Overview of the Shared Task on Multilevel Political Meme Classification in Tamil and Malayalam

Saranya Rajiakodi[1], Shunmuga Priya Muthusamy Chinnan[2], Premjith B[3], Subalalitha CN[4], Rahul Ponnusamy[2], Anshid K A[5], Bhuvaneswari Sivagnanam[1], Jananayagan V[1], Ragavan N[1], Santhini P[1], Bharathi Raja Chakravarthi[2]

[1]Central University of Tamil Nadu, India,
[2]Data Science Institute, University of Galway, Ireland
[3]Amrita Vishwa Vidhyapeetham, Tamil Nadu, India
[4]SRMIST, Tamil Nadu, India, [5]Mar Thoma College, Kerala, India

Abstract

This paper presents an overview of the Multi-Level Political Meme Classification shared task conducted at DravidianLangTech–ACL 2026. The task introduces a hierarchical two-level classification framework for Tamil and Malayalam political memes: Level 1 focuses on stance detection (Support/Praise vs. Troll/Oppose), while Level 2 identifies the political target (individual or party), conditioned on the predicted stance. The dataset was curated from social media platforms and manually annotated with strong inter-annotator agreement. A total of 64 teams registered for the shared task, of which 19 teams submitted official results using diverse multimodal approaches, including transformer-based text encoders, vision models, optical character recognition (OCR) pipelines, and hierarchical architectures. The results show that Level 1 stance detection achieved high macro-F1 scores across both Tamil and Malayalam, with several systems exceeding 0.96 F1, whereas Level 2 target identification remained substantially more challenging, especially for Malayalam. These findings highlight the importance of multimodal fusion, hierarchical reasoning, and robustness to OCR noise, cultural context, and class imbalance in political meme analysis.

Disclaimer: This paper (including figures, memes, and examples) contains political content and may include offensive, harmful, or derogatory language and imagery. All such material is reproduced solely for research, annotation, and educational purposes to study multi-level political meme classification. It does not represent the authors' views or endorsement of any ideology, party, or individual. Reader discretion is advised.

1 Introduction

People use social media platforms for communication and sharing their opinions on different domains including politics. Moreover, visual content, for example memes, play an important role in shaping public opinion. A meme consists of image and text to convey the attitude, humour and criticism on various topics, which have more impact than plain text (Trelles-Villanueva et al., 2025). In political contexts, memes were used as powerful tool for both supporting and ridicule political figures and parties (AlAfnan, 2025), (Anderau and Barbarrusa, 2024). This makes the automated analysis of political memes an important challenge in computational social science.

Significant progress has been made in meme analysis for high-resource languages such as English (Hossain et al., 2022), (Veeramani et al., 2024). Various image and text encoding techniques were employed for analysing memes including GPT-4 to generate contextual information by combining the image caption and the text in the meme, Bootstrapping Language-Image Pre-training (BLIP) to obtain the text describing the meme image with U-net Encapsulated Transformer (Ignacio et al., 2024), and Dual Contrastive Learning (DualCL) approach (Veeramani et al., 2024). However, low-resource Dravidian languages largely underexplored (Vivek et al., 2025). These languages present unique challenges such as the complex scripts (Chakravarthi et al., 2020), usage of culturally embedded political references to express sarcasm and irony (Ponnusamy et al., 2026). Generally, memes derive meaning from both visual and textual elements. A single image may convey different political interpretations depending on the supporting caption. Therefore, multimodal approaches, which can jointly reason over visual and textual content are required for addressing these challenges (Premjith et al., 2022), (Chakravarthi et al., 2024). In addition, the hierarchical reasoning, where the identification of stance is done first and determine the specific political target has not been systematically benchmarked for

Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 101–113
July 3, 2026 ©2026 Association for Computational Linguistics

Tamil and Malayalam.

To address this gap, the Multi-Level Political Meme Classification shared task was organized the shared task on political meme classification at DravidianLangTech at ACL 2026. The task involves a hierarchical two-level classification, where the the first level objective is to classify whether a meme expresses Support/Praise or Troll/Oppose toward a political entity, and the second level objective is to determine whether the meme refers to an individual political figure or a political party/group, conditioned on the stance prediction. This paper presents the overview of the DravidianLangTech shared task on Multi-Level Political Meme Classification, hosted at ACL 2026.

A total of 19 teams participated in the shared task and submitted systems that used various vision models, text encoders, and fusion strategies. The participating teams used a variety of approaches including Late fusion of transformer-based text encoders (XLM-R, IndicBERT, MuRIL, mBERT) with vision encoders (Contrastive Language–Image Pretraining (CLIP), Residual Network (ResNet), ConvNeXt (Convolutional Next Network)), attention-based and gated cross-modal fusion, hierarchical and cascaded classification architectures, OCR-based text extraction pipelines, vision-language models fine-tuned with instruction prompts, and image-only systems leveraging face detection and political symbol matching.

The remainder of the paper is organized as follows. Sections 2 and 3 describe the shared task formulation and dataset construction process. Section 4 summarizes the methodologies adopted by the participating teams. Section 5 presents the experimental results and discussion. Finally, the conclusion, limitations, and future work are discussed in Sections 6, 7, 8.

2 Task Description

The goal of this task is to classify political memes at two levels using both the image and the text present in the meme. Participants are required to build systems that can identify the opinion expressed in a meme and the political entity it refers to.

Level 1 – Stance Classification. At this level, the system must decide whether a meme shows a positive or negative attitude toward a political person, party. This is a binary classification task with two labels: *Support/Praise* and *Troll/Oppose*.

Level 2 – Target Identification. After identifying the stance, the system must determine the specific target of the meme. The target can be either an individual political figure or a political party or group. This level provides more detailed information about who or what the meme is directed at and depends on the prediction made at Level 1.

Together, these two levels form a hierarchical classification structure, as shown in Table 1.

Level 1	Level 2
Support/Praise	Support for an individual person Support for a political party
Troll/Oppose	Troll against an individual person Troll against a political party

Table 1: Representative label combinations in the shared task.

