Robustness of Multi-Source Teacher-Student Learning to Domain Shift in Cross-Lingual NER Versus Pairwise Label Projection

Assignee Research

June 18, 2026

Abstract

To better tackle the named entity recognition (NER) problem on languages with little/no labeled data, cross-lingual NER must effectively leverage knowledge learned from source languages with rich labeled data. Previous works on cross-lingual NER are mostly based on label projection with pairwise texts or direct model transfer. However, such methods either are not applicable if the labeled data in the source languages is unavailable, or do not leverage information contained in unlabeled data in the target language. In this paper, we propose a teacher-student learning method to address such limi

1 Introduction

This paper examines: Single-/Multi-Source Cross-Lingual NER via Teacher-Student Learning on Unlabeled Data in Target Language. Research question: Is multi-source teacher-student learning on unlabeled data more robust to domain shift in cross-lingual NER than pairwise label projection methods?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.5/10.

3 Results

11 papers retrieved. 9 claims extracted; 8 independently verified. Quality review score: 8.5/10.

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
Previous works on cross-lingual NER are mostly based on label projection with pairwise texts or direct model transfer.	✓	0.34
Label projection methods are not applicable if labeled data in the source languages is unavailable.	✓	0.31
Direct model transfer methods do not leverage information contained in unlabeled data in the target language.	✓	0.34
The proposed method uses NER models in source languages as teachers to train a student model on unlabeled data in the ta	✓	0.36
The proposed method works for both single-source and multi-source cross-lingual NER.	✓	0.35
For multi-source cross-lingual NER, the authors propose a similarity measuring method to weight supervision from differe	✓	0.33
Experiments were conducted on 3 target languages using benchmark datasets.	×	0.11
The proposed method outperforms existing state-of-the-art methods for single-source cross-lingual NER.	✓	0.30
The proposed method outperforms existing state-of-the-art methods for multi-source cross-lingual NER.	✓	0.26

References

- https://doi.org/10.18653/v1/2020.repl4nlp-1
- https://doi.org/10.18653/v1/2020.acl-main.581
- https://doi.org/10.1609/aaai.v35i14.17512