

ICON 2021: 18th International Conference on Natural Language  
Processing  
National Institute of Technology - Silchar  
December 16-19, 2021

**Shared Task on Multilingual Gender Biased and Communal  
Language Identification**

**PROCEEDINGS**

Editors:

Ritesh Kumar, Siddharth Singh, Enakshi Nandi, Shyam Ratan,  
Laishram Niranjana Devi, Bornini Lahiri, Akanksha Bansal,  
Akash Bhagat, Yogesh Dawer



## Introduction

Aggression and its manifestations in different forms have taken unprecedented proportions with the tremendous growth of Internet and social media. The research community, especially within the fields of Linguistics and Natural Language Processing, has responded by understanding the pragmatic and structural aspects of such forms of language usage and several other) and developing systems that could automatically detect and handle these.

In the ComMA project, we are working on these different aspects of aggressive and offensive language usage online and its automatic identification. As part of our efforts in the project, in this task, we present a novel multi-label classification task to the research community, in which each sample will be required to be classified as aggressive, gender biased or communally charged in 4 languages - Meitei, Bangla, Hindi and English. The motivation behind this is to understand the intersectionality across these three categories and also explore if using intersectional data could be helpful in the task or not.

We attracted a total of 54 registrations in the task, out of which 11 teams submitted their test runs. In this proceedings, we have included system description papers of 8 teams along with a task overview paper. We hope that the task and its findings will be interesting for researchers working in the different related areas of hate speech, offensive language, abusive language as well more generally in text classification.

Shared task page: <https://sites.google.com/view/comma-at-icon2021/home>

Main conference page: <http://icon2021.nits.ac.in/index.html>

**ComMA@ICON Shared Task Organisers**



## **Organising Committee**

Ritesh Kumar, Dr. Bhimrao Ambedkar University, Agra  
Bornini Lahiri, Indian Institute of Technology-Kharagpur  
Akanksha Bansal, Panlingua Language Processing LLP, New Delhi  
Akash Bhagat, Indian Institute of Technology-Kharagpur  
Enakshi Nandi, Panlingua Language Processing LLP, New Delhi  
Laishram Niranjana Devi, Panlingua Language Processing LLP, New Delhi  
Shyam Ratan, Dr. Bhimrao Ambedkar University, Agra  
Siddharth Singh, Dr. Bhimrao Ambedkar University, Agra  
Yogesh Dawer, Dr. Bhimrao Ambedkar University, Agra

## **Editors**

Ritesh Kumar, Dr. Bhimrao Ambedkar University, Agra  
Siddharth Singh, Dr. Bhimrao Ambedkar University, Agra  
Enakshi Nandi, Panlingua Language Processing LLP, New Delhi  
Shyam Ratan, Dr. Bhimrao Ambedkar University, Agra  
Laishram Niranjana Devi, Panlingua Language Processing LLP, New Delhi  
Bornini Lahiri, Indian Institute of Technology Kharagpur  
Akanksha Bansal, Panlingua Language Processing LLP, New Delhi  
Akash Bhagat, Indian Institute of Technology-Kharagpur  
Yogesh Dawer, Dr. Bhimrao Ambedkar University, Agra



## Table of Contents

<i>ComMA@ICON: Multilingual Gender Biased and Communal Language Identification Task at ICON-2021</i>	
Ritesh Kumar, Shyam Ratan, Siddharth Singh, Enakshi Nandi, Laishram Niranjana Devi, Akash Bhagat, Yogesh Dawer, bornini lahiri and Akanksha Bansal . . . . .	1
<i>Team_BUDDI at ComMA@ICON: Exploring Individual and Joint Modelling Approaches for Detecting Aggression, Communal Bias and Gender Bias</i>	
Anand Subramanian, Mukesh Reghu and Sriram Rajkumar . . . . .	13
<i>Hypers at ComMA@ICON: Modelling Aggressive, Gender Bias and Communal Bias Identification</i>	
Sean Benhur, Roshan Nayak, Kanchana Sivanraju, Adeep Hande, CN Subalalitha, Ruba Priyadharshini and Bharathi Raja Chakravarthi . . . . .	21
<i>Beware Haters at ComMA@ICON: Sequence and Ensemble Classifiers for Aggression, Gender Bias and Communal Bias Identification in Indian Languages</i>	
Deepakindresh Gandhi, Aakash Ambalavanan, Avireddy Rohan and Radhika Selvamani . . . . .	26
<i>DELab@IITSM at ICON-2021 Shared Task: Identification of Aggression and Biasness Using Decision Tree</i>	
Maibam Debina and Navanath Saharia . . . . .	35
<i>LUC at ComMA-2021 Shared Task: Multilingual Gender Biased and Communal Language Identification without Using Linguistic Features</i>	
Rodrigo Cuéllar-Hidalgo, Julio de Jesús Guerrero-Zambrano, Dominic Forest, Gerardo Reyes-Salgado and Juan-Manuel Torres-Moreno . . . . .	41
<i>ARGUABLY at ComMA@ICON: Detection of Multilingual Aggressive, Gender Biased, and Communally Charged Tweets Using Ensemble and Fine-Tuned IndicBERT</i>	
Guneet Kohli, Prabsimran Kaur and Jatin Bedi . . . . .	46
<i>Sdutta at ComMA@ICON: A CNN-LSTM Model for Hate Detection</i>	
Sandip Dutta, Utso Majumder and Sudip Naskar . . . . .	53
<i>MUCIC at ComMA@ICON: Multilingual Gender Biased and Communal Language Identification Using N-grams and Multilingual Sentence Encoders</i>	
Fazlourrahman Balouchzahi, Oxana Vitman, Hosahalli Lakshmaiah Shashirekha, Grigori Sidorov and Alexander Gelbukh . . . . .	58
<i>MUM at ComMA@ICON: Multilingual Gender Biased and Communal Language Identification Using Supervised Learning Approaches</i>	
Asha Hegde, Mudoor Devadas Anusha, Sharal Coelho and Hosahalli Lakshmaiah Shashirekha .	64
<i>BFCAl at ComMA@ICON 2021: Support Vector Machines for Multilingual Gender Biased and Communal Language Identification</i>	
Fathy Elkazzaz, Fatma Sakr, Rasha Orban and Hamada Nayel . . . . .	70





