Contrastive Loss Enhancement of CodeT5 Scaling in Low-Resource Languages on MBXP

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: Does applying contrastive losses to fused encoder features improve the scaling behavior of CodeT5 on low-resource languages within the MBXP dataset. 13 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: A Unified Contrastive Transfer Framework with Propagation Structure for Boosting Low-Resource Rumor Detection. Research question: Does applying contrastive losses to fused encoder features improve the scaling behavior of CodeT5 on low-resource languages within the MBXP dataset?.

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 3.7/10.

3 Results

14 papers retrieved. 13 claims extracted; 0 independently verified. Quality review score: 3.7/10.

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
There are no public benchmarks available for detecting low-resource rumors with propagation tree structure in tweets.	×	0.08
The datasets contain two binary labels: Rumor and Non-rumor.	×	0.05
The statistics of the six datasets are shown in Table 1.	×	0.03
EngCovid, ChiCovid, CanCovid and AraCovid denote the English-COVID19, Chinese-COVID19, Cantonese-COVID19 and Arabic-COVI	×	0.02
The model comparison results are shown in Table 2.	×	0.05
The model comparison results for different target datasets are shown in Table 3.	×	0.04
The model comparison results for different source datasets are shown in Table 4.	×	0.05
The methodology involves transforming each microblog post into a language-independent vector by semantically aligning th	×	0.04
The diffusion of rumors generally follows a propagation structure that provides valuable domain-invariant clues on how a	×	0.08
The methodology presents the conversation propagation thread as an undirected topology, which allows full-duplex interac	×	0.04
A multi-scale Graph Convolutional mechanism is used to catch informative patterns fused from both claim semantics and ev	×	0.11
A novel domain-adaptive contrastive learning paradigm is proposed to minimize the domain discrepancy.	×	0.14
The alignment dedicates identical rumor-indicative features from different domains closer, while the uniformity could he	×	0.07

References

- `http://arxiv.org/abs/2412.10008v1`

- http://arxiv.org/abs/2304.01492v5
- http://arxiv.org/abs/2506.09781v2