations struggled to capture the deep contextual and semantic relationships in fake news content, leading to suboptimal performance. In addition, data augmentation and the use of ensemble models significantly improve their performance. The availability of the computational resources played a major role in determining the performance of the models. The teams who fine-tuned the transformer models using the task achieved better scores compared to the models that did not use it.

For the most part, the best teams use data enhancement techniques, transformer-based architectures, ensemble methods, and computing resources well. The most effective models incorporated oversampling for class balance, transliteration-aware augmentation, hybrid architectures combining transformers and LSTMs, and multi-layered pooling strategies. In contrast, teams that relied on simpler machine learning models, failed to address class imbalance, or lacked data augmentation strategies struggled to achieve competitive results. These findings highlight the importance of adaptive learning techniques and advanced model enhancement strategies for tackling complex NLP tasks like fake news detection.

6 Conclusion

This paper presents a summary of the shared task "Fake News Detection in Dravidian Languages - DravidianLangTech@NAACL 2025," which focuses on the Malayalam language. The task provided an opportunity to assess the efficacy of several machine learning and deep learning algorithms in detecting fake news on social media. Transformer-based architectures, particularly XLM-RoBERTa, consistently outperformed traditional machine learning algorithms, with the highest macro F1-scores in both binary and multi-class classification tasks. These findings give the promise of advanced NLP models in handling fake news and emphasize the significance of continued research and model improvement to enhance accuracy.

Acknowledgments

This work was conducted with the financial support from Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289_P2(Insight_2), supported in part of Science Foundation Ireland Centre for Research Training in Artificial Intelligence under Grant No. 18/CRT/6223.

References

Tewodros Achamaleh, Nida Hafeez, Mikiyas Mebraihtu, Fatima Uroosa, and Grigori Sidorov. 2025. CIC-NLP@DravidianLangTech 2025: Fake News Detection in Dravidian Languages. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Shriya Alladi and Bharathi B. 2025. Cognitext@DravidianLangTech2025: Fake News Classification in Malayalam Using mBERT and LSTM. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Mahfuz Ahmed Anik, Md. Iqramul Hoque, Wahid Faisal, Azmine ToushikWasi, and Md Manjurul Ahsan. 2025. Akatsuki-CIOL@DravidianLangTech 2025: Ensemble-Based Approach Using Pre-Trained Models for Fake News Detection in Dravidian Languages. In *Proceedings of the Fifth Workshop on*

Speech, Vision, and Language Technologies for Dravidian Languages. Association for Computational Linguistics.

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.

Dola Chakraborty, Shamima Afroz, Jawad Hossain, and Mohammed Moshiul Hoque. 2025. One_by_zero@DravidianLangTech 2025: Fake News Detection in Malayalam Language Leveraging Transformer-based Approach. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Sharal Coelho, Asha Hegde, G Kavya, and Hosahalli Lakshmaiah Shashirekha. 2023. Mucs@ dravidianlangtech2023: Malayalam fake news detection using machine learning approach. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 288–292.

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.

Salman Farsi, Asrarul Eusha, Ariful Islam, Hasan Mesbaul Ali Taher, Jawad Hossain, Shawly Ahsan, Avishek Das, and Mohammed Moshiul Hoque. 2024. CUET_Binary_Hackers@ DravidianLangTech EACL2024: Fake News Detection in Malayalam Language Leveraging Fine-tuned MuRIL BERT. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 173–179.

Anusha M D Gowda and Parameshwar R Hegde. 2025. YenCS@DravidianLangTech 2025: Integrating Hybrid Architectures for Fake News Detection in Low-Resource Dravidian Languages. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Fariha Haq, Md. Tanvir Ahammed Shawon, Md Ayon Mia, Golam Sarwar Md. Mursalin, and Muhammad Ibrahim Khan. 2025. KCRL@DravidianLangTech 2025: Multi-Pooling Feature Fusion with XLM-RoBERTa for Malayalam Fake News Detection and Classification. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Md Minhazul Kabir, Md. Mohiuddin, Kawsar Ahmed, and Moshiul Mohammed Hoque. 2025. CUET-NLP_MP@DravidianLangTech 2025: A Transformer and LLM-Based Ensemble Approach for Fake News Detection in Dravidian. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Rohith Gowtham Kodali and Durga Prasad Manukonda. 2025. byteSizedLLM@DravidianLangTech 2025: Fake News Detection in Dravidian Languages Using Transliteration-Aware XLM-RoBERTa and Attention-BiLSTM. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

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.

