Findings of the Shared Task on Multimodal Social Media Data Analysis in Dravidian Languages (MSMDA-DL)@DravidianLangTech 2024

Premjith B[1], **Jyothish Lal G**[1], **Sowmya V**[1], **Bharathi Raja Chakravarthi** [2], **K Nandhini**[3]
Rajeswari Natarajan [4], **Abirami Murugappan**[5], **Bharathi B**[6], **Saranya Rajiakodi**[7]
Rahul Ponnusamy[2], **Jayanth Mohan**[1], **Mekapati Spandana Reddy**[1]
Amrita School of Artificial Intelligence, Coimbatore, Amrita Vishwa Vidyapeetham, India
[2]School of Computer Science, University of Galway, Ireland
[3]School of Mathematics and Computer Sciences, Central University of Tamil Nadu, India
[4]SASTRA University, India [5]Department of Information Science and Technology,
Anna University, India [6]SSN College of Engineering, Tamil Nadu, India
[7]Central University of Tamil Nadu, India

Abstract

This paper presents the findings of the shared task on multimodal sentiment analysis, abusive language detection and hate speech detection in Dravidian languages. Through this shared task, researchers worldwide can submit models for three crucial social media data analysis challenges in Dravidian languages: sentiment analysis, abusive language detection, and hate speech detection. The aim is to build models for deriving fine-grained sentiment analysis from multimodal data in Tamil and Malayalam, identifying abusive and hate content from multimodal data in Tamil. Three modalities make up the multimodal data: text, audio, and video. YouTube videos were gathered to create the datasets for the tasks. Thirty-nine teams took part in the competition. However, only two teams, though, turned in their findings. The macro F1-score was used to assess the submissions.

1 Introduction

Analyzing insights from social media data with several modalities—text, audio, and video—is known as multimodal social media data. Text data from sources like Facebook posts, YouTube comments, and tweets make up conventional social media data. The primary goal of social media data analysis is to extract useful information from text data. However, a study published in (Chakravarthi et al., 2021) considers the variety of content posted on social media sites. Multiple modalities in the data can be analyzed to provide a more thorough knowledge of user behaviour, attitudes, and trends. It is possible to think of the features retrieved from audio and video data as extra information that improves the text features of the input data. Pitch, tone, and the video's facial expressions can all be used to fine-tune the elements used to identify various viewpoints and expressions more accurately. To analyze and identify various data categories, multimodal social media data analysis integrates methods from several fields, such as computer vision (CV), speech processing, and natural language processing (NLP).

The Multimodal Social Media Data Analysis in Dravidian Languages (MSMDA-DL) at the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages (DravidianLangTech-2024) at EACL 2024 has three tasks: multimodal sentiment analysis in Tamil and Malayalam; multimodal detect abusive language in Tamil; and multimodal detect hate speech in Tamil. The primary goal of the shared task is to motivate researchers and academicians worldwide to join and submit their methods and findings to help research in languages with limited resources, such as Tamil and Malayalam.

The shared task on Multimodal Social Media Data Analysis in Dravidian Languages (MSMDA-DL) is summarized in this work. Additionally, the results of the submitted models for the three subtasks are discussed in this study. The shared task was hosted on the CodaLab [1]. All enrolled participants received access to training and validation data to construct their models. The test data without labels were later shared to use the developed models to forecast the future. Thirty-nine teams signed up for the two subtasks. Only two teams, though, turned in their findings.

2 Literature Review

Users on many social media sites write messages and comments in their native tongue and codemixed languages. As a result, machine learning models developed using monolingual datasets are inappropriate for classifying abusive language or deciphering the emotions present in languages with mixed coding. However, scientists are making strides toward creating systems with codemixed

[1]https://codalab.lisn.upsaclay.fr/competitions/16093

Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 56–61

datasets. One major problem is, as mentioned above, the process of gathering and annotating data.

2.1 Datasets

To detect hate speech and offensive content, (Chakravarthi et al., 2020b,a; Hande et al., 2021a; Mandl et al., 2020) provided a few Dravidian language datasets. The 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) have been used to suggest several models. For emotion analysis and offensive language detection tasks, the authors (Chakravarthi et al., 2022) created datasets using the YouTube comments of three Dravidian languages: Malayalam–English (20,000), Tamil–English (44,000), and Kannada–English (7000).

