Findings of the Shared Task on Multimodal Abusive Language Detection and Sentiment Analysis in Tamil and Malayalam

Premjith B[1], Jyothish Lal G[1], Sowmya V[1], Bharathi Raja Chakravarthi[2], Rajeswari Natarajan[3], K Nandhini[4], Abirami Murugappan[5], Bharathi B[6], Kaushik M[1], Prasanth SN[1], R Aswin Raj[1], Vijai Simmon S[1]

[1] Amrita School of Artificial Intelligence, Coimbatore, Amrita Vishwa Vidyapeetham, India
[2]Insight SFI Research Centre for Data Analytics, School of Computer Science, University of Galway, Ireland
[3]SRC Campus, SASTRA Deemed to be University, Kumbakonam, Tamilnadu, India
[4]School of Mathematics and Computer Sciences, Central University of Tamil Nadu
[5]Department of Information Science and Technology, Anna University
[6] SSN college of Engineering, Tamilnadu

Abstract

This paper summarizes the shared task on multimodal abusive language detection and sentiment analysis in Dravidian languages as part of the third Workshop on Speech and Language Technologies for Dravidian Languages at RANLP 2023. This shared task provides a platform for researchers worldwide to submit their models on two crucial social media data analysis problems in Dravidian languages - abusive language detection and sentiment analysis. Abusive language detection identifies social media content with abusive information, whereas sentiment analysis refers to the problem of determining the sentiments expressed in a text. This task aims to build models for detecting abusive content and analyzing fine-grained sentiment from multimodal data in Tamil and Malayalam. The multimodal data consists of three modalities - video, audio and text. The datasets for both tasks were prepared by collecting videos from YouTube. Sixty teams participated in both tasks. However, only two teams submitted their results. The submissions were evaluated using macro F1-score.

1 Introduction

Multimodal social media data analyses the insights from social media data that include multiple modalities such as text, audio, and videos. Conventional social media data involves text data such as tweets, Facebook posts, and YouTube comments, and social media data analysis mainly focuses on processing text data to obtain valuable information. However, multimodal data analysis considers the diverse content shared on various social media platforms (Chakravarthi et al., 2021). Analyzing the data containing multiple modalities can give a more comprehensive understanding of user behavior, opinions, and trends. The features extracted from video and audio data can be considered additional information to enhance the text features of input data. The facial expressions from the video, the pitch, and the tone of the audio signal can help to refine the features for identifying different opinions and expressions more effectively.

Multimodal social media data analysis combines approaches from different domains, including computer vision (CV), speech processing, and natural language processing (NLP), to process and detect different categories of data.

- *Text mining and analysis*: The text data related to the videos posted on different social media platforms can be mined and analysed using NLP techniques to identify the context, sentiment, hate comments, comments with offensive language, and abusive content.
- *Computer vision*: This task involves the analysis of images and videos to understand different categories of data. This analysis can include detecting different objects, scenes, facial expressions, and emotions of people from images and videos.
- *Speech analysis*: Analysis of speech data, specifically spoken words, to extract information, sentiment, or other relevant features.
- *Feature fusion*: Integrating and combining features extracted from different modalities to obtain an extensive understanding of multimodal data.

Third Workshop on Speech and Language Technologies for Dravidian Languages, pages 72–79, Varna, Bulgaria, Sep 7, 2023.
https://doi.org/10.26615/978-954-452-085-4_010

The shared task on multimodal abusive language detection and sentiment analysis in Dravidian languages at DravidianLangTech, organized as part of Recent Advancements in Natural Language Processing 2023, has two subtasks - multimodal abusive language detection in Tamil and multimodal sentiment analysis in Tamil and Malayalam. The main objective of the shared task is to encourage researchers around the globe to participate and submit their approaches and results, with the motivation to support the research in resource-poor languages like Tamil and Malayalam.

