Diffusion-Based vs. Contrastive Code Models in Cross-Language Python-to-Java Transfer

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How do diffusion-based code models compare to contrastive learning methods in cross-language transfer from Python to Java on the MBPP dataset. 16 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Massively Parallel Cross-Lingual Learning in Low-Resource Target Language Translation. Research question: How do diffusion-based code models compare to contrastive learning methods in cross-language transfer from Python to Java on the MBPP dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.2/10.

3 Results

13 papers retrieved. 16 claims extracted; 2 independently verified. Quality review score: 4.2/10.

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 proposed model achieves 60.6% accuracy in qualitative evaluation for named entity translation.	✓	0.17
The Bible corpus is used as a test ground for the proposed extensions because it is the most translated text that exists	×	0.03
The Bible corpus includes both Old Testament and New Testament, unlike many past research works that only use the New Te	×	0.02
The model is trained on twenty-three European languages across eight families on a parallel Bible corpus.	×	0.08
Swedish is treated as the hypothetical low-resource target language, English as the rich-resource language in the single	✓	0.26
The training set contains 23K verses and is massively parallel.	×	0.05
The training, validation, and test sets are sampled according to the 0.75, 0.15, 0.10 ratio.	×	0.01
The control experiments use the WMT'14 French-English dataset together with French and English Bibles.	×	0.04
The WMT baseline contains French and English Bibles in addition to the WMT'14 data.	×	0.03
The experiments use a minibatch size of 64, dropout rate of 0.3, 4 RNN layers of size 1000, a word vector size of 600, a	×	0.03
For single-source single-target translation, 2 RNN layers of size 500, a word vector size of 500, and learning rate of 1	×	0.06
All learning rates are decaying.	×	0.02
The BLEU score for Swedish translation with one fifth of the training data is reasonably good.	×	0.15
An order-preserving lexicon translation method is devised by building a parallel lexicon table across twenty-three Europ	×	0.09
The BLEU scores for different language pairs are provided in the tables.	×	0.04
The BLEU scores vary with the amount of training data, as shown in the tables.	×	0.05

References

- http://arxiv.org/abs/1804.07878v2
- http://arxiv.org/abs/2508.09516v1
- http://arxiv.org/abs/2407.19619v1