

First Workshop on Speech and Language Technologies for Dravidian Languages (DravidianLangTech 2021)

Online
19 April 2021

ISBN: 978-1-7138-2959-1

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© (2021) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

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>Tamil Lyrics Corpus: Analysis and Experiments</i>	
Dhivya Chinnappa and Praveenraj Dhandapani	1
<i>DOSA: Dravidian Code-Mixed Offensive Span Identification Dataset</i>	
Manikandan Ravikiran and Subbiah Annamalai	10
<i>Towards Offensive Language Identification for Dravidian Languages</i>	
Siva Sai and Yashvardhan Sharma	18
<i>Sentiment Classification of Code-Mixed Tweets using Bi-Directional RNN and Language Tags</i>	
Sainik Mahata, Dipankar Das and Sivaji Bandyopadhyay	28
<i>Offensive language identification in Dravidian code mixed social media text</i>	
SUNIL SAUMYA, Abhinav Kumar and Jyoti Prakash Singh	36
<i>Sentiment Analysis of Dravidian Code Mixed Data</i>	
Asrita Venkata Mandalam and Yashvardhan Sharma	46
<i>Unsupervised Machine Translation On Dravidian Languages</i>	
Sai Koneru, Danni Liu and Jan Niehues	55
<i>Graph Convolutional Networks with Multi-headed Attention for Code-Mixed Sentiment Analysis</i>	
Suman Dowlagar and Radhika Mamidi	65
<i>Task-Specific Pre-Training and Cross Lingual Transfer for Sentiment Analysis in Dravidian Code-Switched Languages</i>	
Akshat Gupta, Sai Krishna Rallabandi and Alan W Black	73
<i>Analysis of Uvama Urubugal in Tamil Sangam Literatures</i>	
SUBALALITHA CN	80
<i>Task-Oriented Dialog Systems for Dravidian Languages</i>	
Tushar Kanakagiri and Karthik Radhakrishnan	85
<i>A Survey on Paralinguistics in Tamil Speech Processing</i>	
Anosha Ignatius and Uthayasanker Thayasivam	94
<i>Is this Enough?-Evaluation of Malayalam Wordnet</i>	
Nandu Chandran Nair, Maria-chiara Giangregorio and Fausto Giunchiglia	100
<i>LA-SACo: A Study of Learning Approaches for Sentiments Analysis in Code-Mixing Texts</i>	
Fazlourrahman Balouchzahi and H L Shashirekha	109
<i>Findings of the Shared Task on Machine Translation in Dravidian languages</i>	
Bharathi Raja Chakravarthi, Ruba Priyadharshini, Shubhanker Banerjee, Richard Saldanha, John P. McCrae, Anand Kumar M, Parameswari Krishnamurthy and Melvin Johnson	119
<i>Findings of the Shared Task on Troll Meme Classification in Tamil</i>	
Shardul Suryawanshi and Bharathi Raja Chakravarthi	126

<i>Findings of the Shared Task on Offensive Language Identification in Tamil, Malayalam, and Kannada</i> Bharathi Raja Chakravarthi, Ruba Priyadharshini, Navya Jose, Anand Kumar M, Thomas Mandl, Prasanna Kumar Kumaresan, Rahul Ponnusamy, Hariharan R L, John P. McCrae and Elizabeth Sherly	133
<i>GX@DravidianLangTech-EACL2021: Multilingual Neural Machine Translation and Back-translation</i> Wanying Xie	146
<i>OFFLangOne@DravidianLangTech-EACL2021: Transformers with the Class Balanced Loss for Offensive Language Identification in Dravidian Code-Mixed text.</i> Suman Dowlagar and Radhika Mamidi	154
<i>Simon @ DravidianLangTech-EACL2021: Detecting Offensive Content in Kannada Language</i> Qinyu Que	160
<i>Codewithzichao@DravidianLangTech-EACL2021: Exploring Multilingual Transformers for Offensive Language Identification on Code Mixing Text</i> Zichao Li	164
<i>JudithJeyafreedaAndrew@DravidianLangTech-EACL2021:Offensive language detection for Dravidian Code-mixed YouTube comments</i> Judith Jeyafreeda Andrew	169
<i>professionals@DravidianLangTech-EACL2021: Malayalam Offensive Language Identification - A Minimalistic Approach</i> Srinath Nair and Dolton Fernandes	175
<i>UVCE-IIITT@DravidianLangTech-EACL2021: Tamil Troll Meme Classification: You need to Pay more Attention</i> Siddhanth U Hegde, Adeep Hande, Ruba Priyadharshini, Sajeetha Thavareesan and Bharathi Raja Chakravarthi	180
<i>IIITT@DravidianLangTech-EACL2021: Transfer Learning for Offensive Language Detection in Dravidian Languages</i> Konthala Ysaswini, Karthik Puranik, Adeep Hande, Ruba Priyadharshini, Sajeetha Thavareesan and Bharathi Raja Chakravarthi	187
<i>Hypers@DravidianLangTech-EACL2021: Offensive language identification in Dravidian code-mixed YouTube Comments and Posts</i> Charangan Vasantharajan and Uthayasanker Thayasivam	195
<i>HUB@DravidianLangTech-EACL2021: Identify and Classify Offensive Text in Multilingual Code Mixing in Social Media</i> Bo Huang and Yang Bai	203
<i>HUB@DravidianLangTech-EACL2021: Meme Classification for Tamil Text-Image Fusion</i> Bo Huang and Yang Bai	210
<i>ZYJ123@DravidianLangTech-EACL2021: Offensive Language Identification based on XLM-RoBERTa with DPCNN</i> Yingjia Zhao and Xin Tao	216

