

Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages (DravidianLangTech 2024)

St. Julian's, Malta
22 March 2024

ISBN: 978-1-7138-9453-7

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2024) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>A Few-Shot Multi-Accented Speech Classification for Indian Languages using Transformers and LLM's Fine-Tuning Approaches</i>	
Jairam R, Jyothish Lal G and Premjith B	1
<i>Neural Machine Translation for Malayalam Paraphrase Generation</i>	
Christeena Varghese, Sergey Koshelev and Ivan P. Yamshchikov	10
<i>From Dataset to Detection: A Comprehensive Approach to Combating Malayalam Fake News</i>	
Devika K, Hariprasath .S.B, Haripriya B, Vigneshwar E, Premjith B and Bharathi Raja Chakravarthi	16
<i>Social Media Fake News Classification Using Machine Learning Algorithm</i>	
Girma Yohannis Bade, Olga Kolesnikova, Grigori Sidorov and José Luis Oropeza	24
<i>Exploring the impact of noise in low-resource ASR for Tamil</i>	
Vigneshwar Lakshminarayanan and Emily Prud'hommeaux	30
<i>SetFit: A Robust Approach for Offensive Content Detection in Tamil-English Code-Mixed Conversations Using Sentence Transfer Fine-tuning</i>	
Kathiravan Pannerselvam, Saranya Rajiakodi, Sajeetha Thavareesan, Sathiyaraj Thangasamy and Kishore Kumar Ponnusamy	35
<i>Findings of the First Shared Task on Offensive Span Identification from Code-Mixed Kannada-English Comments</i>	
Manikandan Ravikiran, Ratnavel Rajalakshmi, Bharathi Raja Chakravarthi, Anand Kumar Madasamy and Sajeetha Thavareesan	43
<i>Findings of the Shared Task on Hate and Offensive Language Detection in Telugu Codemixed Text (HOLD-Telugu)@DravidianLangTech 2024</i>	
Premjith B, Bharathi Raja Chakravarthi, Prasanna Kumar Kumaresan, Saranya Rajiakodi, Sai Prashanth Karnati, Sai Rishith Reddy Mangamuru and Chandu Janakiram	49
<i>Findings of the Shared Task on Multimodal Social Media Data Analysis in Dravidian Languages (MSMDA-DL)@DravidianLangTech 2024</i>	
Premjith B, Jyothish Lal G, Sowmya V, Bharathi Raja Chakravarthi, K Nandhini, Rajeswari Natarajan, Abirami Murugappan, Bharathi B, Saranya Rajiakodi, Rahul Ponnusamy, Jayanth Mohan and Mekapati Spandana Reddy	56
<i>Overview of Second Shared Task on Sentiment Analysis in Code-mixed Tamil and Tulu</i>	
Lavanya Sambath Kumar, Asha Hegde, Bharathi Raja Chakravarthi, Hosahalli Lakshmaiah Shashirekha, Rajeswari Natarajan, Sajeetha Thavareesan, Ratnasingam Sakuntharaj, Thenmozhi Durairaj, Prasanna Kumar Kumaresan and Charmathi Rajkumar	62
<i>Overview of the Second Shared Task on Fake News Detection in Dravidian Languages: Dravidian-LangTech@EACL 2024</i>	
Malliga Subramanian, Bharathi Raja Chakravarthi, Kogilavani Shanmugavadivel, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani, Premjith B, Vanaja k, Mithunja S, Devika K, Hariprasath S.B, Haripriya B and Vigneshwar E	71
<i>byteSizedLLM@DravidianLangTech 2024: Fake News Detection in Dravidian Languages - Unleashing the Power of Custom Subword Tokenization with Subword2Vec and BiLSTM</i>	
Rohith Gowtham Kodali and Durga Prasad Manukonda	79

