

The ComMA Dataset V0.2: Annotating Aggression and Bias in Multilingual Social Media Discourse

LREC 2022

Authors:

***Ritesh Kumar¹, Shyam Ratan¹, Siddharth Singh¹, Enakshi Nandi²,
Laishram Niranjana Devi², Akash Bhagat³, Yogesh Dawer¹,
Bornini Lahiri³, Akanksha Bansal², Atul Kr. Ojha^{4,2}***

Institutional Affiliations:

¹Dr Bhimrao Ambedkar University, Agra

²Panlingua Language Processing LLP, New Delhi

Indian ³Institute of Technology-Kharagpur, Kharagpur

⁴DSI, National University of Ireland Galway, Galway

Email id: comma.kmi@gmail.com

The ComM1 Dataset V0.2: Annotating Aggression and Bias in Multilingual Social Media Discourse

Ritesh Kumar¹, Shyam Ratan¹, Siddharth Singh¹, Enakshi Nandi², Laishram Niranjana Devi², Akash Bhagat³, Yogesh Dawer¹, Bornini Lahiri³, Akanksha Bansal², Atul Kr. Ojha^{4,2}
¹Dr Bhimrao Ambedkar University, ²Panlingua Language Processing LLP, ³Indian Institute of Technology-Kharagpur ⁴DSI, National University of Ireland Galway
 New Delhi, Agra, Kharagpur, Galway
 comma.kmi@gmail.com

