# Context-aware conversational models vs. sequence labeling for zero-shot cross-lingual hate speech detection in code-mixed social

Assignee Research

June 15, 2026

**Abstract**

In the current era of the internet, where social media platforms are easily accessible for everyone, people often have to deal with threats, identity attacks, hate, and bullying due to their association with a cast, creed, gender, religion, or even acceptance or rejection of a notion. Existing works in hate speech detection primarily focus on individual comment classification as a sequence labeling task and often fail to consider the context of the conversation. The context of a conversation often plays a substantial role when determining the author's intent and sentiment behind the tweet. Thi

## 1 Introduction

This paper examines: Leveraging Transformers for Hate Speech Detection in Conversational Code-Mixed Tweets. Research question: How do context-aware conversational models and sequence labeling approaches differ in zero-shot cross-lingual transfer accuracy for hate speech detection in code-mixed social media data?.

## 2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.8/10.

## 3 Results

8 papers retrieved. 8 claims extracted; 8 independently verified. Quality review score: 8.8/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## 5 Extracted Claims

| Claim | Verified | Confidence |
|---|---|---|
| The Soft Voting Ensemble approach achieved a score of 0.7682 in precision, 0.7687 in recall, 0.7684 in F1-score, and 76. | ✓ | 0.20 |
| The Hard Voting Ensemble approach achieved a score of 0.7621 in precision, 0.7628 in recall, 0.7624 in F1-score, and 76. | ✓ | 0.19 |
| The submit-2 Soft Voting Ensemble approach achieved a score of 0.7223 in precision, 0.7236 in recall, 0.7222 in F1-score | ✓ | 0.21 |
| The submit-3 Hard Voting Ensemble approach achieved a score of 0.7253 in precision, 0.7267 in recall, 0.7251 in F1-score | ✓ | 0.21 |
| The methodology involves fine-tuning transformer models pre-trained on a massive multilingual corpus. | ✓ | 0.20 |
| The data pre-processing step includes concatenating the tweet and its comments and replies, inserting a separator token | ✓ | 0.18 |
| Data cleaning involves removing hashtags, emojis, URLs, and mentions from the tweets while preserving punctuation and nu | ✓ | 0.22 |
| Hindi text in Roman script is transliterated to Devanagari script using the AI4Bharat library. | ✓ | 0.20 |

## References


- http://arxiv.org/abs/2112.09986v1
- http://arxiv.org/abs/2109.13711v1
- http://arxiv.org/abs/2101.03207v1