MMLow 2026

Fifth Workshop on Multimodal Machine Learning in Low-resource Languages, AACL-IJCNLP 2026

Proceedings of the Conference

November 9, 2026

The MMLow organizers gratefully acknowledge the support from the following sponsors.







©2026 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-445-3

Introduction

Multimodal machine learning has advanced significantly in recent years, driven by rapid progress in computer vision, speech processing, and natural language processing. However, much of this progress has focused on high-resource languages, particularly English. The availability of labelled datasets, benchmarks, and computational tools remains limited for low-resource languages, especially those spoken across Asia, Africa, Europe, and South America. This scarcity presents significant challenges for developing robust multimodal systems capable of understanding and processing content in regional and underrepresented languages. The Fifth Workshop on Multimodal Machine Learning in Low-resource Languages (MMLow 2026), co-located with AACL-IJCNLP 2026 in Hengqin, China, continues a workshop series initiated in 2022. The workshop provides a forum for researchers and practitioners to exchange ideas and present advances in multimodal data fusion, representation learning, and applications involving regional and low-resource languages. It aims to address the persistent scarcity of datasets, tools, and evaluation resources while fostering research directions that support future developments and real-world applications. The workshop welcomed submissions on a broad range of topics, including multimodal sentiment and emotion analysis in regional languages; hate speech, misogyny, offensive content, and troll or political meme detection; multimodal data fusion and representation learning; benchmark datasets and evaluation methodologies for low-resource multimodal tasks; fake news and misinformation detection in regional languages; speech-text and vision-language modelling for low-resource languages; and data collection and annotation methods for safer social media and content moderation. MMLow 2026 also featured two shared tasks: Multi-Level Political Meme Classification, which focused on classifying political memes at multiple levels, and Identifying Misogynistic Videos in Dravidian Languages, which addressed the detection of misogynistic content in videos featuring Dravidian languages. These shared tasks encouraged systematic evaluation and provided opportunities for researchers to explore practical solutions to challenging multimodal language understanding problems. The workshop received 18 submissions, of which 11 papers were accepted following a rigorous review process, corresponding to an acceptance rate of 61.1%. We sincerely thank all authors and shared-task participants for their valuable contributions. We are also grateful to the programme committee members for their timely, and constructive reviews, and to the AACL-IJCNLP 2026 workshop chairs for their guidance and support.

Organizing Committee

Workshop Organizers

Rahul Ponnusamy, University of Galway, Ireland
Prasanna Kumar Kumaresan, University of Galway, Ireland
Saranya Rajiakodi, Central University of Tamil Nadu, Thiruvarur, India
Premjith B, Amrita Vishwa Vidyapeetham, Coimbatore, India
Shunmuga Priya MC, University of Galway, Ireland
Abirami Murugappan, Anna University, Tamil Nadu, India
Jyothish Lal G, Amrita Vishwa Vidyapeetham, Coimbatore, India
Sowmya V, Amrita Vishwa Vidyapeetham, Coimbatore, India
Bharathi Raja Chakravarthi, University of Galway, Ireland

Program Committee

Program Committee

Bhuvaneshwari M, Kumaraguru College of Technology, India
Jairam R, Amrita Vishwa Vidyapeetham, India
Thomas Mandl, Universität Hildesheim
Rajesh Balaji, Costco Wholesale Corporation
Divya Sasidharan, Amrita Vishwa Vidyapeetham, India
Hariharan R L, Vellore Institute of Technology
Shreyash Rawat, Facebook
Anisha Ramakrishna Yarlapati, Adobe Systems
Priyanka C, Amrita Vishwa Vidyapeetham, India
Suryaa Veerabathiran Seran, Adobe Systems
Manjula R, SRM Institute of Science and Technology
Abhishek Anilkumar, Amrita Vishwa Vidyapeetham, India
Archudha Arjunasamy, Amrita Vishwa Vidyapeetham, India
Yokesh Dhanabal, Kumaraguru College of Technology
Anita R, SRM Institute of Science and Technology
Ipsita Mohanty, Amazon
Kathiresan Jayabalan, Sathyabama Institute of Science and Technology
Krishnakumari K, University College of Engineering, Panruti
Ravindrakumar M. Purohit, Dhirubhai Ambani Institute of Information and Communication Technology
Shalmiya Paulraj, Kumaraguru College of Technology, India
Naman Ahuja, Facebook
Rizwana Kallooravi Thandil, Sullamussalam Science College Areekode
Kamal Kishore Chigullarevu Thirumala, Stanford University
Rathipriya Easwaramoorthy, Amrita Vishwa Vidyapeetham, India
Divyabharathi R, Amrita Vishwa Vidyapeetham, India
Bagavathi C, Amrita Vishwa Vidyapeetham, India

Table of Contents

Script Fragmentation and Format: What Drives the English-Bengali Performance Gap in Open LLMs?
Shimanto Bhowmik, Tawsif Tashwar Dipto, Md Sazzad Islam, Sheryl Hsu and Tahsin Reasat . . 1