2.2 Models

An ensemble of multilingual BERT models was used by the authors (Singh and Bhattacharyya, 2020) to identify objectionable content and hate speech in Dravidian languages. For tasks including detecting hate speech and offensive content, they received an F-score of 0.95. In Malayalam comments on YouTube (mixing code and script). F-scores 0.86 and 0.72 were obtained for hate speech and offensive content detection tasks for YouTube or Twitter datasets in Malayalam (codemixed: Tanglish and Manglish). The obtained weighted average F1-score was 0.89. To identify offensive language, transformer-based models and machine learning were employed by (Dave et al., 2021). The authors used pre-trained word embedding and character n-gram to represent the sentences. For Tamil and Malayalam, the F1 scores were 0.71 and 0.95 respectively. Transformer-based models, including BERT, RoBERTa, and MuRiL, have been employed by (Li, 2021; Dowlagar and Mamidi, 2021; Zhao and Tao, 2021; Chen and Kong, 2021; Dave et al., 2021) for the job of identifying offensive languages for Dravidian languages. Text-based datasets and models are available for sentiment analysis, abusive language detection, and hate speech detection in Dravidian languages. However, multimodal dataset research still needs to be improved.

3 Description of the subtasks

The three subtasks - multimodal sentiment analysis in Tamil and Malayalam, multimodal abusive language identification in Tamil and multimodal hate speech detection in Tamil - as well as the dataset utilized, are covered in this section. There are two tasks in the "Multimodal Sentiment Analysis in Tamil and Malayalam" subtask: one in Tamil and one in Malayalam. Both tasks are modelled as a multiclass classification task with. The subtask "Multimodal abusive language detection in Tamil" is a binary class classification task, whereas the third task, "multimodal hate speech detection in Tamil", is a multiclass classification task. All of the tasks mentioned above included text, audio, and video data, and participants could use any combination of modalities to create their models.

Table 1: 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

3.1 Multimodal Sentiment Analysis in Tamil and Malayalam

This is the third edition of this subtask (Premjith et al., 2023), (Premjith et al., 2022). There are two sections in this subtask: one for Tamil and another for Malayalam (Chakravarthi et al., 2021). We considered YouTube to gather data for this work. We gave the participants test, validation, and training data. Data for training and validation were made available simultaneously, and unlabeled test data was provided during the testing stage. The data points were annotated with five labels in both languages: Highly Positive, Positive, Neutral, Negative, and Highly Negative.

The 64 data samples in the Tamil data were divided into training, validation, and test data in a 22:5:5 ratio. There were 70 data samples in the Malayalam corpus, of which 50 were used for training and 10 for each validation and testing. The divide is explained in depth in Table 2, which 2 shows the class-wise distribution of the data points in both languages. The class-wise distribution of the data points utilized in the training, validation, and test datasets is provided in the Tables 3 and 4. It is clear from the dataset specifics that there is an issue with high-class imbalance. There are substantially more data points in the positive category.

Table 2: 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

Table 3: 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 4: 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

Table 5: 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 Abusive Language Detection in Tamil

This is the second edition of this task (Premjith et al., 2023). We supplied test and training data for this competition. Seventy YouTube videos, both abusive and non-abusive, comprise the training data. The dataset collection process is similar to that of the sentiment analysis task. Following that, 88 films were classified as abusive or non-abusive using the assistance of qualified native speakers (Ashraf et al., 2021).

The dataset was split into test and training subsets. Eighteen videos were in the test dataset, and seventy in the training dataset. Both datasets contain text and audio data in addition to videos. There are 38 videos in the abusive category and 32 in the non-abusive category in the training dataset. The quantity of data points in every class indicates a little issue with class imbalance. There are nine abusive and nine non-abusive videos in the set of 18 test videos. We sent the test data without labels for the competition's testing phase. However, following the competition's conclusion, we made the test data with labels available. Table 5 details the training and testing data.

3.3 Multimodal Hate Speech Detection in Tamil

The hate speech detection task is a multiclass classification problem, where the data points are labelled into four categories: caste, offensive, racist, and sexist. The training data comprised 40 samples, whereas the test data comprised 12. The class-wise distribution of data in both train and test data is shown in Tables 6 and 7.