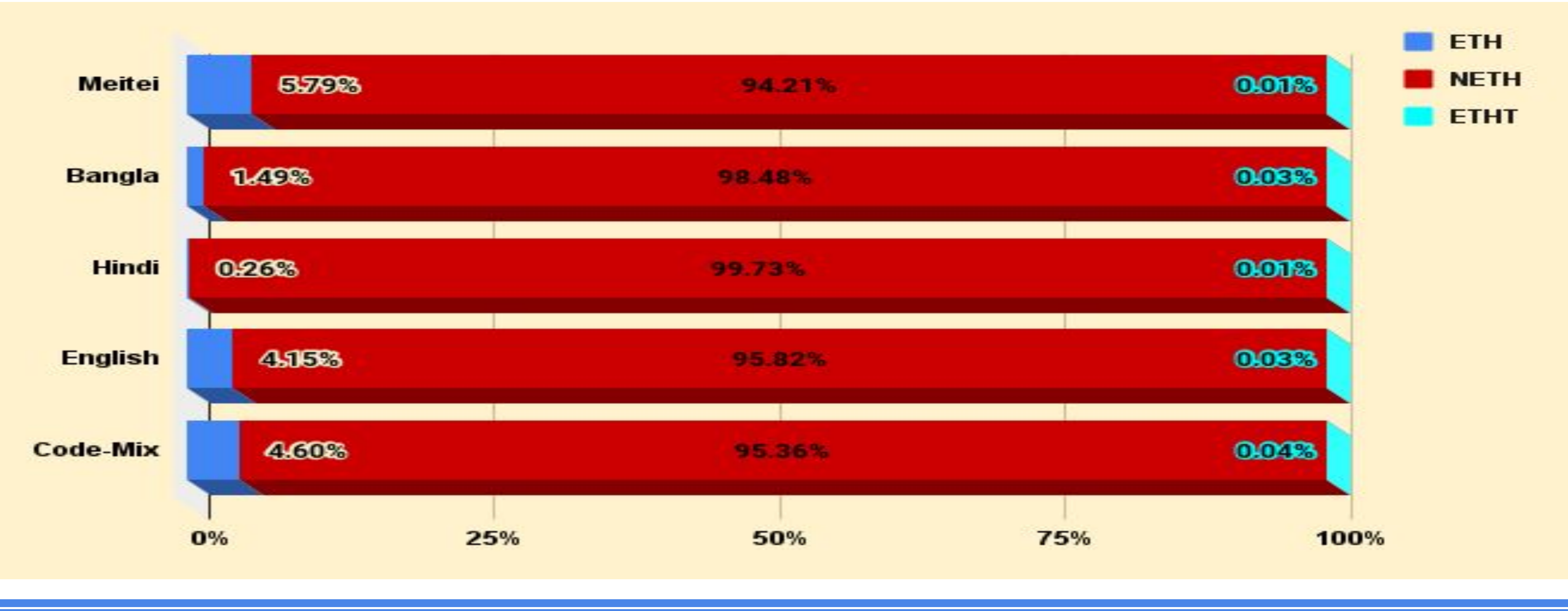
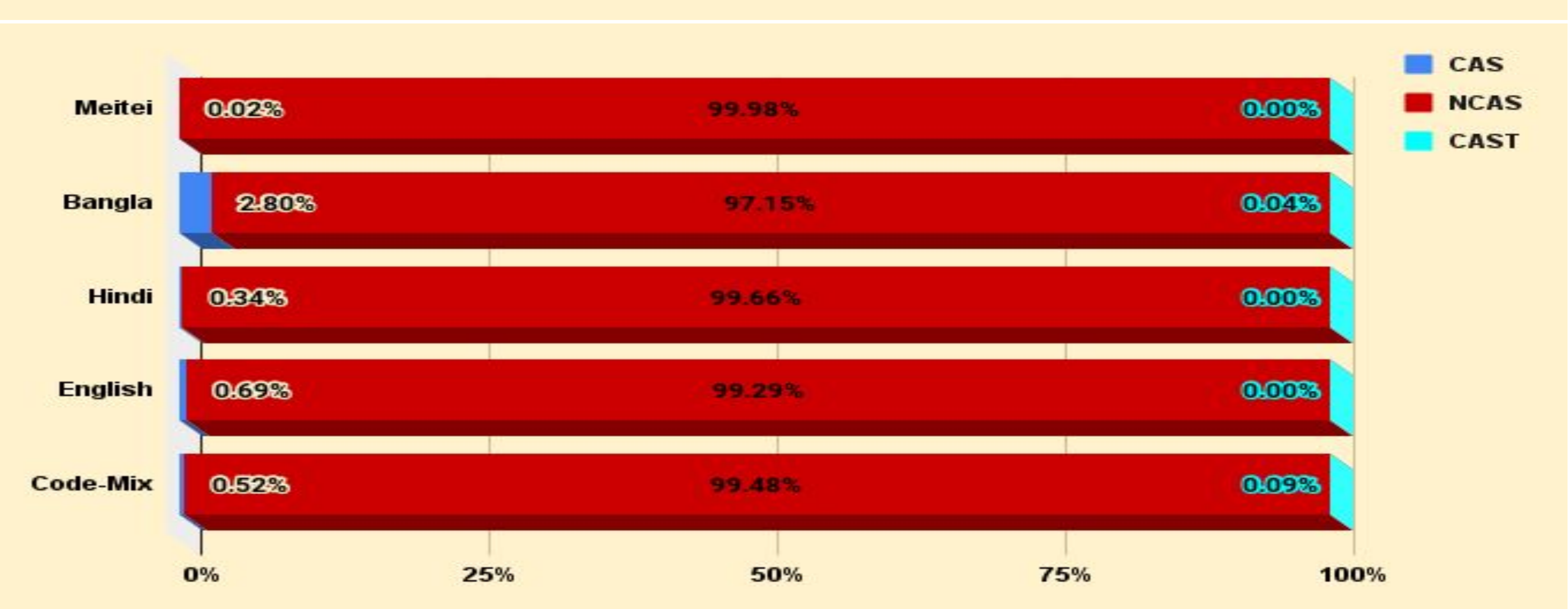
