

Performance comparison of projection-based cross-lingual NER models and XLM-R on MasakhanNER F1 scores

Assignee Research

July 19, 2026

Abstract

Cross-lingual Named Entity Recognition (NER) leverages knowledge transfer between languages to identify and classify named entities, making it particularly useful for low-resource languages. We show that the data-based cross-lingual transfer method is an effective technique for crosslingual NER and can outperform multilingual language models for low-resource languages. This paper introduces two key enhancements to the annotation projection step in cross-lingual NER for low-resource languages. First, we explore refining word alignments using back-translation to improve accuracy. Second, we pres

1 Introduction

This paper examines: Revisiting Projection-based Data Transfer for Cross-Lingual Named Entity Recognition in Low-Resource Languages. Research question: How does the performance of projection-based cross-lingual NER models compare to XLM-R when evaluated on the F1 score for named entity recognition in African low-resource languages using the MasakhanNER dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.5/10.

3 Results

12 papers retrieved. 10 claims extracted; 7 independently verified. Quality review score: 7.5/10.

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 evaluation was performed across a total of 57 languages using the XTREME (39 languages) and MasakhaNER2 datasets (18	✓	0.15
The XLM-R-Large model, fine-tuned on the English split of the CONLL2003, served as both the source model and for target	✓	0.20
The NLLB200-3.3B model was employed as a translation model for all experiments.	×	0.12
The evaluation involved full pipelines, considering both translation and source NER model performance.	×	0.12
The proposed approach involving n-gram candidates extraction provides comparable or superior results while offering grea	✓	0.21
The NER model-based extraction (N) consistently delivers strong performance.	×	0.11
Projection-based data transfer can outperform multilingual language models for low-resource languages.	✓	0.28
Data-based methods automate labelling through translation and annotation projection processes while leveraging advanceme	✓	0.30
Models trained in high-resource languages can identify and classify named entities in other languages without additional	✓	0.20
Categorization can be done through two approaches: translate-test, which labels original sentences in zero-shot settings	✓	0.34

References

- <http://arxiv.org/abs/2303.09306v2>

- <http://arxiv.org/abs/2501.18750v1>
- <http://arxiv.org/abs/1908.10261v1>