4 System Description

We received two submissions; only one team participated in all three tasks. Each team could submit up to three runs. The run with the highest macro F1 score was considered when creating the rank list. Below are the system descriptions that were submitted for the shared tasks.

Table 6: Details of the training dataset used for the shared task on multimodal hate speech detection in Tamil

Class	# Data points
Caste	12
Offensive	13
Racist	12
Sexist	3
Total	40

Table 7: Details of the test dataset used for the shared task on multimodal hate speech detection in Tamil

Class	# Data points
Caste	3
Offensive	4
Racist	4
Sexist	1
Total	12

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

Team	Macro F1	Rank
Wit Hub	0.2444	1

Table 9: Ranklist for the shared task on multimodal abusive language detection in Tamil

Team	Macro F1	Rank
Binary_Beasts	0.7143	1
Wit Hub	0.4156	2

Table 10: Ranklist for the shared task on multimodal hate speech detection in Tamil

Team	Macro F1	Rank
Wit Hub	0.2881	1

4.1 Binary_Beasts

The team (Rahman et al., 2024) participated only in abusive language detection. The team used ConvLSTM for the video dataset, and for the audio dataset, they used BiLSTM and multinomial naive Bayes for the text dataset. They did not use any external data to build the model. This team developed three models. From these models, they counted the majority-based result for the final output.

4.2 Wit Hub

The team (HS et al., 2024) used three different models for each subtask. They considered only text data for the analysis. For the sentiment analysis task, the team used LSTM, K-means, KNN and logistic regression models for classification. TF-IDF features and Multinomial Naive Bayes classifier were used to train the model for categorising data into abusive and non-abusive categories. In contrast, TF-IDF and random forest feature-classifier combination were considered for the hate speech identification task. This team did not use any external data to train the model.

The rank lists for multimodal sentiment analysis in Tamil, multimodal abusive language detection in Tamil, and multimodal hate speech detection in Tamil are shown in Tables 8, 9 and 10, respectively.

Among the submitted models, only Binary_Beats used three modalities to develop their model. The ConvLSTM model used by this team captures the features from the image frames and the LSTM can model the sequentiality of the image frames, which is an interesting approach. They considered LSTM for capturing the sequential properties of other modalities. The second team, Wit Hub relied on text data and extracted TF-IDF features for classification. From the results, it is evident that TF-IDF features are not good enough to classify the data into different categories for the data used in this shared task.

5 Conclusion

The results of the shared task on Multimodal Social Media Data Analysis in Dravidian Languages (MSMDA-DL) are presented in this study. The task dataset was made up of transcripts and audio that went along with videos that were gathered from YouTube. Thirty-nine people signed up for each of the three subtasks. Nevertheless, only two teams turned in their predictions for the test data that the participants were given. To create the rank

list and evaluate the performance of the submitted forecasts, we employed the macro F1-score.

6 Acknowledgement

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

References

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

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, 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*.

Anierudh HS, Abhishek R, Ashwin V Sundar, Amrit Krishnan, and Bharathi B. 2024. Wit Hub@DravidianLangTech-2024:Multimodal Social Media Data Analysis in Dravidian Languages using Machine Learning Models. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, Malta. European Chapter of the Association for Computational Linguistics.

Kushal Kedia and Abhilash Nandy. 2021. indic-nlp@ 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.

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.

B Premjith, V Sowmya, Bharathi Raja Chakravarthi, Rajeswari Natarajan, K Nandhini, Abirami Murugappan, B Bharathi, M Kaushik, Prasanth Sn, et al.

2023. Findings of the shared task on multimodal abusive language detection and sentiment analysis in Tamil and Malayalam. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 72–79.

Md. Tanvir Rahman, Abu Bakkar Siddique Raihan, Tanzim Rahman, Shawly Ahsan, Jawad Hossain, Avishek Das, and Mohammed Moshiul Hoque. 2024. Binary_Beasts@DravidianLangTech@EACL 2024: Multimodal Abusive Language Detection in Tamil based on Integrated Approach of Machine Learning and Deep Learning Techniques. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, Malta. European Chapter of the Association for Computational Linguistics.

Sara Renjit and Sumam Mary Idicula. 2020. CUSATNLP@HASOC-Dravidian-CodeMix-FIRE2020:Identifying Offensive Language from Manglish Tweets. *arXiv preprint arXiv:2010.08756*.

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.

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.

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.