Japanese Word Sense Disambiguation Using Gloss Information of a Japanese Dictionary

Article authors: Hiroki Okemoto, Minoru Sasaki

Presenter's name and affiliation:

Hiroki Okemoto, Major in Computer and Information Sciences Graduate School of Science and Engineering, Ibaraki University

e-mail: 21nm716g@vc.ibaraki.ac.jp

A short resume of the presenter

➢Hiroki Okemoto graduated from Ibaraki University.

➢He is currently a Master's student in Ibaraki University.

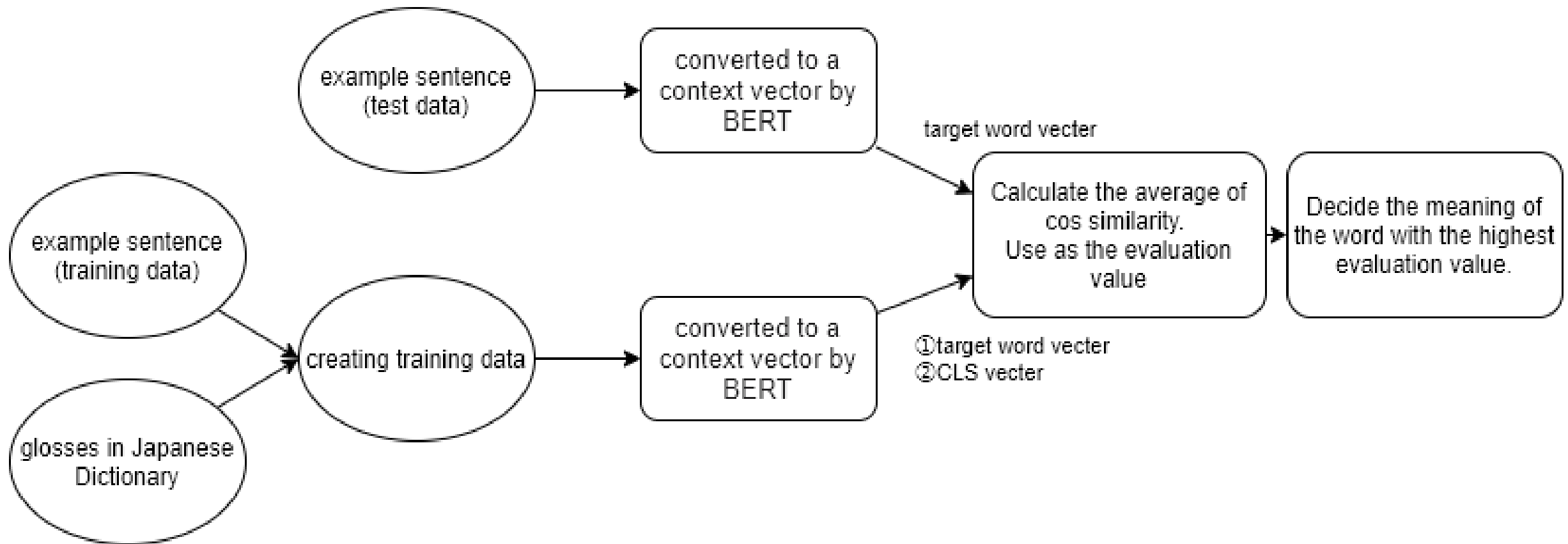➢His research interests include natural language processing.

Outline

1. Purpose of Research

2. System Overview

3. Input Data

4. Experimental Methods

5. Baseline Method

6. Results

7. Discussion

8. Conclusion and Future Work

1. Purpose of Research

➤ Analysis of the effectiveness of Japanese WSD using the gloss information of Japanese dictionary.

  ➤ There is no research that has used it effectively with Deep Learning methods.

➤ Improving the data volume of training data by combining gloss information and example sentences.

  ➤ Using Japanese evaluation data to analyze effectiveness

2. System Overview

example sentence (test data) → converted to a context vector by BERT → target word vecter →

example sentence (training data), glosses in Japanese Dictionary → creating training data → converted to a context vector by BERT → ①target word vecter ②CLS vecter →

Calculate the average of cos similarity. Use as the evaluation value → Decide the meaning of the word with the highest evaluation value.

3. Training Data

➢When creating training data, extract glosses and example sentences from dictionary definitions.

> Headword あう【合う・会う・遭う・遇う・逢う】
>
> ⋮
>
> 166-0-1-3-2 <イ>一致する。「意見が—」。理にかなう。「答えが—」
>
> ⋮

➢glosses　　　　　　　→ "一致する。", "理にかなう。"

➢example sentences　　→ "意見が合う", "答えが合う"

　➢The "-" part is completed by the Headword.

