Third Workshop on Speech and Language Technologies for Dravidian Languages (DravidianLangTech 2023)

Varna, Bulgaria
7 September 2023

ISBN: 978-1-7138-9043-0

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© (2023) 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

On the Errors in Code-Mixed Tamil-English Offensive Span Identification
Manikandan Ravikiran and Bharathi Raja Chakravarthi 1

Hate and Offensive Keyword Extraction from CodeMix Malayalam Social Media Text Using Contextual Embedding
Mariya Raphel, Premjith B, Sreelakshmi K and Bharathi Raja Chakravarthi 10

Acoustic Analysis of the Fifth Liquid in Malayalam
Punnoose A K 19

Transformer-based Context Aware Morphological Analyzer for Telugu
Priyanka Dasari, Abhijith Chelpuri, Nagaraju Vuppala, Mounika Marreddy, Parameshwari Krishnamurthy and Radhika Mamidi 25

Improving Reinfocement Learning Agent Training using Text based Guidance: A study using Commands in Dravidian Languages
Nikhil Chowdary Paleti, Sai Aravind Vadlapudi, Sai Aashish Menta, Sai Akshay Menta, Vishnu Vardhan Gorantla V N S L, Janakiram Chandu, Soman K P and Sachin Kumar S 33

Social Media Data Analysis for Malayalam YouTube Comments: Sentiment Analysis and Emotion Detection using ML and DL Models
Abeera V P, Dr. Sachin Kumar and Dr. Soman K P 43

Findings of the Second Shared Task on Offensive Span Identification from Code-Mixed Tamil-English Comments
Manikandan Ravikiran, Ananth Ganesh, Anand Kumar M, R Rajalakshmi and Bharathi Raja Chakravarthi 52

Overview of the shared task on Fake News Detection from Social Media Text
Malliga S, Bharathi Raja Chakravarthi, Kogilavani S V, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani and Muskaan Singh 59

Findings of the Shared Task on Sentiment Analysis in Tamil and Tulu Code-Mixed Text
Asha Hegde, Bharathi Raja Chakravarthi, Hosahalli Lakshmaiah Shashirekha, Rahul Ponnusamy, SUBALALITHA CN, Lavanya S K, Thenmozhi D., Martha Karunakar, Shreya Shreeram and Sarah Aymen 64

Findings of the Shared Task on Multimodal Abusive Language Detection and Sentiment Analysis in Tamil and Malayalam
Premjith B, Jyothish Lal G, Sowmya V, Bharathi Raja Chakravarthi, Rajeswari Natarajan, Nandhini K, Abirami Murugappan, Bharathi B, Kaushik M, Prasanth SN, Aswin Raj R and Vijai Simmon S . . . 72

Overview of Shared-task on Abusive Comment Detection in Tamil and Telugu
Ruba Priyadharshini, Bharathi Raja Chakravarthi, Malliga S, SUBALALITHA CN, Kogilavani S V, Premjith B, Abirami Murugappan and Prasanna Kumar Kumaresan 80

CoPara: The First Dravidian Paragraph-level n-way Aligned Corpus
Nikhil E, Mukund Choudhary and Radhika Mamidi 88

ChatGPT_Powered_Tourist_Aid_Applications__Proficient_in_Hindi__Yet_To_Master_Telugu_and_Kannada
Sanjana Kolar and Rohit Kumar 97

Enhancing Telugu News Understanding: Comparative Study of ML Algorithms for Category Prediction
Manish Rama Gopal Nadella, Venkata Krishna Rayalu Garapati, Eswar Sudhan S.K., Gouthami Jangala, Soman K.P. and Sachin Kumar 108

Revisiting Automatic Speech Recognition for Tamil and Hindi Connected Number Recognition
Rahul Mishra, Senthil Raja Gunaseela Boopathy, Manikandan Ravikiran, Shreyas Kulkarni, Mayurakshi Mukherjee, Ananth Ganesh and Kingshuk Banerjee 116

Poorvi@DravidianLangTech: Sentiment Analysis on Code-Mixed Tulu and Tamil Corpus
Poorvi Shetty 124

NLP_SSN_CSE@DravidianLangTech: Fake News Detection in Dravidian Languages using Transformer Models
Varsha Balaji, Shahul Hameed T and Bharathi B 133

AbhiPaw@DravidianLangTech: Multimodal Abusive Language Detection and Sentiment Analysis
Abhinaba Bala and Parameswari Krishnamurthy 140

Athena@DravidianLangTech: Abusive Comment Detection in Code-Mixed Languages using Machine Learning Techniques
Hema M, Anza Prem, Rajalakshmi Sivanaiah and Angel Deborah S 147

AlphaBrains@DravidianLangTech: Sentiment Analysis of Code-Mixed Tamil and Tulu by Training Contextualized ELMo Word Representations
Toqeer Ehsan, Amina Tehseen, Kengatharaiyer Sarveswaran and Amjad Ali 152

HARMONY@DravidianLangTech: Transformer-based Ensemble Learning for Abusive Comment Detection
Amrish Raaj P, Abirami Murugappan, Lysa Packiam R S and Deivamani M 160

Avalanche@DravidianLangTech: Abusive Comment Detection in Code Mixed Data Using Machine Learning Techniques with Under Sampling
Rajalakshmi Sivanaiah, Rajasekar S, Srilakshmisai K, Angel Deborah S and Mirnalinee ThankaNadar 166