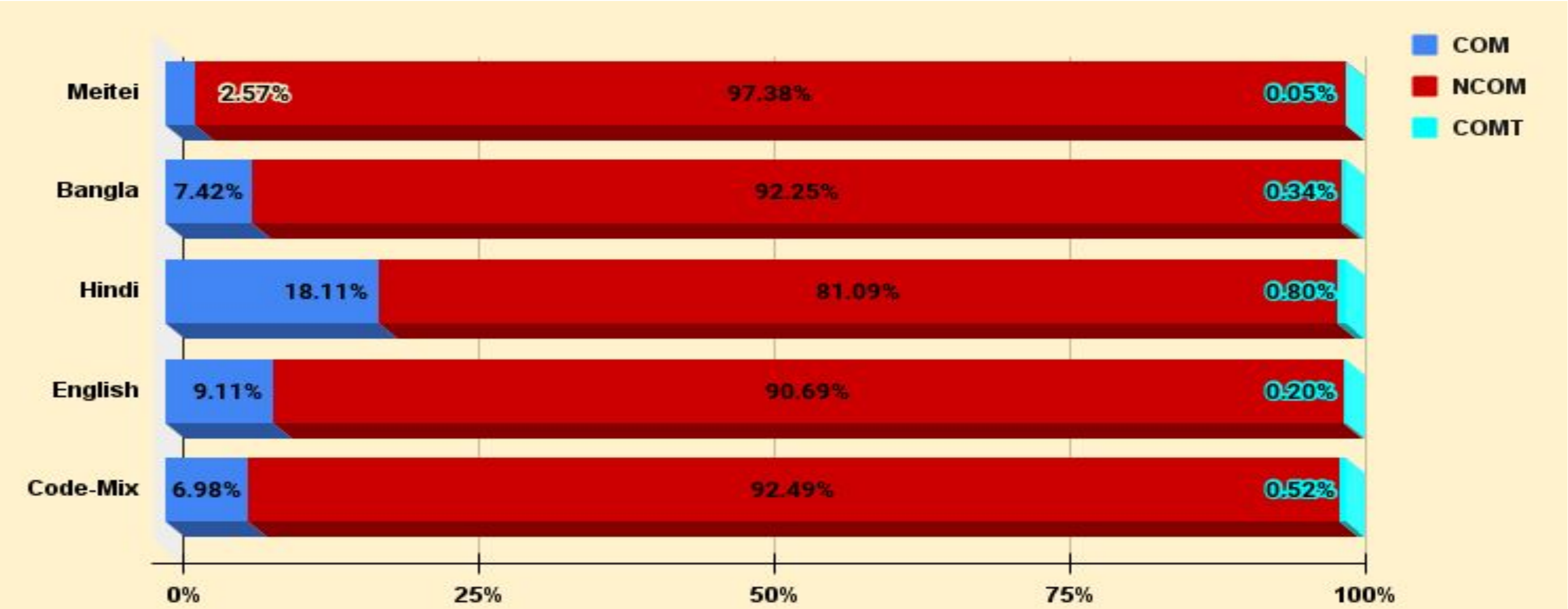
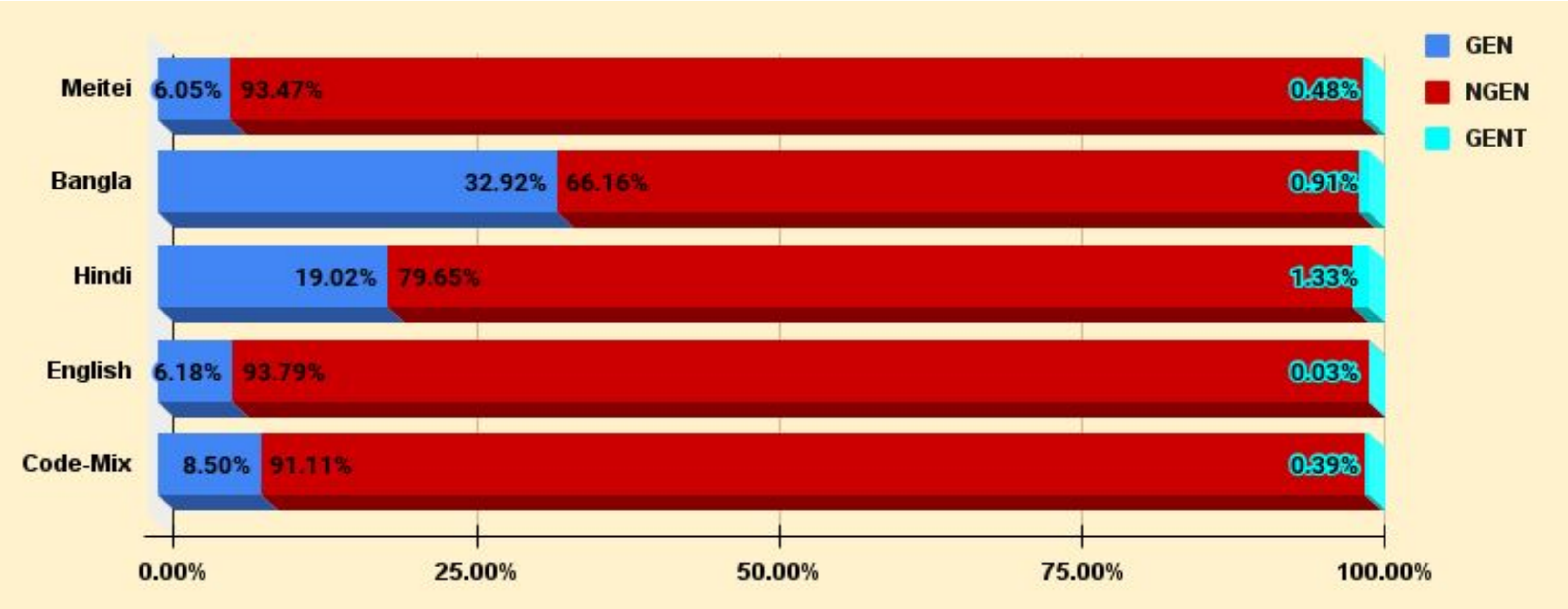