This task involves several challenges. The meaning of a meme often comes from the combination of the image and the text, and using only one of them is usually not sufficient. In Tamil and Malayalam political memes (Fig 2), sarcasm, cultural context, and indirect expressions are common, which makes interpretation difficult. In addition, the dataset contains more trolling memes than supportive ones, leading to class imbalance. Finally, the same image can convey different meanings when paired with different text, so both image and text must be analyzed together. The shared task was hosted on the Codabench platform. Participants were provided with labelled training data and unlabelled test data for both Tamil and Malayalam. Systems were evaluated using macro-averaged F1 score across all label combinations.

3 Dataset Description

Language	Train	Test	Total
Tamil	802	201	1003
Malayalam	500	100	600

Table 2: Multimodal Meme Dataset Statistics

The dataset comprises two South Indian language corpora of political memes. The Tamil dataset consists of 1003 memes sourced from publicly available platforms, divided into 802 training instances and 201 test instances. Similarly, the Malayalam dataset consists of 600 political memes collected from publicly available sources, including Facebook, Instagram, and web searches via Google, divided into 500 training instances and

100 test instances to support supervised learning and evaluation. Fig 1 illustrates the steps involved in data curation, from collection through annotation to the final dataset.

3.1 Data Collection and Pre Processing

The dataset consists of political memes collected from Facebook, Instagram, and Google through automated web data extraction using the BeautifulSoup4 Python library, in combination with official platform APIs. Memes were identified through political hashtags and keywords. The collected images were cleaned before annotation, wherein duplicate images were removed and memes that were too unclear to interpret were discarded. To ensure privacy, user details, if any, were anonymized, and all images were converted to JPG format.

3.2 Data Annotation

Inter-annotator agreement (IAA) was measured to assess the consistency among annotators for both Tamil and Malayalam subsets. Two standard metrics were employed: Krippendorff's alpha and Fleiss' kappa.

Krippendorff's Alpha measures agreement among any number of annotators on any type of data, accounting for chance agreement as in Eq. (1):

$$\alpha = 1 - \frac{D_o}{D_e} \tag{1}$$

where D_o is the observed disagreement and D_e is the expected disagreement by chance.

Fleiss' Kappa measures agreement among a fixed number of annotators assigning items to categorical labels, corrected for chance as in Eq. (2):

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e} \tag{2}$$

where $\bar{P}$ is the mean observed agreement and $\bar{P}_e$ is the mean expected agreement by chance.

Both metrics were applied across two hierarchical annotation levels. For Tamil, agreement at Level 1 was high ($\alpha = 0.876$, $\kappa = 0.876$), indicating strong consistency in identifying the general stance of memes. At Level 2, agreement was moderate ($\alpha = 0.606$, $\kappa = 0.606$), reflecting the greater difficulty in agreeing on more fine-grained political categories.

For Malayalam, Level 1 agreement was similarly strong ($\alpha = 0.817$), showing that annotators were largely consistent in their general assessments. Level 2 achieved perfect agreement ($\alpha = 1.000$), meaning annotators were in complete agreement when identifying specific political intent.

Language	Level	Krippendorff's α	Fleiss' κ
Tamil	Level 1	0.876	0.876
	Level 2	0.606	0.606
Malayalam	Level 1	0.817	–
	Level 2	1.000	–

Table 3: Inter-annotator agreement scores for Tamil and Malayalam.

Overall, both languages show reliable annotation quality, with particularly strong agreement at the broader classification level across both subsets.

3.2.1 Annotator Demographics

The annotation was carried out by two separate teams for Tamil and Malayalam. For Tamil, three postgraduate students aged between 20 and 25, comprising both male and female annotators, completed the annotation. For Malayalam, the annotation was divided across two levels: Level 1 was handled by four annotators (two male, two female) aged 21 to 27, while Level 2 was completed by two male annotators aged 27 to 37. All annotators were proficient in English along with their respective languages. Table 4 provides the combined demographic details.

Annotation guidelines. Annotators were instructed to jointly interpret both the textual and visual components of each meme before assigning labels. For Level 1 (stance classification), memes were categorized as *Support/Praise* or *Troll/Oppose* based on the overall political stance expressed toward a political entity such as a politician, party, or group. For Level 2 (target identification), annotators identified the primary political target of the meme, including political leaders, political parties, organizations, or communities.Annotators were additionally instructed to consider sarcasm, symbolic imagery, implicit political messaging, and culturally grounded references while labeling memes. Ambiguous or borderline cases were resolved through discussion with an additional annotator to improve annotation consistency.

4 Participant Methodology

A total of 19 teams submitted systems for the shared task covering Tamil and Malayalam political meme classification. Most participants built

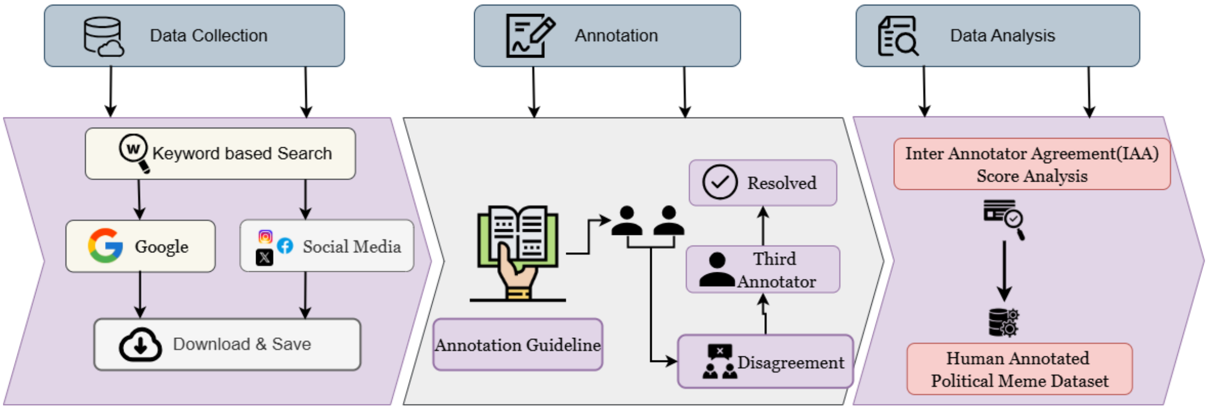

Figure 1: Steps Involved in Data Curation.

