

Scaling Intermediate Tasks for Robust Zero-Shot Cross-Lingual Transfer in XTREME-R

Assignee Research

June 19, 2026

Abstract

Zero-shot cross-lingual transfer learning has been shown to be highly challenging for tasks involving a lot of linguistic specificities or when a cultural gap is present between languages, such as in hate speech detection. In this paper, we highlight this limitation for hate speech detection in several domains and languages using strict experimental settings. Then, we propose to train on multilingual auxiliary tasks – sentiment analysis, named entity recognition, and tasks relying on syntactic information – to improve zero-shot transfer of hate speech detection models across languages. We sh

1 Introduction

This paper examines: Multilingual Auxiliary Tasks Training: Bridging the Gap between Languages for Zero-Shot Transfer of Hate Speech Detection Models. Research question: What is the impact of scaling the number of intermediate training tasks on the robustness of zero-shot cross-lingual transfer performance across low-resource languages in the XTREME-R benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.1/10.

3 Results

14 papers retrieved. 11 claims extracted; 8 independently verified. Quality review score: 7.1/10.

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 study analyzes the training effect of adding different auxiliary tasks on top of XLM-R jointly with monolingual hate	×	0.09
Results are computed as deltas between the baseline system (no auxiliary task) and the augmented system.	×	0.04
Nozza (2021) demonstrates the limitation of cross-lingual transfer for domain-specific hate speech, particularly hate sp	✓	0.19
Nozza (2021) explains limitations in cross-lingual transfer by showing examples of cultural variation between languages.	✓	0.19
Some notable hate speech vocabulary in one language may be used as an intensifier in another language.	✓	0.15
Stappen et al. (2020) perform zero- and few-shot cross-lingual transfer on some of the datasets used in this paper using	✓	0.21
Stappen et al. (2020) do not distinguish between hate speech targets.	✓	0.20
The hate speech corpora used in this study are in three languages: English (en), Spanish (es), and Italian (it).	✓	0.17
The hate speech corpora cover two domains: hate speech towards immigrants and hate speech towards women.	✓	0.20
For English and Spanish, the study uses the dataset from the HatEval shared task on hate speech against immigrants and w	✓	0.22
For the Italian corpora, the study uses automatic misogyny identification data.	×	0.05

References

- <http://arxiv.org/abs/2210.13029v2>
- <http://arxiv.org/abs/2104.08645v2>
- <http://arxiv.org/abs/2605.10714v1>