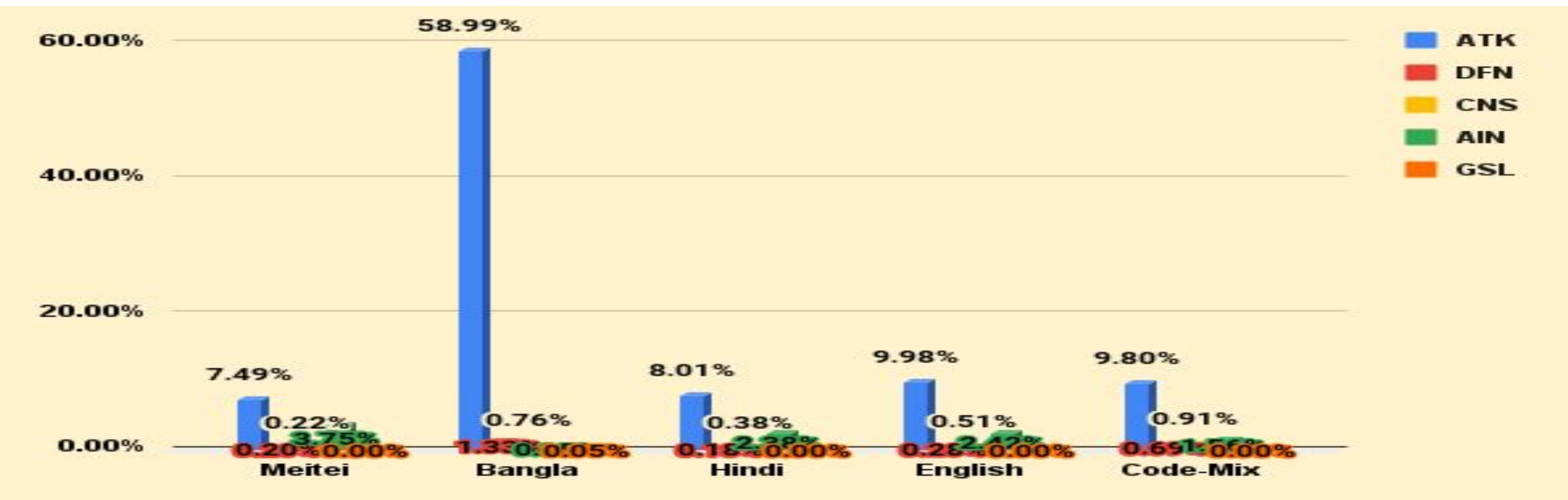
