TartuNLP @ AXOLOTL-24: Leveraging Classifier Output for New Sense Detection in Lexical Semantics

Aleksei Dorkin [1] Kairit Sirts [1]

[1]*University of Tartu, Tartu, Estonia*

Methodology

Adapters-Based Binary Classification

- Adopted the GlossBERT approach with some modifications;
- The system is a binary classification model that matches usage examples with glosses;
- The model is a cross-encoder that processes usage examples and sense definitions simultaneously, predicting probabilities of a match for each pair;
- If all definitions for a word in a usage examples have low probability, then we assume this to be the example of a novel sense;
- Fine-tuned XLM-RoBERTa individually for each language using bottleneck adapters;
- Positive examples are pairs where the gloss correctly defines the meaning of the word in the usage example, while negative examples are mismatched pairs.

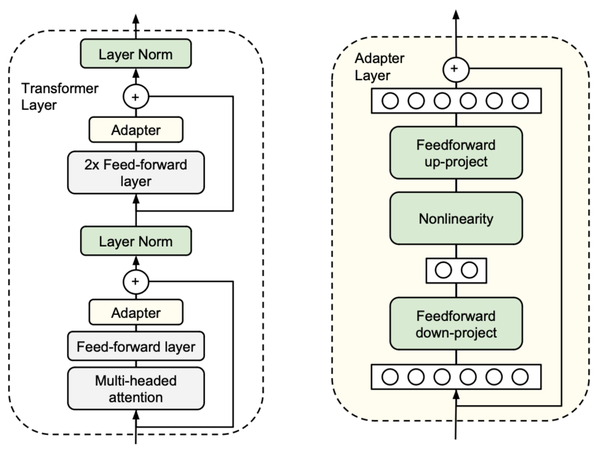

Figure 1. An illustration of the Bottleneck Adapter. The left side demonstrates how a bottleneck adapter is added to a single transformer layer, while the structure of an individual adapter layer is on the right. Only elements in green are trained, while the rest remains frozen.

Cross-Lingual Knowledge Transfer

- No training data for German;
- Used the *old* period from the test data for training;
- Continued training from the checkpoint trained on Finnish.

Data Handling and Preprocessing

Crafting Training Examples

- Positive examples are available in the provided data;
- Negative examples necessary for classification are obtained by sampling the other definitions of the same word;
- Negative examples are deemed to be hard (at least in case of polysemy, where different senses are expected to be somewhat related).

TL;DR

- First place on *Subtask 2: Definition Generation for Novel Word Senses*;
- Not generating anything;
- Sentence pair classification inspired by GlossBERT;
- Binary classification of gloss/usage example pairs;
- Each usage is paired with all available definitions;
- If the *classifier probability* of all pairs is low, then this is a new sense;
- Identified new senses are matched against Wiktionary definitions for the same word using the same model.

Analysis

Strengths

- Simple and straightforward approach to the solution;
- Interpretable and intuitive;
- Low computational resource requirements;
- Strong performance in novel sense definition matching.

Challenges

- Variability in quality of definitions both in the provided data and Wiktionary;
- Variability in the amount of content in different versions of Wiktionary;
- Variability in time periods considered old and new;
- BERTScore might be misleading due to the lack of re-scaling;
- BLEU is limited in capturing nuanced semantic changes.

Results

Subtask 1: Bridging Diachronic Word Uses and a Synchronic Dictionary

- Achieved competitive scores in F1 and ARI;
- Binary classification model to match glosses and usage examples;
- Classification model's probability is used to identify novel senses.

Team	ARI	F1
deep-change	**0.413**	**0.750**
Holotniekat	0.312	0.641
TartuNLP (ours)	0.310	0.590
IMS_Stuttgart	0.287	0.487
ABDN-NLP	0.221	0.431
WooperNLP	0.187	0.316
Baseline	0.041	0.207

Table 1. Overall results on the Subtask 1.

Subtask 2: Definition Generation for Novel Word Senses

- First place in the subtask;
- System matched identified novel sense examples with definitions sourced from Wiktionary;
- Reused the same model from the first subtask.

Team	Overall	BLEU	BERTScore
TartuNLP (ours)	**0.467**	**0.208**	**0.726**
WooperNLP	0.340	0.020	0.660
ABDN-NLP	0.253	0.045	0.461
baseline	0.218	0.013	0.423

Table 2. Overall results on the Subtask 2.

Limitations

- Novel sense definition matching is limited to a given word's sense inventory;
- The approach itself is dependent on the availability of external lexical resources;
- High score in Russian partially caused by accidentally matching the source of definitions.

Future Directions

- Explore bi-encoder architectures for retrieval style definition matching;
- Refine threshold determination for new sense identification.