## Shared Task Program

**Sunday December 19, 2021**

**10:00–11:18 Paper Session I**

Chair: TBD

10:00–10:13 *ComMA@ICON: Multilingual Gender Biased and Communal Language Identification Task at ICON-2021*

Ritesh Kumar, Shyam Ratan, Siddharth Singh, Enakshi Nandi, Laishram Niranjana Devi, Akash Bhagat, Yogesh Dawer, bornini lahiri and Akanksha Bansal

10:13–10:26 *Team\_BUDDI at ComMA@ICON: Exploring Individual and Joint Modelling Approaches for Detecting Aggression, Communal Bias and Gender Bias*

Anand Subramanian, Mukesh Reghu and Sriram Rajkumar

10:26–10:39 *Hypers at ComMA@ICON: Modelling Aggressive, Gender Bias and Communal Bias Identification*

Sean Benhur, Roshan Nayak, Kanchana Sivanraju, Adeep Hande, CN Subalalitha, Ruba Priyadharshini and Bharathi Raja Chakravarthi

10:39–10:52 *Beware Haters at ComMA@ICON: Sequence and Ensemble Classifiers for Aggression, Gender Bias and Communal Bias Identification in Indian Languages*

Deepakindresh Gandhi, Aakash Ambalavanan, Avireddy Rohan and Radhika Selvamani

10:52–11:05 *DELab@IIITSM at ICON-2021 Shared Task: Identification of Aggression and Biasness Using Decision Tree*

Maibam Debina and Navanath Saharia

11:05–11:18 *LUC at ComMA-2021 Shared Task: Multilingual Gender Biased and Communal Language Identification without Using Linguistic Features*

Rodrigo Cuéllar-Hidalgo, Julio de Jesús Guerrero-Zambrano, Dominic Forest, Gerardo Reyes-Salgado and Juan-Manuel Torres-Moreno

**11:18–11:28 Break**

**Sunday December 19, 2021 (continued)**

**11:28–12:33 Paper Session II**

Chair: TBD

11:28–11:41 *ARGUABLY at ComMA@ICON: Detection of Multilingual Aggressive, Gender Biased, and Communally Charged Tweets Using Ensemble and Fine-Tuned IndicBERT*

Guneet Kohli, Prabsimran Kaur and Jatin Bedi

11:41–11:54 *Sdutta at ComMA@ICON: A CNN-LSTM Model for Hate Detection*

Sandip Dutta, Utso Majumder and Sudip Naskar

11:54–12:07 *MUCIC at ComMA@ICON: Multilingual Gender Biased and Communal Language Identification Using N-grams and Multilingual Sentence Encoders*

Fazlourrahman Balouchzahi, Oxana Vitman, Hosahalli Lakshmaiah Shashirekha, Grigori Sidorov and Alexander Gelbukh

12:07–12:20 *MUM at ComMA@ICON: Multilingual Gender Biased and Communal Language Identification Using Supervised Learning Approaches*

Asha Hegde, Mudoor Devadas Anusha, Sharal Coelho and Hosahalli Lakshmaiah Shashirekha

12:20–12:33 *BFCAI at ComMA@ICON 2021: Support Vector Machines for Multilingual Gender Biased and Communal Language Identification*

Fathy Elkazzaz, Fatma Sakr, Rasha Orban and Hamada Nayel