(a) Level 1: Support/Praise
Level 2: Support for Individual Person (Malayalam)

(b) Level 1: Troll/Oppose
Level 2: Troll against Individual Person (Malayalam)

(c) Level 1: Support/Praise
Level 2: Support for Party (Tamil)

(d) Level 1: Troll/Oppose
Level 2: Troll against Individual Person (Tamil)

Figure 2: Example memes with multi-level classification labels.

Language	Level	Gender	Age	Qualification
Tamil	1, 2	Female	20–25	PG
		Female	20–25	PG
		Male	20–25	PG
Malayalam	1	Female	21–27	UG
		Female	21–27	UG
		Male	21–27	UG
		Male	21–27	PG
	2	Male	27–37	PG
		Male	27–37	PhD

Table 4: Demographic profile of the annotation teams for Tamil and Malayalam (UG: Undergraduate; PG: Postgraduate).

systems that process both the text and image content of memes, which reflects the nature of the task where meaning often arises from the combination of visual and textual elements. Table 5 shows a high level overview of participant approaches. This section summarises the main approaches taken by participants.

4.1 Combining Text and Image Features

Most systems processed text and image content separately and combined the resulting representations before classification. The most common approach was to concatenate the text and image feature vectors and pass the combined representation through classification layers. Codecrafters, CUET-2567 (Dutta et al., 2026), CUET-DoubleA, Memeocracy, SJK, SSN-CSE-DiNa, susmitha, nitc-hsr, and RMS all followed this design due to its simplicity and consistent performance.

Some teams went beyond simple concatenation. CUETEuropa used an attention mechanism where the text representation guided which parts of the image representation to focus on, followed by a gating step that balances the contribution of each modality. Meme_Alchemists fed the stance prediction from Level 1 as an additional input to the Level 2 classifier so that the target prediction is informed by the predicted stance. MindBridge applied label constraints during training and inference to prevent the model from producing outputs that are inconsistent across the two levels. HAG Signals, IndiLangTech, Semantica, and Synapse (K P et al., 2026) used a single pre-trained vision-language model to process both text and image together without a separate combination step.

4.2 Text Representations

XLM-RoBERTa was the most frequently used model for encoding meme text. Codecrafters, Memeocracy, SSN-CSE-DiNa, susmitha, and CUETEuropa (for Malayalam) used it for its coverage of multiple languages including Tamil and Malayalam. IndicBERT, which was trained on a large collection of Indian language text, was used by CUET-DoubleA, CUETEuropa (for Tamil), SJK, and Team_One. MuRIL was used by Meme_Alchemists, and mBERT was used by CUET-2567 (Dutta et al., 2026) and CUET-DoubleA.

Semantica fine-tuned a seven-billion-parameter vision-language model using a low-rank adaptation technique with instruction-style prompts for both languages. Synapse (K P et al., 2026) applied a similar adaptation method to a two-billion-parameter model and deliberately gave more weight to the text content than the image during training, as Tamil political memes often carry their primary meaning through language. IndiLangTech built a two-stage system using a four-billion-parameter instruction-tuned model, where the first stage predicts the stance of a meme and its output is passed to the second stage to identify the target, allowing each stage to focus on a well-defined sub-task.

4.3 Image Representations

CLIP was the most widely used model for extracting image features. Codecrafters, MindBridge, nitc-hsr, Meme_Alchemists, susmitha, and Cyberpunk (Abir, 2026) used CLIP either alone or alongside other encoders. Memeocracy used ConvNeXt-Tiny, Team_One and SSN-CSE-DiNa used ResNet-18, SJK used VGG16, DLRG used EfficientNet-B3, and CUETEuropa combined CLIP, ResNet, and EfficientNet.

Cyberpunk (Abir, 2026) designed a system that does not use any text extraction and relies entirely on image content. It used CLIP image embeddings combined with features derived from detected faces and the visual similarity of party logos and symbols collected from public sources. This approach is not affected by the quality of text in the image, which can be unreliable in political memes that use stylised or low-resolution fonts.

4.4 Text Extraction from Meme Images

CUET-2567, CUET-DoubleA, CUETEuropa, Meme_Alchemists, MindBridge, Memeocracy,

Team	Lang.		Text Model	Vision Model	Fusion Strategy	PT	IH	OCR	MT	HC
	Ta	Ml								
Codecrafters	✓	✓	XLM-RoBERTa	CLIP / ViT	Multi-modal late fusion	✓	✗	✗	✗	✗
CUET-2567 (Dutta et al., 2026)	✓	✓	mBERT	ViT (patch16-224)	Late fusion (concat.)	✓	✗	✓	✗	✗
CUET-DoubleA	✓	✓	mBERT, IndicBERT, RoBERTa	ViT, ResNet, CLIP	Late fusion (concat.)	✓	✓	✓	✗	✓
CUETEuropa	✓	✓	IndicBERT, Qwen / XLM-R	CLIP, Qwen VL / ResNet, EfficientNet	Cross-attn. + gating	✓	✓	✓	✓	✓
Cyberpunk	✓	✓	— (OCR-free)	OpenCLIP	Image-only (CLIP+face/logo)	✓	✓	✗	✗	✓
DLRG	✓	✗	—	EfficientNet-B3	Image-only (CNN)	✓	✗	✗	✗	✗
HAG Signals	✓	✓	VLM (zero-shot)	VLM (zero-shot)	Zero-shot VLM	✓	✗	✗	✗	✓
IndiLangTech	✓	✓	Gemma 3 4B (Agent-1,2)	Gemma 3 4B (multimodal)	Agentic hierarchical (instruct.)	✓	✗	✗	✗	✓
Meme_Alchemists	✓	✓	MuRIL	CLIP ViT-B/32	Late fusion + L1→L2 cond.	✓	✓	✓	✓	✓
Memeocracy	✓	✓	XLM-RoBERTa	ConvNeXt-Tiny	Late fusion (concat.)	✓	✓	✓	✓	✗
MindBridge	✓	✓	CLIP (text enc.)	CLIP (vision enc.)	CLIP + hierarchical masking	✓	✓	✓	✓	✓
nitc-hsr	✓	✓	CLIP (text enc.)	CLIP (vision enc.)	Multimodal CLIP fusion	✓	✗	✗	✗	✗
RMS	✓	✓	Transformer enc.	Pretrained vision backbone	Concat. + gated fusion	✓	✓	✓	✓	✗
Semantica	✓	✓	Qwen2.5-VL-7B (QLoRA)	Qwen2.5-VL-7B (QLoRA)	VLM fine-tuning (instruct.)	✓	✓	✗	✗	✓
SJK	✓	✓	IndicBERT	VGG16	Late fusion (concat.)	✓	✗	✓	✓	✗
SSN-CSE-DiNa	✓	✓	XLM-RoBERTa	ResNet-18	Late fusion + MLP	✓	✗	✓	✗	✗
susmitha	✓	✓	XLM-RoBERTa	CLIP ViT-B/32	Late fusion (concat.)	✓	✗	✓	✓	✗
Synapse (K P et al., 2026)	✓	✗	Qwen3-VL-2B (LoRA)	Qwen3-VL-2B (LoRA)	VLM text-primacy PEFT (LoRA)	✓	✗	✗	✗	✗
Team_One	✓	✓	IndicBERT / XLM-R	None / ResNet-18	Text-only (Ta); Late fusion (Ml)	✓	✓	✗	✗	✗