Sentiment Analysis (opinion mining) (Medhat et al., 2014), (Chakravarthi et al., 2020b), (Priyadharshini et al., 2021) identifies sentiments or opinions by analyzing the underlying subjective information in the text data. Generally, sentiment analysis involves understanding the emotional tone present in a text to classify it as positive, negative, neutral, or more fine-grained categories. This analysis utilizes NLP embedding algorithms to represent text as a feature vector and classify the text into different categories using machine learning (ML) or deep learning (DL) algorithms. The multimodal sentiment analysis from video data involves the analysis of video or image sequence analysis and speech analysis in addition to the text analysis to identify the sentiments expressed in the data. In this task, the data are in Tamil and Malayalam, two low-resource languages. Both languages are morphologically rich and agglutinative, which makes the analysis more complex. In addition, the code-mixing property, which usually occurs in social media data, increases the complexity as it covers words from multiple languages.

Abusive language detection from social media data is identifying comments with abusive content (Nobata et al., 2016), (Priyadharshini et al., 2022b), (Priyadharshini et al., 2022a), (Chakravarthi et al., 2023), (Prasanth et al., 2022). Generally, in text-based analysis, the algorithms look for the words or phrases (sequence of words) or similar words/phrases that define the abusiveness of the input text. Multimodal abusive language detection includes detecting videos containing abusive information by analyzing video, speech, and text data. This subtask is primarily focused on building models for multimodal data in Tamil.

This paper summarises the shared task on multimodal abusive language detection in Tamil and multimodal sentiment analysis in Tamil and Malayalam. Besides, this paper discusses the findings from the models submitted to the two subtasks mentioned above. The shared task was hosted on CodaLab [1]. We shared the training data and validation data with all the registered participants for building their models. Later, the test data without labels were shared to make predictions using the models built. Sixty teams registered for both subtasks. However, only two teams submitted their results.

The structure of the paper is as follows: Section 2 discusses the research papers published in the domains of multimodal abusive language detection and multimodal sentiment analysis, Section 3 describes the shared tasks, followed by Section 4, which summarizes the systems submitted by participating teams. The paper concludes in Section 5.

2 Related Works

People post messages and comments on various social media platforms in their mother tongue or code-mixed languages. Hence, the machine learning models built with monolingual datasets are unsuitable for identifying abusive language or analyzing sentiments from code-mixed languages. However, researchers are now progressing to develop systems using code-mixed datasets. As stated above, the data collection and annotation process is one key challenge. (Chakravarthi et al., 2020b,a; Hande et al., 2021a; Mandl et al., 2020) released a few Dravidian language datasets for offensive language and hate speech detection. Many models have been proposed using the above-cited datasets (Saumya et al., 2021; Yasaswini et al., 2021; Hande et al., 2021b; Kedia and Nandy, 2021; Renjit and Idicula, 2020; Chakravarthi et al., 2022). The authors (Singh and Bhattacharyya, 2020) have employed an ensemble of multilingual BERT models to detect hate speech and offensive content for Dravidian languages. They have achieved an F-score of 0.95 for hate speech and offensive content detection tasks. in YouTube comments in Malayalam language(code-mixed, script-mixed). For hate speech and offensive content detection tasks for YouTube or Twitter datasets in Malayalam (code-mixed: Tanglish and Manglish), F-scores of 0.86 and 0.72 were achieved, respectively. (Ranasinghe et al., 2020) experimented with multinomial Naive Bayes, support vector machines and random for-

[1] https://codalab.lisn.upsaclay.fr/competitions/11092

est methods to identify the offensive comments in code-mixed Malayalam YouTube comments. They have also used cross-lingual contextual word embeddings and transfer learning models to predict hate and offensive speech in Malayalam data. A weighted average F1-score of 0.89 was achieved.