DeepBlueAI@DravidianLangTech-RANLP 2023
Zhipeng Luo and Jiahui Wang 171

Selam@DravidianLangTech:Sentiment Analysis of Code-Mixed Dravidian Texts using SVM Classification
Selam Kanta and Grigori Sidorov 176

LIDOMA@DravidianLangTech: Convolutional Neural Networks for Studying Correlation Between Lexical Features and Sentiment Polarity in Tamil and Tulu Languages
Moein Tash, Jesus Armenta-Segura, zahra ahani, olga kolesnikova, Grigori Sidorov and Alexander Gelbukh 180

nlpt malayalm@DravidianLangTech : Fake News Detection in Malayalam using Optimized XLM-RoBERTa Model
Eduri Raja, Badal Soni and Sami Kumar Borgohain 186

ML&AI_IIITRanchi@DravidianLangTech: Fine-Tuning IndicBERT for Exploring Language-specific Features for Sentiment Classification in Code-Mixed Dravidian Languages
Kirti Kumari, Shirish Shekhar Jha, Zarikunte Kunal Dayanand and Praneesh Sharma 192

ML&AI_IIITRanchi@DravidianLangTech:Leveraging Transfer Learning for the discernment of Fake News within the Linguistic Domain of Dravidian Language
Kirti Kumari, Shirish Shekhar Jha, Zarikunte Kunal Dayanand and Praneesh Sharma 198

NITK-IT-NLP@DravidianLangTech: Impact of Focal Loss on Malayalam Fake News Detection using Transformers
Hariharan R L and Anand Kumar M 207

VEL@DravidianLangTech: Sentiment Analysis of Tamil and Tulu
Kishore Kumar Ponnusamy, Charmathi Rajkumar, Prasanna Kumar Kumaresan, Elizabeth Sherly and Ruba Priyadharshini 211

hate-alert@DravidianLangTech: Multimodal Abusive Language Detection and Sentiment Analysis in Dravidian Languages
Shubhankar Barman and Mithun Das 217

Supernova@DravidianLangTech 2023@Abusive Comment Detection in Tamil and Telugu - (Tamil, Tamil-English, Telugu-English)
Ankitha Reddy, Pranav Moorthi and Ann Maria Thomas 225

AbhiPaw@ DravidianLangTech: Abusive Comment Detection in Tamil and Telugu using Logistic Regression
Abhinaba Bala and Parameswari Krishnamurthy 231

AbhiPaw@ DravidianLangTech: Fake News Detection in Dravidian Languages using Multilingual BERT
Abhinaba Bala and Parameswari Krishnamurthy 235

Habesha@DravidianLangTech: Utilizing Deep and Transfer Learning Approaches for Sentiment Analysis.
Mesay Gemeda Yigezu, Tadesse Kebede, Olga Kolesnikova, Grigori Sidorov and Alexander Gelbukh 239

Habesha@DravidianLangTech: Abusive Comment Detection using Deep Learning Approach
Mesay Gemeda Yigezu, Selam Kanta, Olga Kolesnikova, Grigori Sidorov and Alexander Gelbukh 244

SADTech@DravidianLangTech: Multimodal Sentiment Analysis of Tamil and Malayalam
Abhinav Patil, Sam Briggs, Tara Wueger and Daniel D. O'Connell 250

MUCS@DravidianLangTech2023: Sentiment Analysis in Code-mixed Tamil and Tulu Texts using fastText
Rachana K, Prajnashree M, Asha Hegde and H. L Shashirekha 258

MUCS@DravidianLangTech2023: Leveraging Learning Models to Identify Abusive Comments in Code-mixed Dravidian Languages
Asha Hegde, Kavya G, Sharal Coelho and Hosahalli Lakshmaiah Shashirekha 266

MUNLP@DravidianLangTech2023: Learning Approaches for Sentiment Analysis in Code-mixed Tamil and Tulu Text
Asha Hegde, Kavya G, Sharal Coelho, Pooja Lamani and Hosahalli Lakshmaiah Shashirekha . 275

MUCSD@DravidianLangTech2023: Predicting Sentiment in Social Media Text using Machine Learning Techniques
Sharal Coelho, Asha Hegde, Pooja Lamani, Kavya G and Hosahalli Lakshmaiah Shashirekha . 282

MUCS@DravidianLangTech2023: Malayalam Fake News Detection Using Machine Learning Approach
Sharal Coelho, Asha Hegde, Kavya G and Hosahalli Lakshmaiah Shashirekha 288

KEC_AI_NLP@DravidianLangTech: Abusive Comment Detection in Tamil Language
Kogilavani Shanmugavadivel, Malliga Subramanian, Shri Durga R, SRIGHA S, Sree Harene J S and Yasvanth Bala P 293

KEC_AI_NLP@DravidianLangTech: Sentiment Analysis in Code Mixture Language
Kogilavani Shanmugavadivel, Malliga Subaramanian, VetriVendhan S, Pramoth Kumar M, Karthickeyan S and Kavin Vishnu N 300

CSSCUTN@DravidianLangTech:Abusive comments Detection in Tamil and Telugu
Kathiravan Pannerselvam, Saranya Rajiakodi, Rahul Ponnusamy and Sajeetha Thavareesan . . . 306