Legend: Ta=Tamil, Ml=Malayalam, PT=Pre-trained Models, IH=Imbalance Handling, OCR=OCR Used, MT=Multi-task Learning, HC=Hierarchical Classification

Table 5: Overview of Participants System

RMS, SJK, SSN-CSE-DiNa, and susmitha extracted visible text from meme images using optical character recognition tools, including EasyOCR, PaddleOCR, and Tesseract. Several of these teams noted that text extraction introduced noise into their pipeline, particularly for Tamil script, because memes often use non-standard fonts and image compression that reduces text clarity. Some teams applied text cleaning steps after extraction to reduce this noise before encoding.

4.5 Handling Class Imbalance

The datasets for both languages contained a much larger number of troll-class examples compared to support-class examples, with the imbalance being particularly severe in the Malayalam dataset. Teams addressed this in different ways. Team_One and Semantica increased the number of training examples for the minority class by repeating samples. Team_One additionally lowered the decision threshold at inference time to produce more support-class predictions. Memeocracy, MindBridge, CUETEuropa, and RMS assigned higher penalties to errors on the minority class during training using weighted loss functions. Semantica also addressed the tendency of the model to misclassify sarcastic troll memes as support by adding explicit instructions in the prompt describing how sarcasm and irony should be interpreted. CUET-DoubleA generated additional Tamil training examples by translating existing samples through other Indian and high-resource languages and back into Tamil.

4.6 Two-Level Classification

Several teams specifically addressed the relationship between the two classification levels in their system design. Level 1 requires the system to predict whether a meme expresses support or opposition, while Level 2 requires identifying whether the target is an individual or a political party. IndiLangTech separated the two levels into distinct fine-tuned stages, with the output of the first stage passed as context to the second. Meme_Alchemists trained a shared model with two output heads and fed Level 1 predictions directly into the Level 2 head. MindBridge applied masking to prevent logically contradictory label combinations. Cyberpunk (Abir, 2026) used a two-stage prediction process specifically for the Malayalam target category to handle the rare intersection class. CUET-DoubleA applied a logical mapping after prediction to ensure consistency between the two levels.

4.7 Language Coverage

Seventeen of the 19 teams submitted results for both Tamil and Malayalam. DLRG and Synapse (K P et al., 2026) participated only in the Tamil subtask. Most teams trained separate models for each language to account for differences in dataset size and class distribution, while keeping the overall system design the same across both languages.

5 Results and Discussion

Table 6 and Table 7 present the official results for the Malayalam and Tamil sub-tasks respectively. Systems are ranked by the average macro-F1 score across both levels. Results of all runs of all systems are in Appendix 10, 11

The experimental results across both Tamil and Malayalam sub-tasks reveal clear trends in system performance, model design choices, and the challenges inherent to multi-level political meme classification. Across all submissions, performance on Level 1, which focuses on stance detection, is consistently higher than on Level 2, which requires identifying the political target. This difference reflects the increased complexity of target identification, where cues are often implicit, culturally grounded, or conveyed through sarcasm and visual symbolism rather than explicit textual mentions.

Systems generally achieve macro-F1 scores above 0.90 for Level 1 in both languages, indicating that current multimodal representations are effective at capturing overall sentiment or stance. In contrast, Level 2 performance is notably lower, particularly for Malayalam. This gap can be attributed to several factors, including stronger class imbalance, subtler visual metaphors, and a heavier reliance on contextual political knowledge in Malayalam memes. Tamil systems tend to perform better on Level 2, likely due to a larger dataset size and more explicit linguistic or visual markers that signal the intended target.

The strongest systems demonstrate that architectural design and task decomposition play a critical role in overall performance. Cyberpunk, which ranks highest for Malayalam, achieves competitive results without relying on OCR or textual encoders, instead focusing entirely on visual information such as faces and political symbols. This suggests that for languages where OCR quality is unreliable due to stylised fonts or low-resolution text, image-centric approaches can offer a robust alternative. In contrast, the best-performing Tamil systems, particularly IndiLangTech (Kumar et al., 2026) and CUETEuropa, rely heavily on textual and cross-modal interactions, highlighting the language-specific nature of effective feature extraction strategies.

Several teams (Appendix Table 7, 9) explicitly modeled the hierarchical relationship between Level 1 and Level 2 labels, either by cascading predictions or by constraining the output space. These approaches generally lead to more consistent predictions across levels and reduce logically invalid label combinations. However, improvements in consistency do not always translate into large gains in macro-F1, especially when errors at Level 1 propagate to Level 2 in cascaded systems. This trade-off indicates that while hierarchical modeling is beneficial, it must be carefully designed to avoid amplifying early-stage errors.

A common source of error across many systems arises from the use of OCR for text extraction. Although OCR enables access to valuable linguistic information, especially in Tamil, it often introduces noise due to non-standard fonts, image compression, and creative typography common in memes. Several teams reported that noisy OCR output adversely affected downstream text encoders unless extensive cleaning was applied. This observation explains why some OCR-free or image-dominant systems remain competitive despite ignoring textual content altogether.