An ensemble of multilingual transformer networks like XLMRoBERTa has been proposed for offensive speech detection in code-mixed and Romanised variants of three Dravidian languages-Tamil, Malayalam, and Kannada (Sai and Sharma, 2021). Two DL frameworks, namely CNN and Bi-LSTM (Saumya et al., 2021), have been used in parallel to extract the contextual features from the text. These features were concatenated and presented to a fully connected layer for final prediction. They also used conventional machine learning approaches, such as naive Bayes, random forests, support vector machine and transformer-based models to predict content from Dravidian code-mixed scripts. (Yasaswini et al., 2021) explored various transformer models to detect the offensive language in social media posts in Tamil, Malayalam, and Kannada. They achieved F1-scores of 0.9603 and 0.7895 on the Malayalam and Tamil datasets by using the ULMFiT model, respectively. For the Kannada dataset, they obtained an F1-score of 0.7277 by using the distilmBERT model. Two multi-task learning approaches (Hande et al., 2021a) was developed to identify the offensive language for Kannada, Malayalam, and Tamil and analyse the sentiments. The model obtained a weighted F1-score of (59% and 70%), (66.8% and 90.5%), and (62.1% and 75.3%) for Malayalam, Kannada, and Tamil on sentiment analysis and offensive language identification tasks, respectively. An offensive content classification model (Kedia and Nandy, 2021) was built with transformer-based models, namely BERT and RoBERTa, for Dravidian code-mixed languages like Kannada, Malayalam and Tamil. The authors achieved a weighted average F1-scores of 0.72, 0.77, and 0.97 for Kannada-English, Tamil-English, and Malayalam-English datasets.

The authors (Chakravarthi et al., 2022) developed datasets using the YouTube comments of three Dravidian languages, namely Malayalam–English (20,000), Tamil–English (44,000), and Kannada–English (7000) for sentiment analysis and offensive language identification tasks. (Dave et al., 2021) used machine learning and transformer-based models for offensive language identification. The authors represented the sentences using character n-gram and pre-trained word embedding. The F1-scores were 0.95 and 0.71 for Malayalam and Tamil, respectively. (Li, 2021; Dowlagar and Mamidi, 2021; Zhao and Tao, 2021; Chen and Kong, 2021; Dave et al., 2021) have used transformer-based models, such as BERT, RoBERTa and MuRiL for offensive language identification task for Dravidian languages. Besides, research articles (Dowlagar and Mamidi, 2021; Andrew, 2021) have shown the performances of offensive language detection from code-mixed languages using machine learning models such as logistic regression, K-nearest neighbour, support vector machine, decision trees and random forests. (Zhao and Tao, 2021; Chen and Kong, 2021; Sreelakshmi et al., 2021; Sharif et al., 2021) used deep learning techniques for offensive language identification task. All the above-discussed articles cover various methods and evaluations for offensive language identification and hate speech detection. They provide insights into this research's ensemble strategies and deep learning models. Nevertheless, research has to progress deeper and wider to develop a robust model for hate speech and offensive language identification tasks using Dravidian languages. Even though text-based datasets and models are available for abusive language detection and sentiment analysis in Dravidian languages, research on multimodal datasets is yet to kickstart.

3 Task Description

This section discusses the two subtasks - multimodal abusive language detection in Tamil and multimodal sentiment analysis in Tamil and Malayalam, including the dataset used. The subtask, "Multimodal abusive language detection in Tamil", aims to encourage researchers to develop AI/machine learning/deep learning models for classifying video data posted on YouTube into abusive and non-abusive categories. Multimodal sentiment analysis in Tamil and Malayalam" subtask has two tasks - one in Tamil and another in Malayalam. In both tasks, the objective is to develop AI/machine learning/deep learning models for classifying video data into highly positive, positive, neutral, negative, and highly negative categories. In all the tasks mentioned above, video, audio, and text data were provided, and the participants were free to use any combination of modalities to build their models.

The dataset used for conducting both tasks is insufficient for training a machine learning or deep learning model. However, the availability of pre-trained models can help the participants to resolve this problem by generating meaningful feature representations.

3.1 Multimodal Abusive Language Detection in Tamil

The competition was hosted on the CodaLab platform. We provided training data and test data for this competition. Training data contains 70 videos collected from YouTube with abusive and non-abusive content. We extracted audio signals from the video and prepared the transcripts using the Google automatic speech recognition (ASR) module. The errors in the transcripts were corrected manually in the postprocessing step. After that, 88 videos were labeled with the help of qualified native speakers into two categories - abusive and non-abusive (Ashraf et al., 2021).

