A Few-shot Learning Approach for Lexical Semantic Change Using GPT-4

Zhengfei Ren   Annalina Caputo   Gareth J.F. Jones

School of Computing, Dublin City University, Ireland

CRT·AI

SFI Centre for Research Training
in Artificial Intelligence

Introduction

- We report a systematic evaluation of a Large Lange Modelling (LLM) method for a Lexical Semantic Change Detection (LSCD) task is required.
- We examine the few-shot learning ability of LLMs with hand-picked or machine selected examples using GPT-4 to for LSCD.

Our Approach

Basic Prompt

Consider the use of target word in two contexts of sentences, determine whether the target word has changed its semantic meaning between those sentences. Do they refer to the Same, different but Related, distantly Linked or unrelated objects.

[demonstration examples here]

#1. [sentence1]
#2. [sentence2]

Answer: choose from (same, related, linked, distinct) without any further explanation.

GPT - 4

Semantic
Change Answer

Prompt Engineering

- Zero-shot, one-shot and few-shot learning models with different number of examples.
- Strategies for selecting Language Change Demonstrations:
  - Manually selected
    - Select several typical semantic change examples with related or similar context sentence pairs to show the change detection results.
  - Machine selection
    - Retrieve similar semantic change example pairs with highest cosine similarity from the sample dataset computed by cosine similarity of the BERT embedding representations. Examples retrieved from a sample dataset from the ChiWUG dataset.

ChiWUG dataset

- Our dataset used is based on the ChiWUG, a Chinese versions dataset following the DWUG framework.
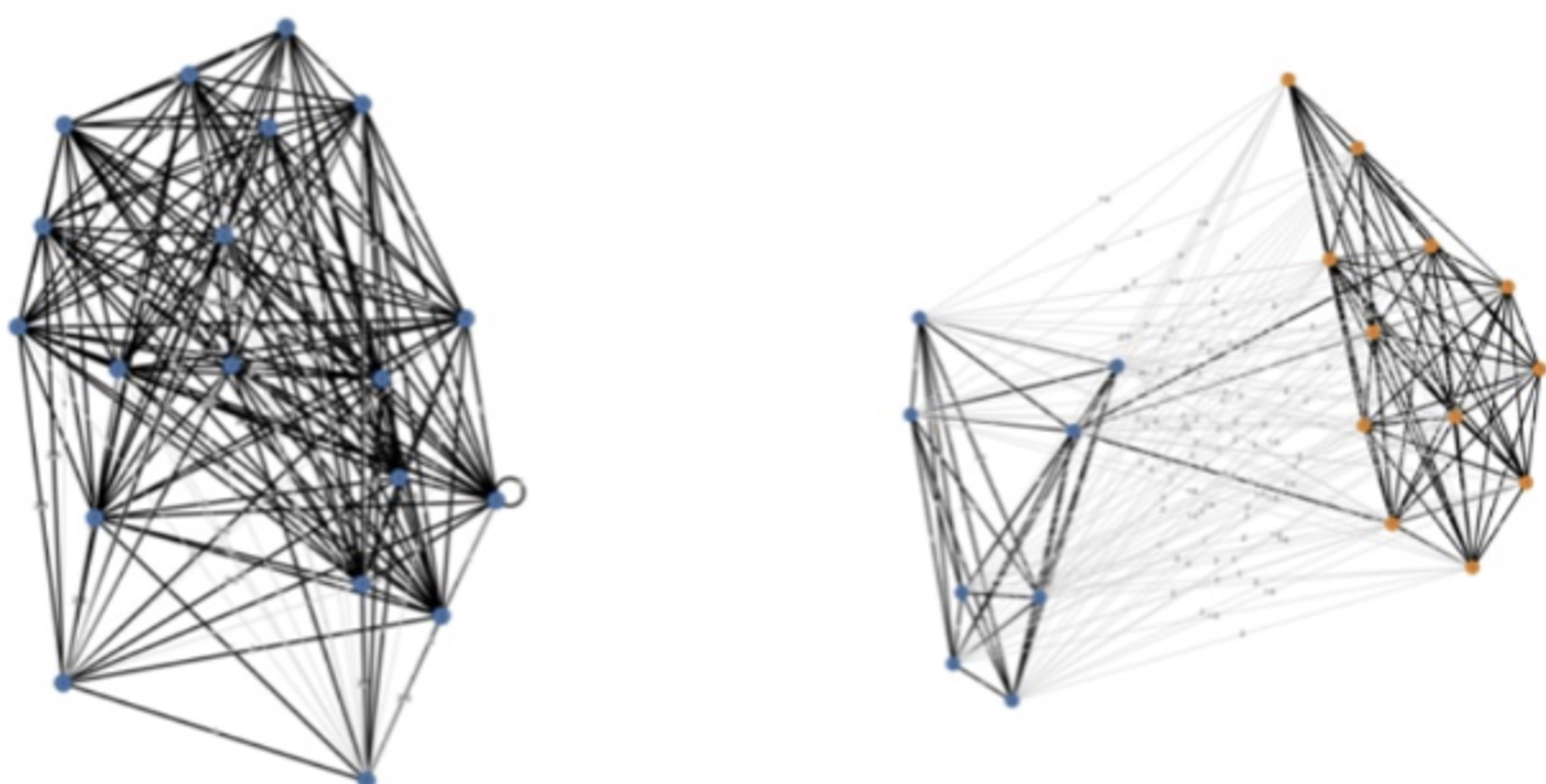

Figure 1 DWUG for the target word, 下海 xiahai, nodes and edges represent the word usage and relations between them, semantic change is represented by the four degree of changes from 1 to 4 (unrelated to same).

Hand-picked Example

- Example of hand-picked for target word 下海 xiahai (go into sea or start business) demonstrating the semantic change *Related*.

1. 原本在大学教授生物学的她，决定下海创办了一家生物学科技公司

   A professor of biology in a university, decided to set up a biotechnology company.

2. 她曾是一名成功的时尚设计师，后来选择下海，开设了自己的时装品牌

   She was as a successful fashion designer, before she chose to go to business and start her own fashion brand.

Results & Discussion

- Results show the different models with zero-shot, one-shot and few-shot learning with two or three examples.
- The one-shot Retrieval model uses example selected by the machine.

Approach	Binary Change	GCD
XL-LEXEME	/	0.73
Zero-shot	0.70	0.65
One-shot	0.70	0.72
Two-shot	0.65	0.73
Three-shot	**0.83**	**0.79**
One-shot Retrieval	0.70	0.70

Table 1 GCD and binary change predictions with different zero-shot or few-shot GPT-4 models. XL-LEXEME is the previous best-performing model on the ChiWUG evaluation dataset.

1. Result shows that more examples in the prompt lead to higher performance

   -> three-shot learning performs best among our systems.

2. One-shot learning with or without machine selected examples shows similar results. There is no indication that example retrieval provides any further improvement in performance.

Conclusions and Further Work

- We show that our three shot learning performs well on the LSCD task, and outperforms previous best-performing pre-trained model XL-LEXEME.
- Further exploration is needed the potential for different combinations of manually selected or machine selected examples to improve performance.
- Further study should also include applying other prompt engineering techniques to increase language change detection performance.