Class imbalance presents another major challenge, particularly for Level 2 categories in Malayalam, where troll-related classes dominate the dataset. While weighted loss functions, oversampling, and threshold adjustment help mitigate this issue, many systems still exhibit reduced recall for minority classes. This suggests that data-level solutions alone are insufficient and that more targeted modeling strategies are needed to handle rare but semantically important categories.

An interesting pattern emerges when comparing fusion strategies. Simple late fusion through feature concatenation remains a strong baseline and performs competitively against more complex attention-based architectures. This indicates that representation quality and task-specific training often outweigh the benefits of sophisticated fusion mechanisms, especially under limited data conditions. At the same time, systems using attention or instruction-tuned vision–language models show advantages in handling nuanced cases, particularly for Level 2 target identification in Tamil. Eventually, the results highlight that effective political meme classification requires balancing multimodal integration, hierarchical reasoning, and robustness to noisy inputs. While current systems demonstrate strong performance for stance detection, the persistent difficulty of target identification underscores the need for future research on culturally informed representations, sarcasm-aware modeling, and improved annotation of implicit political references in low-resource languages.

6 Conclusion

This shared task introduced a benchmark for hierarchical political meme classification in low-resource Dravidian languages, namely Tamil and

Team	Level 1				Level 2				Avg-F1	Rank
	ACC	P	R	F1	ACC	P	R	F1		
Cyberpunk (Abir, 2026)	0.9600	0.9698	0.9600	0.9638	0.6300	0.6367	0.6300	0.6222	**0.7930**	1
CUET-2567	0.9600	0.9216	0.9600	0.9404	0.6300	0.6124	0.5900	0.5148	0.7276	2
Semantica	0.8400	0.9680	0.8400	0.8861	0.5300	0.6253	0.5300	0.5494	0.7178	3
IndiLangTech (Kumar et al., 2026)	0.9400	0.9208	0.9400	0.9303	0.4900	0.3933	0.4900	0.4302	0.6802	4
Synapse	0.9200	0.9200	0.9200	0.9200	0.5000	0.4442	0.5000	0.4256	0.6728	5
Meme_Alchemists	0.9600	0.9216	0.9600	0.9404	0.5000	0.4675	0.5000	0.4048	0.6726	6
MindBridge	0.9600	0.9216	0.9600	0.9404	0.4700	0.3592	0.4700	0.4006	0.6705	7
DLRG	0.9200	0.9200	0.9200	0.9404	0.4500	0.3921	0.4500	0.4095	0.6647	8
CUETEuropa_Rathijit	0.9600	0.9216	0.9600	0.9404	0.4900	0.3740	0.4900	0.3819	0.6612	9
CUETEuropa_Joy	0.9600	0.9216	0.9600	0.9404	0.4900	0.3740	0.4900	0.3819	0.6612	9
SJK	0.9600	0.9216	0.9600	0.9404	0.4600	0.3439	0.4600	0.3742	0.6573	10
Memeocracy	0.9600	0.9506	0.9600	0.9535	0.3800	0.4470	0.3800	0.3545	0.6540	11
Codecrafters	0.9600	0.9216	0.9600	0.9404	0.3500	0.3055	0.3500	0.3115	0.6260	12
Team_One	0.8700	0.9313	0.8700	0.8979	–	–	–	–	0.4490	13

Table 6: Results for the Malayalam sub-task. Systems ranked by average macro-F1 across Level 1 and Level 2. ACC=Accuracy, P=Precision, R=Recall. Team_One did not submit Level 2 predictions for Malayalam.

Team	Level 1				Level 2				Avg-F1	Rank
	ACC	P	R	F1	ACC	P	R	F1		
IndiLangTech (Kumar et al., 2026)	0.9204	0.9147	0.9204	0.9159	0.6915	0.6930	0.6915	0.6898	**0.8029**	1
CUETEuropa_Rathijit	0.9204	0.9161	0.9204	0.9175	0.6716	0.6571	0.6716	0.6619	0.7897	2
CUETEuropa_Joy	0.9204	0.9161	0.9204	0.9175	0.6716	0.6571	0.6716	0.6619	0.7897	2
DLRG	0.9055	0.8986	0.9055	0.8883	0.7015	0.6740	0.7015	0.6842	0.7862	3
Semantica	0.8706	0.8979	0.8706	0.8805	0.6667	0.6936	0.6667	0.6785	0.7795	4
CUET-2567	0.9005	0.8921	0.9005	0.8806	0.6965	0.6652	0.6965	0.6556	0.7681	5
Cyberpunk (Abir, 2026)	0.9254	0.9295	0.9254	0.9271	0.5771	0.6558	0.5771	0.6060	0.7666	6
MindBridge	0.8856	0.8802	0.8856	0.8826	0.6965	0.6408	0.6965	0.6471	0.7648	7
SJK	0.8955	0.8822	0.8955	0.8765	0.6766	0.6521	0.6766	0.6487	0.7626	8
CUET-DoubleA	0.9055	0.9147	0.9055	0.8807	0.6617	0.6883	0.6617	0.6403	0.7605	9
Memeocracy	0.9055	0.8957	0.9055	0.8968	0.6020	0.6582	0.6020	0.6202	0.7585	10
RMS	0.8756	0.8554	0.8756	0.8608	0.6318	0.6151	0.6318	0.6156	0.7382	11
Meme_Alchemists	0.8806	0.8660	0.8806	0.8403	0.6766	0.5977	0.6766	0.5905	0.7154	12
Susmitha	0.8706	0.7580	0.8706	0.8104	0.6816	0.4646	0.6816	0.5525	0.6814	13
Synapse	0.7512	0.8414	0.7512	0.7840	0.4428	0.6117	0.4428	0.4885	0.6362	14
Codecrafters	0.8706	0.7580	0.8706	0.8104	0.1841	0.0339	0.1841	0.0572	0.4338	15

Table 7: Results for the Tamil sub-task. Systems ranked by average macro-F1 across Level 1 and Level 2. ACC=Accuracy, P=Precision, R=Recall.

Malayalam. Political memes present a challenging multimodal problem as they combine textual content, visual cues, and socio-political context to convey stance, criticism, sarcasm, and ideological positioning. To address this challenge, we proposed a two-level hierarchical framework where systems first predict the stance of the meme (Support/Praise vs. Troll/Oppose) and subsequently identify the political target (Individual vs. Party/Group). A total of 19 teams participated in the shared task and explored diverse modeling strategies, including multilingual transformer-based text encoders, visual backbone networks, OCR-enhanced pipelines, cross-modal fusion architectures, and vision–language models. The best-performing systems achieved macro F1-scores of 0.8029 for Tamil and 0.7930 for Malayalam, demonstrating the potential of multimodal approaches for political meme understanding in low-resource languages. The results highlight the importance of combining textual and visual signals for effective meme interpretation. However, performance variations across systems indicate that multimodal political meme analysis remains a challenging problem. We hope that this shared task dataset, benchmark, and system analyses will facilitate future research on multimodal political discourse, sarcasm detection, and stance analysis in low-resource languages.

7 Limitations

This section discusses the limitations of this work. Here, the dataset was collected from social media platforms, which may introduce sampling bias and may not fully represent the diversity of political discourse across different demographic groups and regions. Social media data often reflects platform-specific communication styles and meme cultures. Variations in spelling, grammar, and transliteration may introduce additional challenges for automated systems. Another limitation is that, political memes often rely on subtle multimodal cues such as sarcasm, satire, symbolic imagery, and cultural references. Models may struggle to capture these nuanced signals, particularly when the meaning emerges from the interaction between text and visual content. The relatively limited size of the dataset and the scarcity of large-scale pretrained multimodal resources for Tamil and Malayalam is another limitation of this work.

8 Future Work

Future work may explore larger multilingual political meme datasets covering additional Indic languages and political contexts. Cross-lingual transfer learning between Tamil and Malayalam memes using multilingual vision-language models represents another promising direction. Further improvements may also be achieved through OCR-aware multimodal architectures, sarcasm-aware reasoning mechanisms, and explainable multimodal systems capable of identifying fine-grained political symbolism and contextual relationships. Future work could also explore larger datasets, richer multimodal representations, and improved annotation strategies for capturing complex political discourse.

9 Ethical Statement

This shared task involves political memes that may contain satire, trolling, or potentially offensive language. The data set was collected from publicly accessible sources and user identifiers were anonymized to protect privacy.Although strong IAA was achieved, the interpretation of political memes can be subjective. Models developed for political meme classification should be used responsibly. This benchmark data set is intended solely for academic research in multimodal NLP. Researchers are encouraged to ensure transparency, fairness, and respect for democratic principles when deploying related technologies.

Usage of AI Declaration

We used AI-based tools to support language editing and formatting of this manuscript (e.g., improving clarity, grammar, and LaTeX presentation). All technical content, experimental results, interpretations, and final wording were reviewed and verified by the authors. No AI system was used to generate or alter the dataset, annotations, or official leaderboard results.

Acknowledgements

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

Shahad Abir. 2026. CYBER-PUNK@DravidianLangTech 2026: Multimodal Political Meme Classification using CLIP and Logo Similarity. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Mohammad Awad AlAfnan. 2025. The role of memes in shaping political discourse on social media. *Studies in Media and Communication*, 13(2):1–10.

Glenn Anderau and Daniel Barbarrusa. 2024. The function of memes in political discourse. *Topoi*, 43(5):1529–1546.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Shardul Suryawanshi, Navya Jose, Elizabeth Sherly, and John P McCrae. 2020. Overview of the track on sentiment analysis for dravidian languages in code-mixed text. In *Proceedings of the 12th Annual Meeting of the Forum for Information Retrieval Evaluation*, pages 21–24.

Bharathi Raja Chakravarthi, Saranya Rajiakodi, Rahul Ponnusamy, Kathiravan Pannerselvam, Ramachandran Rajalakshmi, Hariharan LekshmiAmmal, Anshid Kizhakkeparambil, Susminu S Kumar, Bhuvaneswari Sivagnanam, Charmathi Rajkumar, and 1 others. 2024. Overview of shared task on multitask meme classification - unraveling misogynistic and trolls in online memes. In *Proceedings of the Fourth Workshop on Language Technology for Equality, Diversity, Inclusion*, pages 139–144.

Arka Dutta, Anindya Majumder, Adnan Faisal, and Hasan Murad. 2026. CUET-2567@DravidianLangTech-ACL 2026: Multimodal Stance and Target Identification in Dravidian Political Memes. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Eftekhar Hossain, Omar Sharif, and Mohammed Moshiul Hoque. 2022. Memosen: A multimodal dataset for sentiment analysis of memes. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 1542–1554.

Md. Ajwad Hossain. 2026. RMS@DravidianLangTech 2026: Multimodal Gated Fusion for Hierarchical Tamil Political Meme Classification. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Marvin John Ignacio, Thanh Tin Nguyen, Hulin Jin, and Yong-Guk Kim. 2024. Meme analysis using llm-based contextual information and u-net encapsulated transformer. *IEEE Access*, 12:125993–126005.

Suriya K P, Durai Singh K, Gnanasabesan G, Ganesh Sundhar, Hari Krishnan N, and Jyothish Lal G. 2026. SYNAPSE@DravidianLangTech 2026: Multi-Level Political Meme Classification for Tamil and Malayalam. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Saurabh Kumar, Vivekananda G, Sanasam Ranbir Singh, and Sukumar Nandi. 2026. IndiLangTech@DravidianLangTech 2026: Hierarchical Modeling for Multi-Level Political Meme Classification. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Rahul Ponnusamy, Saranya Rajiakodi, Bhuvaneswari Sivagnanam, Anshid Kizhakkeparambil, Dhruv Sharma, Paul Buitelaar, and Bharathi Raja Chakravarthi. 2026. Fuse-md: A culturally-aware multimodal model for detecting misogyny memes. *Natural Language Processing Journal*, page 100197.

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.

Alicia Trelles-Villanueva, Rafael Carrasco-Polaino, and Miguel Ángel Sánchez de la Nieta. 2025. Humor, persuasion, and politics: The impact of memes on x during the 2024 u.s. presidential election campaign. *VISUAL REVIEW. International Visual Culture Review / Revista Internacional de Cultura Visual*, 17(4):313–324.

Junain Uddin, Rahul Datta, Taha Ibne Abdullah, and Hasan Murad. 2026. Semantica@DravidianLangTech 2026: Vision-Language Models for Hierarchical Political Meme Classification in Tamil and Malayalam. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Hariram Veeramani, Surendrabikram Thapa, Mohammad Salman, Mohammad Ali Moni, and Usman Naseem. 2024. Umml: Layout-aware text-image fusion for unified multilingual multimodal learning. In *2024 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 174–182. IEEE.

Karnati Vivek, Sai Mani Garrepalli, PVS Ajay Krishna, Inti Charan Kumar, Ravi Teja Gedela, and Sasibhushara Rao Pappu. 2025. Telmemood: A multimodal dataset for sentiment analysis of telugu memes. In *2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*, pages 893–899. IEEE.

A Appendix

Table 8: Tamil: Classification examples with sample images.

<table>
<tr><th>Image</th><th>Actual (Level 1 → Level 2)</th><th>Team</th><th>Pred Level 1</th><th>Pred Level 2</th></tr>
<tr><td rowspan="11">
</td><td rowspan="11">Troll/Oppose → Against Party</td><td>INDILANGTECH (Kumar et al., 2026) ✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>CUETEUROPA_RATHIJIT ✗</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>CUETEUROPA_JOY ✗</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>DLRG ✗</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>SEMANTICA (Uddin et al., 2026) ✗</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>CUET-2567 ✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>Cyberpunk (Abir, 2026) ✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>MINDBRIDGE ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>SJK ✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>CUET-DOUBLEA ✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>MEMOCRACY✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td rowspan="11">
</td><td rowspan="11">Troll/Oppose → Against Person</td><td>INDILANGTECH (Kumar et al., 2026)✗</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>CUETEUROPA_RATHIJIT ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>CUETEUROPA_JOY ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>DLRG ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>SEMANTICA ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>CUET-2567 ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>Cyberpunk (Abir, 2026) ✓</td><td>Troll/Oppose</td><td>Against Party</td></tr>
<tr><td>MINDBRIDGE ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>SJK ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>CUET-DOUBLEA ✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
<tr><td>MEMOCRACY✓</td><td>Troll/Oppose</td><td>Against Person</td></tr>
</table>

Table 9: Malayalam: Classification examples with sample images from top 6 teams.

Image	Actual (Level 1 → Level 2)	Team	Pred Level 1	Pred Level 2
	Support/Praise → Support for Person	Cyberpunk (Abir, 2026) ✗	Troll/Oppose	Against Person
		CUET-2567 ✗	Troll/Oppose	Against Person
		SEMANTICA ✓	Support/Praise	Support for Person
		INDILANGTECH (Kumar et al., 2026) ✗	Troll/Oppose	Against Person
		SYNAPSE ✗	Troll/Oppose	Against Person
		ALCHEMIST ✗	Troll/Oppose	Against Person
		MINDBRIDGE ✗	Troll/Oppose	Against Person
		DLRG ✗	Troll/Oppose	Against Person
		CUETEUROPA RATHIJIT✗	Troll/Oppose	Against Person
		CUETEUROPA JOY✗	Troll/Oppose	Against Person
		SJK✗	Troll/Oppose	Against Person
	Troll/Oppose → Against Person	Cyberpunk (Abir, 2026) ✓	Troll/Oppose	Against Person
		CUET-2567 ✓	Troll/Oppose	Against Person
		SEMANTICA ✗	Support/Praise	Support for Party
		INDILANGTECH (Kumar et al., 2026) ✗	Support/Praise	Support for Person
		SYNAPSE ✓	Troll/Oppose	Against Person
		ALCHEMIST ✓	Troll/Oppose	Against Person
		MINDBRIDGE ✗	Troll/Oppose	Against Person
		DLRG ✗	Troll/Oppose	Against Person
		CUETEUROPA RATHIJIT✗	Troll/Oppose	Against Party
		CUETEUROPA JOY✗	Troll/Oppose	Against Party
		SJK✗	Troll/Oppose	Against Person