- **Abusive**: Abuse content encompasses any communication that
    1. targets an individual, group, or community
    2. is disrespectful, sexist, crude, or obscene
    3. pertains to human shortcomings, aims to provoke offense towards an individual or group, or suggests condescension or victim-blaming.
- **Non-abusive**: Anything that do not belong to the abusive category.

We divided the dataset into training and test data. The training dataset consisted of 70 videos, and the test data contained 18 videos. In addition to videos, both datasets are composed of audio and text data. The training dataset consists of 38 videos in the abusive category and 32 in the non-abusive category. The number of data points in each class shows a slight class imbalance problem. The test data consists of 18 videos, of which nine are abusive, and nine are non-abusive. During the testing phase of the competition, we provided the test data without labels. However, we released the test data with labels after the closure of the competition. Table 1 describes the training and testing data details.

Despite sixty registrations, only two teams submitted their results through the Google form provided.

Table 1: Details of the dataset used for the shared task on multimodal abusive language detection in Tamil

Dataset	Abusive	Non-abusive
Training	38	32
Test	9	9

3.2 Multimodal Sentiment Analysis in Tamil and Malayalam

This subtask is also hosted on the CodaLab platform. This is the second edition of this subtak (Premjith et al., 2022). This subtask has two sections - one for Tamil and another for Malayalam. Like the multimodal abusive language detection task, we collected data for this task from YouTube. The same procedure was followed for collecting the data and annotation. Unlike the abusive language detection task, we provided the participants with training, validation, and test data. Training and validation data were released together, and test data without labels were supplied during the testing phase. The test data without labels were uploaded to the CodaLab after completing the task.

We considered five fine-grained labels for annotating each data point in both languages - Highly Positive, Positive, Neutral, Negative, and Highly Negative.

- **Highly Positive**: A video featuring a reviewer using exaggerated language or expressions.
- **Positive**: A video in which the reviewer delivers positive reviews while maintaining subtle facial expressions
- **Neutral**: There are no direct or indirect indications of the speaker's emotional state. Examples include requests for likes or subscriptions and inquiries about a movie's release date or dialogue.
- **Negative**: Videos featuring the utilization of negative words and sarcastic comments, coupled with understated facial expressions.
- **Highly Negative**: Videos where excessively negative words are employed, accompanied by a gloomy facial expression and a stressed voice.

Training, validation, and test data in both languages contain data from three modalities - video, speech, and text. Tamil data consisted of 64 data

Table 2: Details of the dataset used for the shared task on multimodal Sentiment Analysis in Tamil and Malayalam

Dataset	Tamil	Malayalam
Training	44	50
Validation	10	10
Test	10	10

Table 3: Class-wise distribution of the dataset used for the shared task on multimodal Sentiment Analysis in Tamil and Malayalam

Category	Tamil	Malayalam
Highly Positive	8	9
Positive	38	39
Neutral	8	8
Negative	5	12
Highly Negative	5	2
Total	64	70

samples split into training, validation, and test data in the ratio of 22:5:5. Malayalam corpus had 70 data samples, of which 50 were used as training data, ten each as validation and test data. A detailed description of the split is given in Table 3. Class-wise distribution of the data points in both languages is provided in Table 2. The tables 4 and 5 give the class-wise distribution of the data points used in the training, validation, and test datasets. From the dataset details, it is evident that there is a high-class imbalance problem. The positive category has significantly more data points than other categories.

Similar to the previous task, we had 60 registrations for this task, and only two teams submitted their predictions.

Table 4: Distribution of training, validation, and test datasets used for the shared task on multimodal Sentiment Analysis in Tamil

Category	Train	Validation	Test
Highly Positive	5	3	1
Positive	29	4	5
Neutral	4	2	2
Negative	3	1	1
Highly Negative	3	0	1
Total	44	10	10

Table 5: Distribution of training, validation, and test datasets used for the shared task on multimodal Sentiment Analysis in Malayalam

Category	Train	Validation	Test
Highly Positive	5	2	2
Positive	31	5	3
Neutral	5	1	2
Negative	8	2	2
Highly Negative	1	0	1
Total	50	10	10

4 System Description

We received two submissions for both subtasks. Each team was allowed to submit a maximum of three runs. The run with the highest macro F1 score was considered for preparing the rank list. The descriptions of the systems submitted for the shared tasks are given below.