Durga Prasad Manukonda and Rohith Gowtham Kodali. 2025. byteSizedLLM@DravidianLangTech 2025: Fake News Detection in Dravidian Languages Using Transliteration-Aware XLM-RoBERTa and Transformer Encoder-Decoder. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Hamada Nayel, Mohammed Aldawsari, and Hosahalli Lakshmaiah Shashirekha. 2025. NAYEL@DravidianLangTech-2025: Character N-gram and Machine Learning Coordination for Fake News Detection in Dravidian Languages. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Syeda Alisha Noor, Sadia Anjum, Syed Ahmad Reza, and Md Rashadur Rahman. 2025. Celestia@DravidianLangTech 2025: Malayalam-BERT and m-BERT based transformer models for Fake News Detection in Dravidian Languages. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Md Osama, Kawsar Ahmed, Hasan Mesbaul Ali Taher, Jawad Hossain, Shawly Ahsan, and Mohammed Moshiul Hoque. 2024. CUET_NLP_GoodFellows@ DravidianLangTech EACL2024: A Transformer-Based Approach for Detecting Fake News in Dravidian Languages. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 187–192.

Tanzim Rahman, Abu Raihan, Md Rahman, Jawad Hossain, Shawly Ahsan, Avishek Das, and Mohammed Moshiul Hoque. 2024. CUET_DUO@ DravidianLangTech EACL2024: Fake News Classification Using Malayalam-BERT. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 223–228.

Eduri Raja, Badal Soni, and Sami Kumar Borgohain. 2023a. nlpt malayalm@ DravidianLangTech: Fake news detection in Malayalam using optimized XLM-RoBERTa model. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 186–191.

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

Eduri Raja, Badal Soni, and Samir Kumar Borgohain. 2024. Fake news detection in Dravidian languages using multiscale residual CNN_BiLSTM hybrid model. *Expert Systems with Applications*, 250:123967.

Nazmus Sakib, Md. Refaj Hossan, Alamgir Hossain, Jawad Hossain, and Mohammed Moshiul Hoque. 2025. CUET-NLP_Big_O@DravidianLangTech 2025: A BERT-based Approach to Detect Fake News from Malayalam Social Media Texts. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Kogilavani Shanmugavadivel, Malliga Subramanian, Vishali K S, Priyanka B, and Naveen Kumar K. 2025a. KEC_AI_DATA_DRIFTERS@DravidianLangTech 2025: Fake News Detection in Dravidian Languages. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Kogilavani Shanmugavadivel, Malliga Subramanian, Aiswarya M, Aruna T, and Jeevaananth S. 2025b. BlueRay@DravidianLangTech-2025: Fake News Detection in Dravidian Languages. 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.

Nafisa Tabassum, Sumaiya Aodhora, Rowshon Akter, Jawad Hossain, Shawly Ahsan, and Mohammed Moshiul Hoque. 2024. Punny_punctuators@ dravidianlangtech-eacl2024: Transformer-based approach for detection and classification of fake news in malayalam social media text. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 180–186.

Srihari V K, Vijay Karthick Vaidyanathan, and Thenmozhi Durairaj. 2025. NLP_goats@DravidianLangTech 2025: Detecting Fake News in Dravidian Languages: A Text Classification Approach. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

M Zamir, M Tash, Z Ahani, A Gelbukh, and G Sidorov. 2024. Tayyab@ dravidianlangtech 2024: detecting fake news in malayalam lstm approach and challenges. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 113–118.