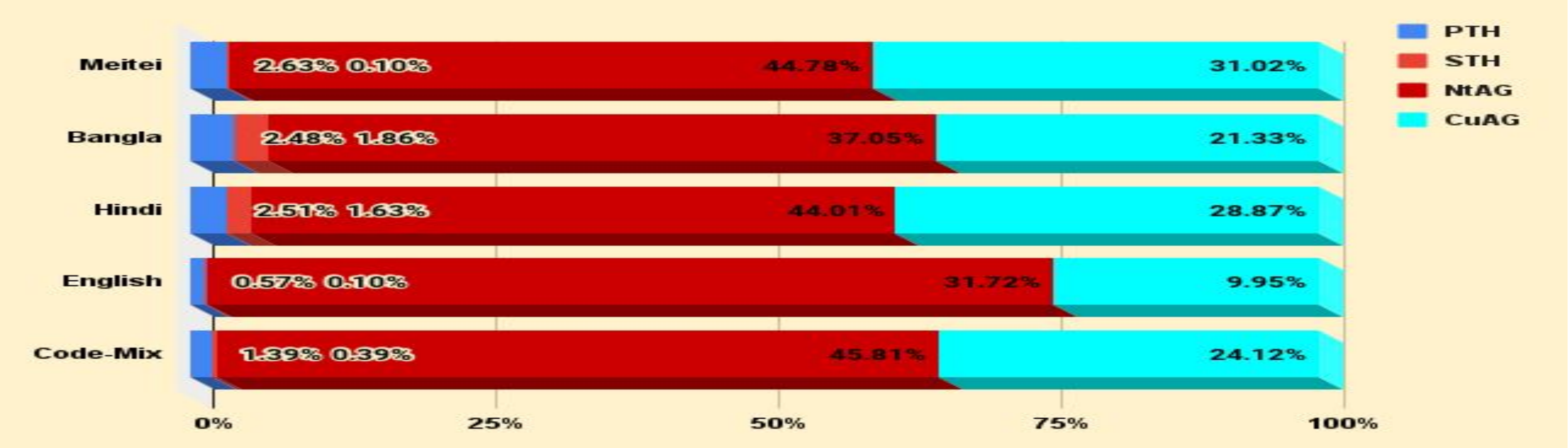
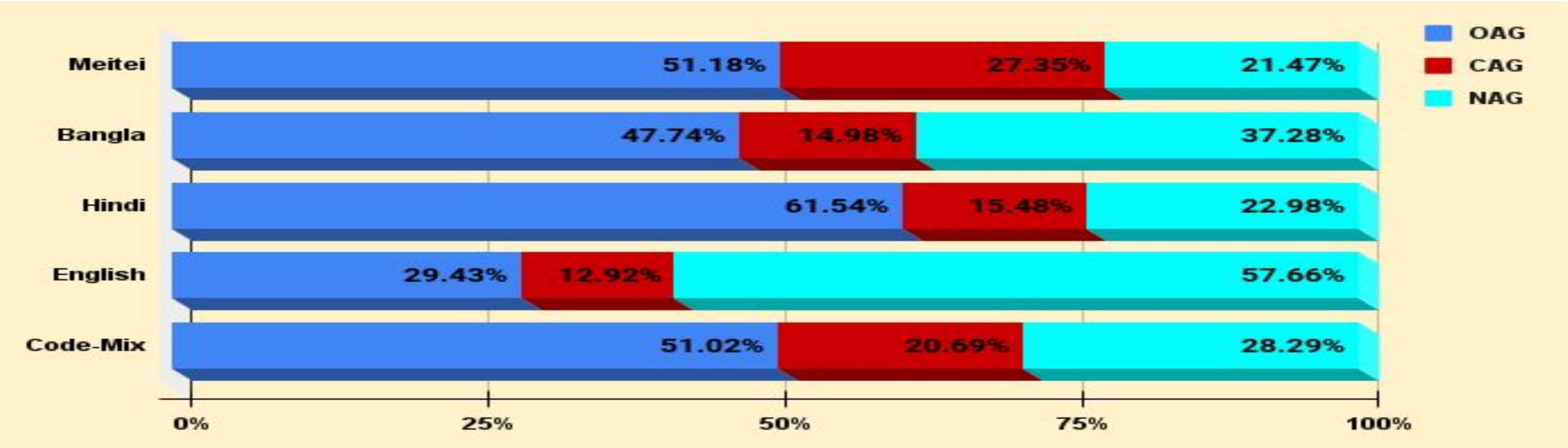
Abstract