A Foundational Study of Multimodal Speech–Text Alignment in Malayalam
Meharuniza Nazeem, Navaneeth S, Adithyakiran A S, Rajeev R R and Ashbi Biju 18

A Retrieval-Augmented Generation Approach for Enhancing Pedagogical Support in Education
Ramkumar N and Rajendran S 28

One Head, Three Transcripts: A Controlled Oracle-to-ASR Audit of Bangla Speech–Text Sentiment Analysis
Md Raihan, Nilam Rajak, Kuntal Ghosh and Sathyanarayanan Dhorali 40

MOEZHI-GURU: Multi-Perspective Misogynistic Meme Detection via Multimodal Mixture of Experts
Sekar M, Pranava Mohan and A. Santhanavijayan 55

Axiom@MMLow 2026: Cascaded Prompting with QLoRA-Tuned Qwen2-VL for Hierarchical Tamil Political Meme Classification
Md. Ajwad Hossain 64

TVKS@MMLow 2026: Factorized Stance × Target Classification of Tamil Political Memes with an Axis-Wise Multimodal Ensemble
Kevin selva K and Thilagavathy R 70

Projukti Kavalan@MMLow 2026: Weighted Visual Frame and Cross-modal Fusion for Misogynistic Video Detection in Dravidian Language
Palaniraj I and Dwaipayan Dutta 76

KEC@MMLow 2026: Multimodal Fusion for Hierarchical Tamil Political Meme Classification
Asvitha RM and Amirthamozhi VS 86

CPSLab2@MMLow 2026: A Hierarchical Multimodal Framework with Stratified K-Fold Training for Political Meme Classification
Anika Srivastava, Vishal Kumar, Harshita Sharma, Nitisha Aggarwal, Geetika Jain Saxena, Amit Pundir and Sanjeev Singh 91

CPS Lab 1@MMLow 2026: Hierarchical Multimodal Classification of Tamil Political Memes via Logit-Conditioned Stance-Target Prediction
Vishal Kumar, Anika Srivastava, Harshita Sharma, Nitisha Aggarwal, Geetika Jain Saxena, Amit Pundir and Sanjeev Singh 97

Program

Monday, November 9, 2026

09:00 - 09:15 *Opening Remarks*

09:15 - 10:30 *Session 1: Oral*

Script Fragmentation and Format: What Drives the English-Bengali Performance Gap in Open LLMs?
Shimanto Bhowmik, Tawsif Tashwar Dipto, Md Sazzad Islam, Sheryl Hsu and Tahsin Reasat

A Foundational Study of Multimodal Speech–Text Alignment in Malayalam
Meharuniza Nazeem, Navaneeth S, Adithyakiran A S, Rajeev R R and Ashbi Biju

One Head, Three Transcripts: A Controlled Oracle-to-ASR Audit of Bangla Speech–Text Sentiment Analysis
Md Raihan, Nilam Rajak, Kuntal Ghosh and Sathyanarayanan Dhorali

A Retrieval-Augmented Generation Approach for Enhancing Pedagogical Support in Education
Ramkumar N and Rajendran S

MOEZHI-GURU: Multi-Perspective Misogynistic Meme Detection via Multimodal Mixture of Experts
Sekar M, Pranava Mohan and A. Santhanavijayan

10:30 - 11:00 *Coffee Break*

11:00 - 12:00 *Session 2: Poster*

Axiom@MMLow 2026: Cascaded Prompting with QLoRA-Tuned Qwen2-VL for Hierarchical Tamil Political Meme Classification
Md. Ajwad Hossain

TVKS@MMLow 2026: Factorized Stance $\times$ Target Classification of Tamil Political Memes with an Axis-Wise Multimodal Ensemble
Kevin selva K and Thilagavathy R

Projukti Kavalan@MMLow 2026: Weighted Visual Frame and Cross-modal Fusion for Misogynistic Video Detection in Dravidian Language
Palaniraj I and Dwaipayan Dutta

Monday, November 9, 2026 (continued)

KEC@MMLow 2026: Multimodal Fusion for Hierarchical Tamil Political Meme Classification
Asvitha RM and Amirthamozhi VS

CPSLab2@MMLow 2026: A Hierarchical Multimodal Framework with Stratified K-Fold Training for Political Meme Classification
Anika Srivastava, Vishal Kumar, Harshita Sharma, Nitisha Aggarwal, Geetika Jain Saxena, Amit Pundir and Sanjeev Singh

CPS Lab 1@MMLow 2026: Hierarchical Multimodal Classification of Tamil Political Memes via Logit-Conditioned Stance-Target Prediction
Vishal Kumar, Anika Srivastava, Harshita Sharma, Nitisha Aggarwal, Geetika Jain Saxena, Amit Pundir and Sanjeev Singh

12:00 - 12:30 *Closing Remarks*

12:30 - 14:00 *Lunch Break*