<i>Fida @DravidianLangTech 2024: A Novel Approach to Hate Speech Detection Using Distilbert-base-multilingual-cased</i>	
Fida Ullah, Muhammad Tayyab Zamir, Muhammad Arif, M. Ahmad, E Felipe-Riveron and Alexander Gelbukh	85
<i>Selam@DravidianLangTech 2024:Identifying Hate Speech and Offensive Language</i>	
Selam Abitte Kanta, Grigori Sidorov and Alexander Gelbukh	91
<i>Tewodros@DravidianLangTech 2024: Hate Speech Recognition in Telugu Codemixed Text</i>	
Tewodros Achamaleh, Lemlem Eyob Kawo, Ildar Batyrshini and Grigori Sidorov	96
<i>Lidoma@DravidianLangTech 2024: Identifying Hate Speech in Telugu Code-Mixed: A BERT Multilingual</i>	
Muhammad Tayyab Zamir, Moein Shahiki Tash, Zahra Ahani, Alexander Gelbukh and Grigori Sidorov	101
<i>Zavira@DravidianLangTech 2024:Telugu hate speech detection using LSTM</i>	
Z. Ahani, M. Shahiki Tash, M. T. Zamir and I. Gelbukh	107
<i>Tayyab@DravidianLangTech 2024:Detecting Fake News in Malayalam LSTM Approach and Challenges</i>	
M. T. Zamir, M. S. Tash, Z. Ahani, A. Gelbukh and G. Sidorov	113
<i>IIITDWD_SVC@DravidianLangTech-2024: Breaking Language Barriers; Hate Speech Detection in Telugu-English Code-Mixed Text</i>	
Chava Srinivasa Sai, Rangoori Vinay Kumar, Sunil Saumya and Shankar Biradar	119
<i>Beyond Tech@DravidianLangTech2024 : Fake News Detection in Dravidian Languages Using Machine Learning</i>	
Kogilavani Shanmugavadivel, Malliga Subramanian, Sanjai R, Mohammed Sameer B and Motheeswaran K.	124
<i>Code_Makers@DravidianLangTech-EACL 2024 : Sentiment Analysis in Code-Mixed Tamil using Machine Learning Techniques</i>	
Kogilavani Shanmugavadivel, Sowbharanika Janani J S, Navbila K and Malliga Subramanian	129
<i>IIITDWD-zk@DravidianLangTech-2024: Leveraging the Power of Language Models for Hate Speech Detection in Telugu-English Code-Mixed Text</i>	
Zuhair Hasan Shaik, Sai Kartheek Reddy Kasu, Sunil Saumya and Shankar Biradar	134
<i>DLRG-DravidianLangTech@EACL2024 : Combating Hate Speech in Telugu Code-mixed Text on Social Media</i>	
Ratnavel Rajalakshmi, Saptharishree M, Hareesh Teja S, Gabriel Joshua R and Varsini SR ...	140
<i>MIT-KEC-NLP@DravidianLangTech-EACL 2024: Offensive Content Detection in Kannada and Kannada-English Mixed Text Using Deep Learning Techniques</i>	
Kogilavani Shanmugavadivel, Sowbarnigaa K S, Mehal Sakthi M S, Subhadevi K and Malliga Subramanian	146
<i>Transformers@DravidianLangTech-EACL2024: Sentiment Analysis of Code-Mixed Tamil Using RoBERTa</i>	
Kriti Singhal and Jatin Bedi	151
<i>Habesha@DravidianLangTech 2024: Detecting Fake News Detection in Dravidian Languages using Deep Learning</i>	
Mesay Gameda Yigezu, Olga Kolesnikova, Grigori Sidorov and Alexander Gelbukh	156