4.1 Team: hate-alert

Their work involved utilizing three distinct models to extract features from various modalities, such as text, audio, and video (Barman and Das, 2023). Firstly, the authors used the BERT model to extract text-based features from the video. By employing the BERT model, they capture and analyze the textual information and characteristics embedded within the video. Additionally, they incorporated the Mel-frequency cepstral coefficients (MFCC) based feature extraction technique to extract audio-based features from the videos. The advantage of MFCCs is that they will aid in capturing the shape of the vocal tract, giving distinct characteristics to the sound or audio. Moreover, the successful extraction of video-based features was achieved by utilizing the Vision Transformer model. It is a specialized deep learning model designed for processing visual data. This model demonstrated its efficacy in handling visual tasks, enabling us to obtain valuable visual features from the videos. The extracted features included significant visual patterns, objects, and informative content. Furthermore, these features significantly contributed to a comprehensive understanding of the visual aspects of the video data. The proposal entailed applying a fusion-based system or a multimodal model after extracting the specific features from each modality (text, audio, and video). The model utilizes the complementary information or features received from different modalities to make accurate predic-

Table 6: Ranklist for the shared task on multimodal abusive language detection in Tamil

Team	Macro F1	Rank
hate-alert	0.5786	1
AbhiPaw	0.3333	2

Table 7: Ranklist for the shared task on multimodal sentiment analysis in Tamil

Team	Macro F1	Rank
hate-alert	0.1429	1
AbhiPaw	0.1333	2

tions. With this approach, the authors categorized the video based on sub-tasks, considering the amalgamated information from text, audio, and video features. Overall, their approach demonstrates a comprehensive and robust video analysis and classification framework.

4.2 Team: AbhiPaw

In this work (Bala and Krishnamurthy, 2023), the authors created multimodal Transformer-based architecture to make accurate predictions. The proposal drew inspiration from two primary sources: (a) ”MISA: Modality-Invariant and -Specific Representations for Multimodal Sentiment Analysis” presented at ACM MM 2020 (Hazarika et al., 2020), and (b) ”How you feelin'? Learning Emotions and Mental States in Movie Scenes” presented at CVPR 2023 (Srivastava et al., 2023). They employed different models with varying dimensions to extract features from all three modalities (video, audio, and text). They used the MVIT model for video data, the openl3 library for audio data, and the BERT-based multilingual model for text data. These models were selected based on their respective strengths and capabilities in processing their corresponding modalities. Further, they designed a 2-layer transformer-based architecture for the primary model. Additionally, they incorporated type embeddings for the features, drawing inspiration from the work (b). These types of embeddings were used to signify the modality of each feature, distinguishing between video, audio, and text. The model's raw output consisted of logits corresponding to the predicted classes.

5 Conclusion

This paper reports the findings of the shared task on multimodal abusive language detection in Tamil

Table 8: Ranklist for the shared task on multimodal sentiment analysis in Malayalam

Team	Macro F1	Rank
hate-alert	0.1889	1
AbhiPaw	0.0923	2

and multimodal sentiment analysis in Tamil and Malayalam. The task dataset consisted of videos collected from YouTube and corresponding audio and transcripts. We received sixty registrations for both subtasks. However, only two teams submitted the predictions for the test data provided to the participants. We used macro F1-score to assess the submitted predictions' performance and prepare the rank list.

Acknowledgments

The author Bharathi Raja Chakravarthi was supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289_P2(Insight_2).

References

Judith Jeyafreeda Andrew. 2021. Judithjeyafreedaandrew@ dravidianlangtech-eacl2021: offensive language detection for dravidian code-mixed youtube comments. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 169–174.

Noman Ashraf, Arkaitz Zubiaga, and Alexander Gelbukh. 2021. Abusive language detection in youtube comments leveraging replies as conversational context. *PeerJ Computer Science*, 7:e742.