3. Training Data

➤Input data format

"gloss「example sentence」"
  ➤gloss→dictionary
  ➤Example sentence→dictionary, training data

➤e.g.）"偶然（人や物に）出会う。「駅でばったり知人と会った」"

➤The table on the right shows how to separate the example sentences

➤Input data is converted to context vector using NWJC-BERT

	How to separate example sentences
A	Not separate
B	Separate by punctuation. And the part that contains the target word.
C	A total of 7 words, including 3 words before and after the target word
D	A total of 11 words, including 5 words before and after the target word
E	A total of 15 words, including 7 words before and after the target word
F	A total of 19 words, including 9 words before and after the target word

4.Experimental Methods

Comparison of test data and training data (two ways)

● CLS vector and target word vector

● target word vector

➢ We compared test data with the training data, determined the evaluation value, determined the meaning of the training data which became the maximum value as the meaning of the test data.
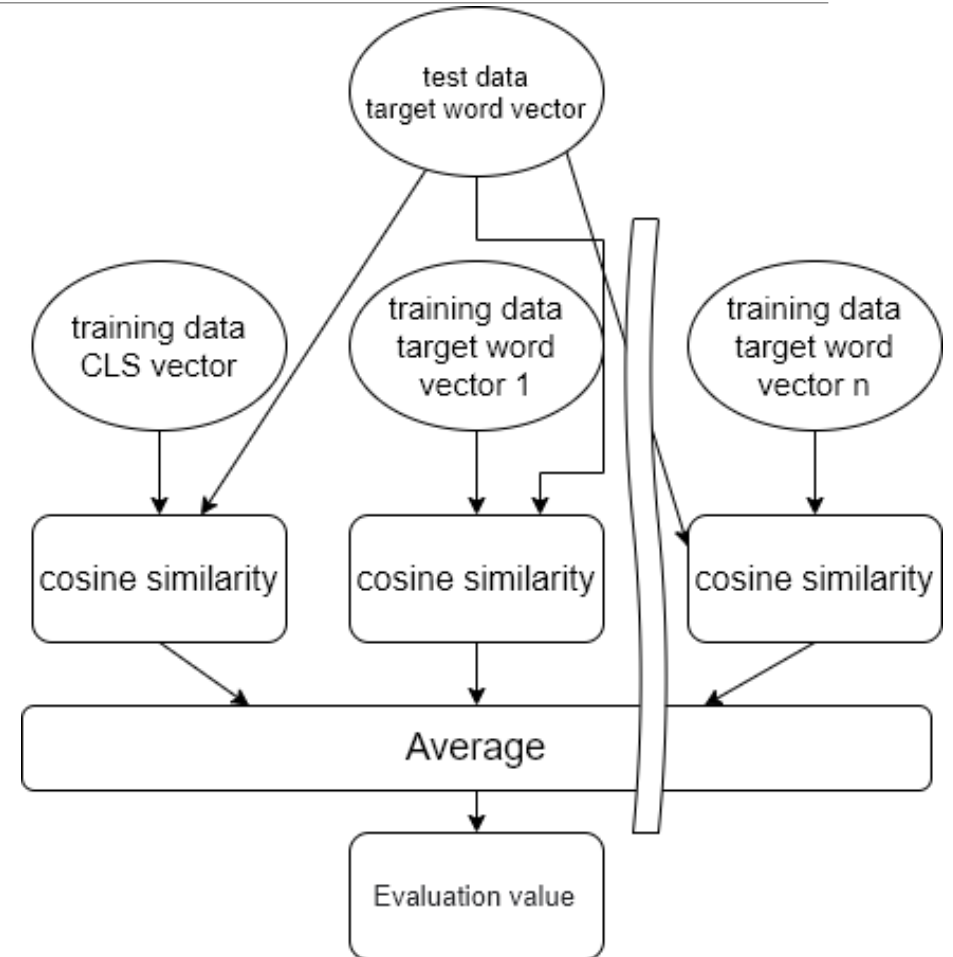
4.Experimental Methods

Evaluation Methods

➢Using SemEval2010 Japanese WSD task data

➢There are 50 target words.

➢There are 50 test data and 50 training data per word.

➢Find the average of the percentage of correct answers for 50 words of data.

5. Baseline Methods

➢Using a supervised learning approach without using gloss information of a Japanese dictionary

➢Using a Multi-Layer Perceptron (MLP) as a classification model.

6. Results

Gloss information	Case	CLS vec + Target word vec	Target word vec
Used	A	78.88%	78.32%
	B	**79.28%**	79.16%
	C	78.36%	79.24%
	D	78.52%	79.16%
	E	79.20%	**79.68%**
	F	78.28%	79.44%
Unused	A	78.68%	79.00%

Baseline method	Correct answer rate
MLP	64%

7. Discussion

➢Experimental results show that the using gloss information of Japanese dictionaries is effective for WSD.

➢Although the amount of data in the training data has been improved, there are some words with extremely small data depending on their meanings, such as those with a small number of both example sentences and definition sentences.

➢Depending on the evaluation method, there was a difference in the position of the separation of example sentences that resulted in the maximum percentage of correct answers.

➢CLS vec + Target word vec　　▪▪▪Case B

➢Target word vec　　　　　　　▪▪▪Case E

8. Conclusion and Future Work

➢We have shown the effectiveness of Japanese WSD using gloss information of Japanese dictionary.

➢Future Work
  ➢Review of the evaluation method to determine the meaning of words such as MLP.
  ➢Examination of word sense judgment models other than cosine similarity.
  ➢Increase the data volume of glosses and example sentences.
  ➢Revise the position of separation of example sentences combined with glossary.