<i>WordWizards@DravidianLangTech 2024:Fake News Detection in Dravidian Languages using Cross-lingual Sentence Embeddings</i>	
Akshatha Anbalagan, Priyadharshini T, Niranjana A, Shreedevi Seluka Balaji and Durairaj Thenmozhi	162
<i>Sandalphon@DravidianLangTech-EACL2024: Hate and Offensive Language Detection in Telugu Code-mixed Text using Transliteration-Augmentation</i>	
Nafisa Tabassum, Mosabbir Hossain Khan, Shawly Ahsan, Jawad Hossain and Mohammed Moshiul Hoque	167
<i>CUET_Binary_Hackers@DravidianLangTech EACL2024: Fake News Detection in Malayalam Language Leveraging Fine-tuned MuRIL BERT</i>	
Salman Farsi, Asrarul Hoque Eusha, Ariful Islam, Hasan Mesbaul Ali Taher, Jawad Hossain, Shawly Ahsan, Avishek Das and Mohammed Moshiul Hoque	173
<i>Punny_Punctuators@DravidianLangTech-EACL2024: Transformer-based Approach for Detection and Classification of Fake News in Malayalam Social Media Text</i>	
Nafisa Tabassum, Sumaiya Rahman Aodhora, Rowshon Akter, Jawad Hossain, Shawly Ahsan and Mohammed Moshiul Hoque	180
<i>CUET_NLP_GoodFellows@DravidianLangTech EACL2024: A Transformer-Based Approach for Detecting Fake News in Dravidian Languages</i>	
Md Osama, Kawsar Ahmed, Hasan Mesbaul Ali Taher, Jawad Hossain, Shawly Ahsan and Mohammed Moshiul Hoque	187
<i>CUET_Binary_Hackers@DravidianLangTech EACL2024: Hate and Offensive Language Detection in Telugu Code-Mixed Text Using Sentence Similarity BERT</i>	
Salman Farsi, Asrarul Hoque Eusha, Jawad Hossain, Shawly Ahsan, Avishek Das and Mohammed Moshiul Hoque	193
<i>TechWhiz@DravidianLangTech 2024: Fake News Detection Using Deep Learning Models</i>	
Madhumitha M, Kunguma Akshatra M, Tejashri J and Jerin Mahibha C	200
<i>CUET_Binary_Hackers@DravidianLangTech-EACL 2024: Sentiment Analysis using Transformer-Based Models in Code-Mixed and Transliterated Tamil and Tulu</i>	
Asrarul Hoque Eusha, Salman Farsi, Ariful Islam, Jawad Hossain, Shawly Ahsan and Mohammed Moshiul Hoque	205
<i>Binary_Beasts@DravidianLangTech-EACL 2024: Multimodal Abusive Language Detection in Tamil based on Integrated Approach of Machine Learning and Deep Learning Techniques</i>	
Md. Tanvir Rahman, Abu Bakkar Siddique Raihan, Tanzim Rahman, Shawly Ahsan, Jawad Hossain, Avishek Das and Mohammed Moshiul Hoque	212
<i>WordWizards@DravidianLangTech 2024: Sentiment Analysis in Tamil and Tulu using Sentence Embedding</i>	
Shreedevi Seluka Balaji, Akshatha Anbalagan, Priyadharshini T, Niranjana A and Durairaj Thenmozhi	218
<i>CUET_DUO@DravidianLangTech EACL2024: Fake News Classification Using Malayalam-BERT</i>	
Tanzim Rahman, Abu Bakkar Siddique Raihan, Md. Tanvir Rahman, Jawad Hossain, Shawly Ahsan, Avishek Das and Mohammed Moshiul Hoque	223
<i>Wit Hub@DravidianLangTech-2024:Multimodal Social Media Data Analysis in Dravidian Languages using Machine Learning Models</i>	
Anierudh H S, Abhishek R, Ashwin V Sundar, Amrit Krishnan and Bharathi B	229

<i>CUETSentimentSillies@DravidianLangTech-EACL2024: Transformer-based Approach for Sentiment Analysis in Tamil and Tulu Code-Mixed Texts</i>	
Zannatul Fardaush Tripty, Md. Arian Al Nafis, Antu Chowdhury, Jawad Hossain, Shawly Ahsan, Avishek Das and Mohammed Moshiul Hoque	234
<i>Social Media Hate and Offensive Speech Detection Using Machine Learning method</i>	
Girma Yohannis Bade, Olga Kolesnikova, Grigori Sidorov and José Luis Oropeza	240
<i>CUETSentimentSillies@DravidianLangTech EACL2024: Transformer-based Approach for Detecting and Categorizing Fake News in Malayalam Language</i>	
Zannatul Fardaush Tripty, Md. Arian Al Nafis, Antu Chowdhury, Jawad Hossain, Shawly Ahsan and Mohammed Moshiul Hoque	245
<i>MUCS@DravidianLangTech-2024: Role of Learning Approaches in Strengthening Hate-Alert Systems for code-mixed text</i>	
Manavi K K, Sonali k, Gauthamraj k, Kavya G, Asha Hegde and Hosahalli Lakshmaiah Shashirekha	252
<i>MUCS@DravidianLangTech-2024: A Grid Search Approach to Explore Sentiment Analysis in Code-mixed Tamil and Tulu</i>	
Prathvi b, Manavi K K, Subrahmanyapoojary k, Asha Hegde, Kavya G and Hosahalli Lakshmaiah Shashirekha	257
<i>InnovationEngineers@DravidianLangTech-EACL 2024: Sentimental Analysis of YouTube Comments in Tamil by using Machine Learning</i>	
Kogilavani Shanmugavadivel, Malliga Subramanian, Palanimurugan V and Pavul chinnappan D	262
<i>KEC_HAWKS@DravidianLangTech 2024 : Detecting Malayalam Fake News using Machine Learning Models</i>	
Malliga Subramanian, Jayanthjr J R, Muthu Karuppan P, Keerthibala A T and Kogilavani Shanmugavadivel	266