Rank	Team	Run	Avg F1	L1 F1	L2 F1	L1 Acc	L1 Prec	L1 Rec	L2 Acc	L2 Prec	L2 Rec
1	IndiLangTech (Kumar et al., 2026)	Run3	0.8029	0.9159	0.6898	0.9204	0.9147	0.9204	0.6915	0.6930	0.6915
2	CUETEuropa_RATHIJIT	Run2	0.8027	0.9175	0.6879	0.9204	0.9161	0.9204	0.7363	0.6985	0.7363
2	CUETEuropa - JOY	Run2	0.8027	0.9175	0.6879	0.9204	0.9161	0.9204	0.7363	0.6985	0.7363
3	IndiLangTech (Kumar et al., 2026)	Run1	0.8022	0.9159	0.6885	0.9204	0.9147	0.9204	0.7065	0.6754	0.7065
4	CUETEuropa_RATHIJIT	Run3	0.7897	0.9175	0.6619	0.9204	0.9161	0.9204	0.6716	0.6571	0.6716
4	CUETEuropa - JOY	Run3	0.7897	0.9175	0.6619	0.9204	0.9161	0.9204	0.6716	0.6571	0.6716
5	DLRG	Single	0.7862	0.8883	0.6842	0.9055	0.8986	0.9055	0.7015	0.6740	0.7015
6	Semantica	Run1	0.7795	0.8805	0.6785	0.8706	0.8979	0.8706	0.6667	0.6936	0.6667
6	Semantica	Run2	0.7795	0.8805	0.6785	0.8706	0.8979	0.8706	0.6667	0.6936	0.6667
7	CUET-DoubleA	Run2	0.7744	0.8807	0.6680	0.9055	0.9147	0.9055	0.7164	0.7066	0.7164
8	IndiLangTech (Kumar et al., 2026)	Run2	0.7733	0.9033	0.6432	0.9104	0.9022	0.9104	0.6965	0.6363	0.6965
9	CUETEuropa - JOY	Run1	0.7714	0.9028	0.6400	0.9154	0.9106	0.9154	0.6617	0.6510	0.6617
9	CUETEuropa_RATHIJIT	Run1	0.7714	0.9028	0.6400	0.9154	0.9106	0.9154	0.6617	0.6510	0.6617
10	CUET-2567	Single	0.7681	0.8806	0.6556	0.9005	0.8921	0.9005	0.6965	0.6652	0.6965
11	RMS (Hossain, 2026)	Run1	0.7672	0.8763	0.6581	0.8806	0.8731	0.8806	0.6617	0.6580	0.6617
12	Cyberpunk (Abir, 2026)	Run2	0.7666	0.9271	0.6060	0.9254	0.9295	0.9254	0.5771	0.6558	0.5771
13	MindBridge	Single	0.7648	0.8826	0.6471	0.8856	0.8802	0.8856	0.6965	0.6408	0.6965
14	Cyberpunk (Abir, 2026)	Run1	0.7629	0.9007	0.6252	0.8955	0.9088	0.8955	0.6020	0.6667	0.6020
15	SJK	Single	0.7626	0.8765	0.6487	0.8955	0.8822	0.8955	0.6766	0.6521	0.6766
16	CUET-DoubleA	Run3	0.7605	0.8807	0.6403	0.9055	0.9147	0.9055	0.6617	0.6883	0.6617
17	Memocracy	Single	0.7585	0.8968	0.6202	0.9055	0.8957	0.9055	0.6020	0.6582	0.6020
18	CUET-DoubleA	Run1	0.7571	0.8686	0.6455	0.8905	0.8746	0.8905	0.6866	0.6410	0.6866
19	RMS (Hossain, 2026)	Run2	0.7525	0.8884	0.6165	0.8955	0.8854	0.8955	0.6418	0.6059	0.6418
20	RMS (Hossain, 2026)	Run3	0.7382	0.8608	0.6156	0.8756	0.8554	0.8756	0.6318	0.6151	0.6318
21	Alchemists	Single	0.7154	0.8403	0.5905	0.8806	0.8660	0.8806	0.6766	0.5977	0.6766
22	SUSMITHA	Single	0.6815	0.8104	0.5525	0.8706	0.7580	0.8706	0.6816	0.4646	0.6816
23	Synapse (K P et al., 2026)	Single	0.6362	0.7840	0.4885	0.7512	0.8414	0.7512	0.4428	0.6117	0.4428
24	Codecrafters	Single	0.4338	0.8104	0.0572	0.8706	0.7580	0.8706	0.1841	0.0339	0.1841

Table 10: Tamil: All Runs Results Ranked by Average F1 Score

Rank	Team	Run	L1 F1	L2 F1	Avg F1	L1 Acc	L1 Prec	L1 Rec	L2 Acc	L2 Prec	L2 Rec
1	Cyberpunk (Abir, 2026)	Run2	0.9638	0.6222	0.7930	0.9600	0.9698	0.9600	0.6300	0.6367	0.6300
2	CUET-2567	Single	0.9404	0.5148	0.7276	0.9600	0.9216	0.9600	0.5900	0.6124	0.5900
3	Cyberpunk (Abir, 2026)	Run1	0.9248	0.5126	0.7187	0.9200	0.9297	0.9200	0.5300	0.5286	0.5300
4	SEMANTICA	Run2	0.8861	0.5494	0.7177	0.8400	0.9680	0.8400	0.5300	0.6253	0.5300
5	INDILANGTECH (Kumar et al., 2026)	Run2	0.9248	0.4896	0.7072	0.9200	0.9297	0.9200	0.5600	0.4788	0.5600
6	SEMANTICA	Run1	0.9084	0.4978	0.7031	0.8800	0.9387	0.8800	0.4900	0.5127	0.4900
7	INDILANGTECH (Kumar et al., 2026)	Run3	0.9303	0.4302	0.6803	0.9400	0.9208	0.9400	0.4900	0.3933	0.4900
8	INDILANGTECH (Kumar et al., 2026)	Run1	0.9303	0.4243	0.6773	0.9400	0.9208	0.9400	0.5000	0.4003	0.5000
9	SYNAPSE (K P et al., 2026)	Single	0.9200	0.4256	0.6728	0.9200	0.9200	0.9200	0.5000	0.4442	0.5000
10	ALCHEMIST	Single	0.9404	0.4048	0.6726	0.9600	0.9216	0.9600	0.5000	0.4675	0.5000
11	CUETEUROPA_JOY	Run1	0.9196	0.4242	0.6719	0.9100	0.9294	0.9100	0.4200	0.4316	0.4200
11	CUETEUROPA_RATHIJIT	Run1	0.9196	0.4242	0.6719	0.9100	0.9294	0.9100	0.4200	0.4316	0.4200
13	MINDBRIDGE	Single	0.9404	0.4006	0.6705	0.9600	0.9216	0.9600	0.4700	0.3592	0.4700
14	DLRG	Single	0.9200	0.4095	0.6648	0.9200	0.9200	0.9200	0.4500	0.3921	0.4500
15	CUETEUROPA_RATHIJIT	Run3	0.9404	0.3819	0.6612	0.9600	0.9216	0.9600	0.4900	0.3740	0.4900
15	CUETEUROPA_JOY	Run3	0.9404	0.3819	0.6612	0.9600	0.9216	0.9600	0.4900	0.3740	0.4900
17	SJK	Single	0.9404	0.3742	0.6573	0.9600	0.9216	0.9600	0.4600	0.3439	0.4600
18	MEMOCRACY	Single	0.9535	0.3545	0.6540	0.9600	0.9506	0.9600	0.3800	0.4470	0.3800
19	CODECRAFTERS	Single	0.9404	0.3115	0.6260	0.9600	0.9216	0.9600	0.3500	0.3055	0.3500
20	CUETEUROPA_RATHIJIT	Run2	0.9303	0.2035	0.5669	0.9400	0.9208	0.9400	0.3300	0.4239	0.3300
20	CUETEUROPA_JOY	Run2	0.9303	0.2035	0.5669	0.9400	0.9208	0.9400	0.3300	0.4239	0.3300
21	Team_One	Single	0.8979	-	0.4490	0.8700	0.9313	0.8700	-	-	-

Table 11: Malayalam: All Runs Results Ranked by Average F1 Score