<i>IITK@DravidianLangTech-EACL2021: Offensive Language Identification and Meme Classification in Tamil, Malayalam and Kannada</i>	
Nikhil Ghanghor, Parameswari Krishnamurthy, Sajeetha Thavareesan, Ruba Priyadarshini and Bharathi Raja Chakravarthi	222
<i>cs@DravidianLangTech-EACL2021: Offensive Language Identification Based On Multilingual BERT Model</i>	
shi chen and bing kong	230
<i>CUSATNLP@DravidianLangTech-EACL2021:Language Agnostic Classification of Offensive Content in Tweets</i>	
Sara Renjit and Sumam Mary Idicula	236
<i>IIT_DWD@EACL2021: Identifying Troll Meme in Tamil using a hybrid deep learning approach</i>	
Ankit Kumar Mishra and SUNIL SAUMYA	243
<i>Amrita_CEN_NLP@DravidianLangTech-EACL2021: Deep Learning-based Offensive Language Identification in Malayalam, Tamil and Kannada</i>	
Sreelakshmi K, Premjith B and Soman KP	249
<i>NLP-CUET@DravidianLangTech-EACL2021: Offensive Language Detection from Multilingual Code-Mixed Text using Transformers</i>	
Omar Sharif, Eftekhari Hossain and Mohammed Moshuiul Hoque	255
<i>IRLAB-DAIICT@DravidianLangTech-EACL2021: Neural Machine Translation</i>	
Raj Prajapati, Vedant Vijay Parikh and Prasenjit Majumder	262
<i>IRNLP_DAIICT@DravidianLangTech-EACL2021:Offensive Language identification in Dravidian Languages using TF-IDF Char N-grams and MuRIL</i>	
Bhargav Dave, Shripad Bhat and Prasenjit Majumder	266
<i>Hate-Alert@DravidianLangTech-EACL2021: Ensembling strategies for Transformer-based Offensive language Detection</i>	
Debjoy Saha, Naman Paharia, Debajit Chakraborty, Punyajoy Saha and Animesh Mukherjee . .	270
<i>TrollMeta@DravidianLangTech-EACL2021: Meme classification using deep learning</i>	
Manoj Balaji J and Chinmaya HS	277
<i>Maoqin @ DravidianLangTech-EACL2021: The Application of Transformer-Based Model</i>	
Maoqin Yang	281
<i>Simon @ DravidianLangTech-EACL2021: Meme Classification for Tamil with BERT</i>	
Qinyu Que	287
<i>Bitions@DravidianLangTech-EACL2021: Ensemble of Multilingual Language Models with Pseudo Labeling for offence Detection in Dravidian Languages</i>	
Debapriya Tula, Prathyush Potluri, Shreyas MS, Sumanth Doddapaneni, Pranjal Sahu, Rohan Sukumar and Parth Patwa	291
<i>NLP-CUET@DravidianLangTech-EACL2021: Investigating Visual and Textual Features to Identify Trolls from Multimodal Social Media Memes</i>	
Eftekhari Hossain, Omar Sharif and Mohammed Moshuiul Hoque	300

<i>SJ_AJ@DravidianLangTech-EACL2021: Task-Adaptive Pre-Training of Multilingual BERT models for Offensive Language Identification</i>	
Sai Muralidhar Jayanthi and Akshat Gupta	307
<i>SSNCSE_NLP@DravidianLangTech-EACL2021: Offensive Language Identification on Multilingual Code Mixing Text</i>	
Bharathi B and Agnusimmaculate Silvia A	313
<i>JUNLP@DravidianLangTech-EACL2021: Offensive Language Identification in Dravidian Languages</i>	
Avishek Garain, Atanu Mandal and Sudip Kumar Naskar	319
<i>MUCS@DravidianLangTech-EACL2021:COOLI-Code-Mixing Offensive Language Identification</i>	
Fazlourrahman Balouchzahi, Aparna B K and H L Shashirekha	323
<i>indicnlp@kgp at DravidianLangTech-EACL2021: Offensive Language Identification in Dravidian Languages</i>	
Kushal Kedia and Abhilash Nandy	330
<i>SSNCSE_NLP@DravidianLangTech-EACL2021: Meme classification for Tamil using machine learning approach</i>	
Bharathi B and Agnusimmaculate Silvia A	336
<i>MUCS@ - Machine Translation for Dravidian Languages using Stacked Long Short Term Memory</i>	
Asha Hegde, Ibrahim Gashaw and Shashirekha H.L.	340
<i>OffTamil@DravideanLangTech-EASL2021: Offensive Language Identification in Tamil Text</i>	
Disne Sivalingam and Sajeetha Thavareesan	346
<i>Codewithzichao@DravidianLangTech-EACL2021: Exploring Multimodal Transformers for Meme Classification in Tamil Language</i>	
Zichao Li	352
<i>DLRG@DravidianLangTech-EACL2021: Transformer based approach for Offensive Language Identification on Code-Mixed Tamil</i>	
Ratnavel Rajalakshmi, Yashwant Reddy and Lokesh Kumar	357