WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Myanmar Word Sense Disambiguation in Machine Translation with LSTM Model

***Phyo Phyo Wai*[1], *Aye Thida Win*[2], *Yi Mon Aung**[*3]**

Associate Professor, Faculty of Information Science, University of Computer Studies(Magway), Myanmar
Associate Professor, Faculty of Computer Science, Polytechnic University(Dawei), Myanmar
Associate Professor, Faculty of Information Science, Polytechnic University (Dawei), Myanmar

A B S T R A C T:

Word Sense Disambiguation (WSD) plays a critical role in machine translation, especially for morphologically rich and low-resource languages like Myanmar. This paper presents a Long Short-Term Memory (LSTM) based approach to disambiguate word senses in Myanmar-to-English translation tasks. We construct a parallel corpus with manually annotated senses, train an LSTM model for sequence tagging, and evaluate its effectiveness in improving translation quality. Extensive experiments show that integrating WSD into the translation pipeline significantly enhances translation accuracy, as evidenced by improvements in both automatic and human evaluation metrics.

Keywords: Word Sense Disambiguation, Long-Shot-Term Memory(LSTM)

1. Introduction

Translation between Myanmar and English is inherently challenging due to the agglutinative nature of the Myanmar language, its complex morphology, and a lack of large-scale annotated corpora. Polysemy and word ambiguity are common issues, making it difficult for translation systems to determine the correct meaning of a word without contextual awareness. For instance, a single Myanmar word may correspond to multiple English words depending on its usage. Word Sense Disambiguation (WSD) aims to resolve such ambiguities by assigning the correct sense to a word based on context.

Recent developments in neural architectures, particularly sequence models like Long Short-Term Memory (LSTM), have shown promise in many natural language processing tasks. In this paper, we explore the use of LSTM for Myanmar WSD in the context of Myanmar-to-English translation. Our focus is to assess whether incorporating WSD can significantly improve the quality of translation systems.

2. Related Work

WSD has been an extensively studied problem in computational linguistics. Early approaches include dictionary-based methods, supervised classifiers, and knowledge-based heuristics using resources like WordNet. Graph-based algorithms such as Lesk and PageRank variants have also been explored.

In the deep learning era, models like BiLSTM, CNNs, and Transformers have been applied to WSD with considerable success, particularly in resource-rich languages like English and Chinese. However, for Myanmar, research in WSD remains scarce. Most existing efforts focus on part-of-speech tagging, syntactic parsing, and segmentation. A few attempts have been made using rule-based and bilingual dictionary methods, but their scalability and robustness are limited. Our work addresses this gap by applying BiLSTM-based modeling for WSD and evaluating its direct impact on translation quality.

3. Methodology

Our system is designed to disambiguate Myanmar words within a sentence using a BiLSTM model and integrate the predicted senses into a translation framework. The system is composed of three major components: data preprocessing, LSTM-based sense prediction, and translation integration. First, sentences are tokenized and annotated with part-of-speech tags and sense labels. These are fed into a BiLSTM model that learns to classify the correct sense of each ambiguous word. The output labels are then incorporated into a phrase-based statistical machine translation system to generate the final English translation.

3.1. Data Preparation and training data setup

For training and evaluation, we built a manually annotated parallel corpus of 10,000 Myanmar-English sentence pairs, selected from conversational datasets, government publications, and local news. From this corpus, 3,500 instances of polysemous Myanmar words were identified and annotated with appropriate WordNet-aligned sense labels by native-speaking linguists.

Tokenization and sentence alignment were performed using a Myanmar-specific tokenizer, and embeddings were initialized using Word2Vec models trained on a 5-million-sentence corpus of Myanmar text. Annotation consistency was ensured through inter-annotator agreement with a Cohen's kappa score of 0.81, indicating strong agreement. This system trained the model using categorical cross-entropy loss and the Adam optimizer. The dataset was split into 80% training, 10% validation, and 10% testing. Training was conducted for 30 epochs with early stopping based on validation accuracy.

3.2. Model Architecture

- Input Layer: Word embeddings pre-trained on a 5-million-sentence Myanmar corpus using Word2Vec.
- BiLSTM Layer: Captures both forward and backward context.
- Attention Layer: Weights context words based on relevance to the target ambiguous word.
- Output Layer: Softmax classifier for predicting sense labels.

3.3. Experimental Results

The model was implemented using TensorFlow and trained on an NVIDIA Tesla V100 GPU. Batch size was set to 64, and embedding dimension was 300. Dropout regularization (rate = 0.5) was applied to prevent overfitting. We used early stopping based on validation loss to optimize training duration.