Abhinaba Bala and Parameswari Krishnamurthy. 2023. Abhipaw@dravidianlangtech: Multimodal abusive language detection and sentiment analysis. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Shubhankar Barman and Mithun Das. 2023. hate-alert@dravidianlangtech: Multimodal abusive language detection and sentiment analysis in dravidian languages. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Bharathi Raja Chakravarthi, Anand Kumar M, John P McCrae, B Premjith, KP Soman, and Thomas Mandl. 2020a. Overview of the track on hasoc-offensive language identification-dravidiancodemix. In *FIRE (Working notes)*, pages 112–120.

Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, and John P McCrae. 2020b. Corpus creation for sentiment analysis in code-mixed tamil-english text. *arXiv preprint arXiv:2006.00206*.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Shubanker Banerjee, Manoj Balaji Jagadeeshan, Prasanna Kumar Kumaresan, Rahul Ponnusamy, Sean Benhur, and John Philip McCrae. 2023. Detecting abusive comments at a fine-grained level in a low-resource language. *Natural Language Processing Journal*, 3:100006.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Navya Jose, Shardul Suryawanshi, Elizabeth Sherly, and John P McCrae. 2022. Dravidiancodemix: Sentiment analysis and offensive language identification dataset for dravidian languages in code-mixed text. *Language Resources and Evaluation*, 56(3):765–806.

Bharathi Raja Chakravarthi, KP Soman, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Kingston Pal Thamburaj, John P McCrae, et al. 2021. Dravidianmultimodality: A dataset for multi-modal sentiment analysis in tamil and malayalam. *arXiv preprint arXiv:2106.04853*.

Shi Chen and Bing Kong. 2021. cs@ dravidianlangtech-eacl2021: Offensive language identification based on multilingual bert model. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 230–235.

Bhargav Dave, Shripad Bhat, and Prasenjit Majumder. 2021. Irnlp daiict@ dravidianlangtech-eacl2021:+ offensive language identification in dravidian languages using tf-idf char n-grams and muril. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 266–269.

Suman Dowlagar and Radhika Mamidi. 2021. Offlangone@ dravidianlangtech-eacl2021: Transformers with the class balanced loss for offensive language identification in dravidian code-mixed text. In *Proceedings of the first workshop on speech and language technologies for dravidian languages*, pages 154–159.

Adeep Hande, Siddhanth U Hegde, Ruba Priyadharshini, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Sajeetha Thavareesan, and Bharathi Raja Chakravarthi. 2021a. Benchmarking multi-task learning for sentiment analysis and offensive language identification in under-resourced dravidian languages. *arXiv preprint arXiv:2108.03867*.

Adeep Hande, Karthik Puranik, Konthala Yasaswini, Ruba Priyadharshini, Sajeetha Thavareesan, Anbukkarasi Sampath, Kogilavani Shanmugavadivel, Durairaj Thenmozhi, and Bharathi Raja Chakravarthi. 2021b. Offensive language identification in low-resourced code-mixed dravidian languages using pseudo-labeling. *arXiv preprint arXiv:2108.12177*.

Devamanyu Hazarika, Roger Zimmermann, and Soujanya Poria. 2020. Misa: Modality-invariant and-specific representations for multimodal sentiment analysis. In *Proceedings of the 28th ACM international conference on multimedia*, pages 1122–1131.

Kushal Kedia and Abhilash Nandy. 2021. indicnlp@ kgp at dravidianlangtech-eacl2021: Offensive language identification in dravidian languages. *arXiv preprint arXiv:2102.07150*.

Zichao Li. 2021. Codewithzichao@ dravidianlangtech-eacl2021: Exploring multilingual transformers for offensive language identification on code mixing text. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 164–168.

Thomas Mandl, Sandip Modha, Anand Kumar M, and Bharathi Raja Chakravarthi. 2020. Overview of the hasoc track at fire 2020: Hate speech and offensive language identification in tamil, malayalam, hindi, english and german. In *Forum for information retrieval evaluation*, pages 29–32.

Walaa Medhat, Ahmed Hassan, and Hoda Korashy. 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4):1093–1113.

Chikashi Nobata, Joel Tetreault, Achint Thomas, Yashar Mehdad, and Yi Chang. 2016. Abusive language detection in online user content. In *Proceedings of the 25th international conference on world wide web*, pages 145–153.

SN Prasanth, R Aswin Raj, P Adhithan, B Premjith, and Soman Kp. 2022. Cen-tamil@ dravidianlangtech-acl2022: Abusive comment detection in tamil using tf-idf and random kitchen sink algorithm. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*, pages 70–74.

B Premjith, Bharathi Raja Chakravarthi, Malliga Subramanian, B Bharathi, Soman Kp, V Dhanalakshmi, K Sreelakshmi, Arunaggiri Pandian, and Prasanna Kumaresan. 2022. Findings of the shared task on multimodal sentiment analysis and troll meme classification in dravidian languages. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*, pages 254–260.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Subalalitha Cn, Thenmozhi Durairaj, Malliga Subramanian, Kogilavani Shanmugavadivel, Siddhanth U Hegde, and Prasanna Kumaresan. 2022a. Overview of abusive comment detection in tamil-acl 2022. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*, pages 292–298.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Subalalitha Chinnaudayar Navaneethakrishnan, Thenmozhi Durairaj, Malliga Subramanian, Kogilavani Shanmugavadivel, Siddhanth U Hegde, and

Prasanna Kumar Kumaresan. 2022b. Findings of the shared task on abusive comment detection in tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages. Association for Computational Linguistics*.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Sajeetha Thavareesan, Dhivya Chinnappa, Durairaj Thenmozhi, and Rahul Ponnusamy. 2021. Overview of the dravidiancodemix 2021 shared task on sentiment detection in tamil, malayalam, and kannada. In *Proceedings of the 13th Annual Meeting of the Forum for Information Retrieval Evaluation*, pages 4–6.

Tharindu Ranasinghe, Sarthak Gupte, Marcos Zampieri, and Ifeoma Nwogu. 2020. Wlv-rit at hasoc-dravidian-codemix-fire2020: Offensive language identification in code-switched youtube comments. *arXiv preprint arXiv:2011.00559*.

Sara Renjit and Sumam Mary Idicula. 2020. Cusatnlp@ hasoc-dravidian-codemix-fire2020: identifying offensive language from manglishtweets. *arXiv preprint arXiv:2010.08756*.

Siva Sai and Yashvardhan Sharma. 2021. Towards offensive language identification for dravidian languages. In *Proceedings of the first workshop on speech and language technologies for Dravidian languages*, pages 18–27.

Sunil Saumya, Abhinav Kumar, and Jyoti Prakash Singh. 2021. Offensive language identification in dravidian code mixed social media text. In *Proceedings of the first workshop on speech and language technologies for Dravidian languages*, pages 36–45.

Omar Sharif, Eftekhar Hossain, and Mohammed Moshiul Hoque. 2021. Nlp-cuet@ dravidianlangtech-eacl2021: Offensive language detection from multilingual code-mixed text using transformers. *arXiv preprint arXiv:2103.00455*.

Pankaj Singh and Pushpak Bhattacharyya. 2020. Cfilt iit bombay@ hasoc-dravidian-codemix fire 2020: Assisting ensemble of transformers with random transliteration. In *FIRE (Working Notes)*, pages 411–416.

K Sreelakshmi, B Premjith, and Soman Kp. 2021. Amrita_cen_nlp@ dravidianlangtech-eacl2021: deep learning-based offensive language identification in malayalam, tamil and kannada. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 249–254.

Dhruv Srivastava, Aditya Kumar Singh, and Makarand Tapaswi. 2023. How you feelin'? learning emotions and mental states in movie scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2517–2528.

Konthala Yasaswini, Karthik Puranik, Adeep Hande, Ruba Priyadharshini, Sajeetha Thavareesan, and Bharathi Raja Chakravarthi. 2021. Iiitt@ dravidianlangtech-eacl2021: Transfer learning for offensive language detection in dravidian languages. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 187–194.

Yingjia Zhao and Xin Tao. 2021. Zyj123@ dravidianlangtech-eacl2021: Offensive language identification based on xlm-roberta with dpcnn. In *Proceedings of the first workshop on speech and language technologies for dravidian languages*, pages 216–221.