Here, we discuss the development of a multilingual dataset annotated with a hierarchical, fine-grained tagset marking different types of aggression and the "context" in which they occur. The context, here, is defined by the conversational thread in which a specific comment occurs and also the "type" of discursive role that the comment is performing with respect to the previous comment. The initial dataset, being discussed here consists of a total 59,152 annotated comments in four languages - Meitei, Bangla, Hindi, and Indian English - collected from various social media platforms such as YouTube, Facebook, Twitter and Telegram. As is usual on social media websites, a large number of these comments are multilingual, mostly code-mixed with English. The paper gives a detailed description of the tagset being used for annotation and also the process of developing a multi-label, fine-grained tagset that has been used for marking comments with aggression and bias of various kinds including gender bias, religious intolerance (called communal bias in the tagset), class/caste bias and ethnic/racial bias.

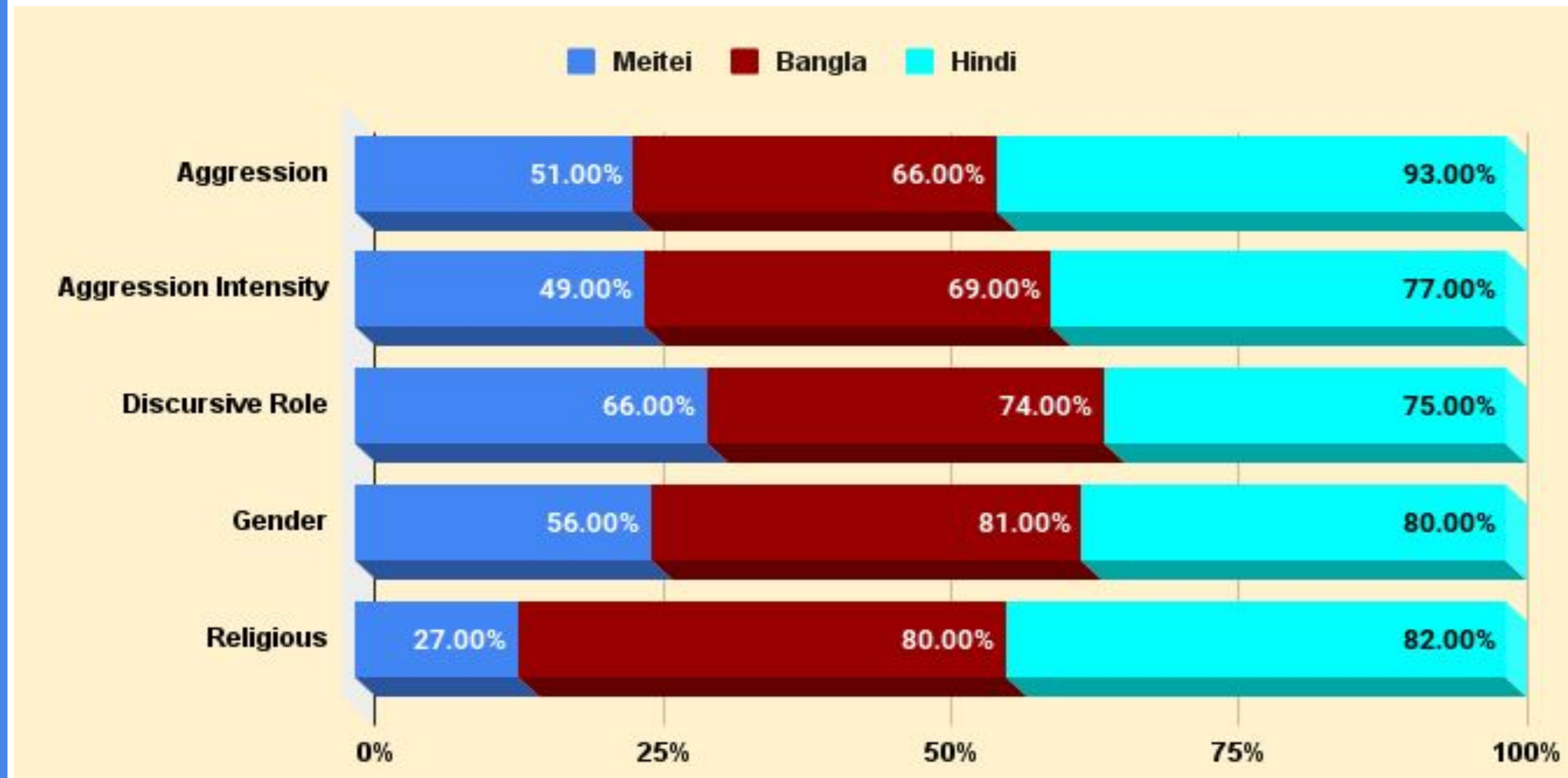
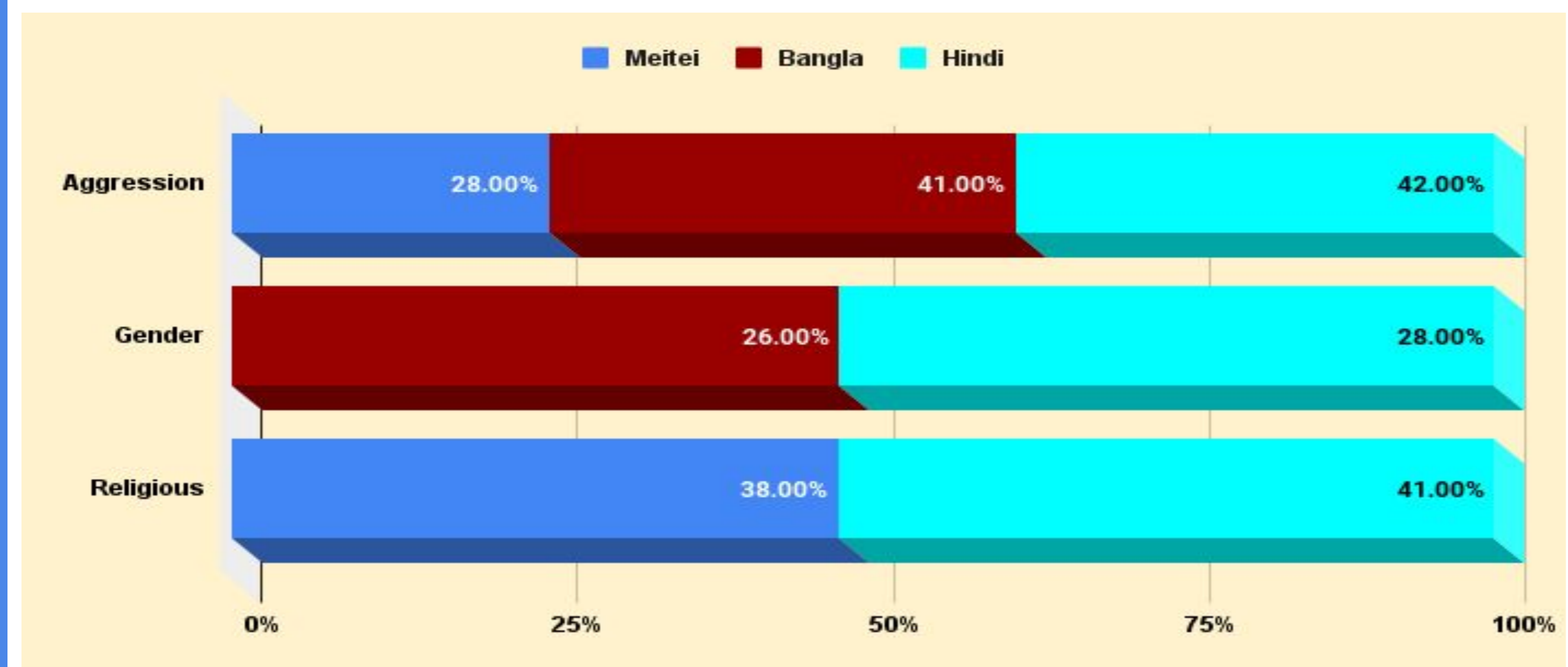
The ComM1 Project Tagset

Aggression	Aggression Level	<ul style="list-style-type: none"> Overtly Aggressive OAG Covertly Aggressive CAG Non Aggressive NAG
Aggression Intensity Level	Attribute	<ul style="list-style-type: none"> Physical Threat PTH Sexual Threat STH Non-threatening Aggression NtAG Curse/Abuse Aggression CuAG
Discursive Role	Attribute	<ul style="list-style-type: none"> Attack ATK Defend DFN Counterspeech CNS Abet and Instigate AIN Gaslighting GSL
The Gender Bias	Attribute	<ul style="list-style-type: none"> Gendered GEN Gendered Threats GENT Non-Gendered Comments NGEN
The Religious Bias	Attribute	<ul style="list-style-type: none"> Communal COM Communal Threats COMT Non-Communal NCOM
The Caste / Class Bias	Attribute	<ul style="list-style-type: none"> Casteist/Classist CAS Casteist/Classist Threats CAST Non-Casteist/Classist NCAS
The Ethnicity / Racial Bias	Attribute	<ul style="list-style-type: none"> Ethnic/Racial ETH Ethnic/Racial Threats ETHT Non-Ethnic/Racial NETH

Findings



Inter-annotator agreement Phase 1 & 2



Top 10 Most Frequent Words In The Dataset & Top 10 Frequent Words In The Dataset For Comments Which Are Either Aggressive/Communal/Gendered Or Combination Of These