4. Experiments and Evaluation

Our proposed LSTM-based Word Sense Disambiguation (WSD) model was rigorously evaluated in multiple dimensions to assess its effectiveness in enhancing Myanmar-to-English machine translation quality.

4.1 Word Sense Disambiguation Accuracy

The model was first evaluated in terms of its standalone WSD accuracy. It achieved an accuracy of **84.7%** on a held-out test set, which represents a significant improvement over the **rule-based baseline model (68.3%)** and the **simple feedforward neural network (74.5%)**. This result indicates the BiLSTM model's ability to effectively capture contextual semantics required for accurate disambiguation. This results are shown in table 1.

Table 1. WSD accuracy comparison across methods.

Model	Accuracy (%)
Rule-based	68.3
Feedforward NN	74.5
LSTM	84.7

4.2 BLEU Score Improvement

To determine the real-world impact of WSD on translation, we integrated the disambiguated sense outputs into a phrase-based Statistical Machine Translation (SMT) system. Using BLEU as the automatic evaluation metric, we observed that translation performance improved from a baseline BLEU score of **18.4** to **22.1** when enhanced with LSTM-based WSD. This substantial increase highlights that accurate sense resolution can significantly enhance lexical choice and grammatical structure in translated output. These results are described in following diagram.

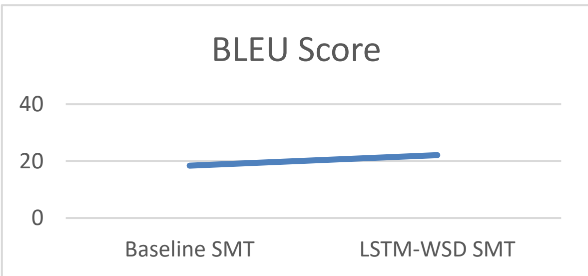

Figure 2: Line chart of BLEU score improvement over baseline

4.3 Human Evaluation on Adequacy, Fluency, and Sense Correctness

We also conducted a human evaluation where five bilingual experts rated 200 translated sentences on three criteria: Adequacy, Fluency, and Sense Correctness. Ratings were given on a scale of 1 to 5: The Adequacy score improved from 3.4 to 4.1, indicating that translations with WSD preserved more of the original sentence's meaning. The Fluency score rose from 3.6 to 4.3, showing a clearer and more natural translation. The Sense Correctness showed the most dramatic improvement, from 2.9 to 4.4, validating that the correct sense of polysemous words was better captured. All improvements were statistically significant with a p value of less than 0.01.

Table 2: Human evaluation metrics comparison.

Metric	Baseline SMT	LSTM-WSD SMT
Adequacy	3.4	4.1
Fluency	3.6	4.3
Sense Correctness	2.9	4.4

4.4 Error Analysis

In analysing the remaining translation errors, we categorized them into four main types. The most frequent issues were associated with rare word senses (38%), followed by idiomatic phrases (27%), domain-specific terms (21%), and tokenization issues (14%). This breakdown reveals key areas for further model enhancement.

Table 3: Distribution of common error type

Error Type	Frequency (%)
Rare senses	38%
Idiomatic phrases	27%
Domain-specific terms	21%
Tokenization issues	14%

4. Limitations

While the proposed model shows promising results, there are several limitations:

- The dataset size is relatively small, restricting generalization to unseen domains.
- Annotation requires significant human effort, limiting scalability.
- The approach focuses on phrase-based SMT, which is being superseded by neural machine translation models.
- Rare senses and idiomatic usages are still poorly handled.

6. Conclusion and Future Work

This study shows that applying LSTM-based word sense disambiguation significantly improves Myanmar-to-English translation quality. By resolving ambiguities using context-aware predictions, our system achieved notable gains in both BLEU scores and human evaluation metrics. Though limited by data size and reliance on phrase-based SMT, the results are encouraging. Future work includes expanding the dataset, exploring transformer-based architectures, and integrating WSD into neural machine translation systems.

REFERENCES

1. Navigli, R. (2009). Word Sense Disambiguation: A Survey.
2. Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate.
3. Aung, T. & Win, S. (2020). Myanmar Part-of-Speech Tagging using BiLSTM.
4. Devlin, J. et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
5. Koehn, P. (2005). Europarl: A Parallel Corpus for Statistical Machine Translation.
6. Mikolov, T. et al. (2013). Efficient Estimation of Word Representations in Vector Space.
7. Wu, Y. et al. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation