Bi-LSTM based metaphor Identification and Interpretation in context

Z.M. Tian [1,*], Q.Z. Wang [1], Z.H. Huang[1]

Supervisor: Prof. W.F. Su

[1] Computer Science and Technology, Division of Science and Technology, BNU-HKBU United International College

** Corresponding student author. Tel: +86-13172650285, E-mail: m730026102@mail.uic.edu.hk*

Abstract

- Metaphor expressions appear broadly in daily life, furnishing vivid and concrete explanations for abstract experience and perception. In light of linguistic theories, a metaphor is identified if the literal meaning of a word contrasts with the meaning that word takes in this context. The computational realization of metaphor identification and interpretation is hence a crucial segment in the NLP translation field. Current word embedding based metaphor identification models may perform a decent performance, whereas the considering aspect of metaphor detection may be single. In this study, we propose to use a Bi-LSTM model. It consists of domain representation of each word from pre-training in big language datasets, which helps to figure fine tone of the model.
- Our model have good performance on some popular metaphor dataset by using GloVe word vector to embedding and training on an Bi-LSTM based model. We also implemented self-attention layer to increase our accuracy.
- Our model adopts three widely used metaphor datasets. U Amsterdam Metaphor Corpus (VUA): the largest publicly available metaphor dataset, MOH-X: sentences from WordNet, and TroFi: consists of sentences from the 1987-89Wall Street Journal Corpus. We evaluate the model with a plethora of data, performing that our model has a decent result.

Introduction

- Formerly, the majority of work on the metaphor processing hinged on limited linguistic forms. For instance, given fix SVO (Subject-Verb-Object) sentence structure "She devoured novels" to the model. In daily condition, we notice that long and sophisticated sentences may provide crucial hints for comprehending the metaphor in a sentence, only considering the main sentence trunk may lead to bias.

1. Rockford teachers are honored for saving a drowning student.
2. You're drowning in student loan debt.

Figure 1. "drowning" in 1 means "die from being submerged in water", "drowning" in 2 means "get into trouble

- Our task is that for a given target verb in a sentence, classify it if the verb is metaphorical or not, interpreting it into the fittest word. We propose a metaphor processing model with Bi-LSTM + GloVe. Specifically, our model is built upon 7.5 F1 on the VU Amsterdam Metaphor Corpus (VUA) for the sequence labelling task, by 2.5 F1 on the VUA verb classification dataset, and by 4.9 F1 on the MOH-X dataset.

Bi-LSTM

- Bidirectional Long Short Term Memory (Bi-LSTM), which is a combination of forward LSTM and backward LSTM is now widely used in the tasks in natural language processing. Using LSTM on understanding sentences, it can better capture the long-distance dependencies since LSTM is capable to learn to memorize or forget the specific information during the training process.

GloVe

- Global Vectors for Word Representation (GloVe) is an unsupervised learning algorithm for obtaining vector representations for words. Training is performed on aggregated global word-word co-occurrence statistics from a corpus, and the resulting representations showcase interesting linear substructures of the word vector space.
- It gives the highlights on "nearest neighbors" of two words, measuring the linguistic or semantic similarity of the corresponding words by Euclidean distance between two-word vectors. It uses "Linear substructures" to maintain the similarity metrics used for nearest neighbor evaluations produce a single scalar that quantifies the relatedness of two words.

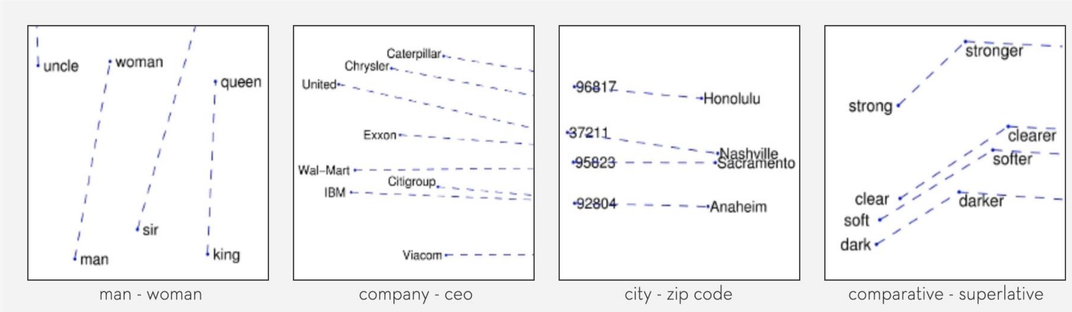

Figure 2. Particle vectors representation in space.

Methodology and Data

- For detection, according to the format of the data set, it provides a sentence and a target word, and label of the target word. Our model is designed as a classification model. Unlike Chinese sentences, we need to do word segmentation. First, we can directly use a 300d glove vector for word embedding, then, use a Bi-LSTM to encode a sentence, producing a contextualized representation, A self-attention layer is added to calculate the weight of target word. Finally, a full connection layer and SoftMax are used to predict the target word 10-fold cross-validation will be used to train our model. We divided the dataset into 10 part, nine of them will be the training set, one of them will be the test set, and repeated train for ten times so that every part of the dataset will be taken as a test set.
- For interpretation, link to an external knowledge base like The Lex-Ecologist or CogBank which contain the attribute of a word, collect the possible meanings of the metaphorical words(target word), put them back into the original text, and compare them with contextual predictions. And choose the fittest literal replacement of the metaphorical word to the model.

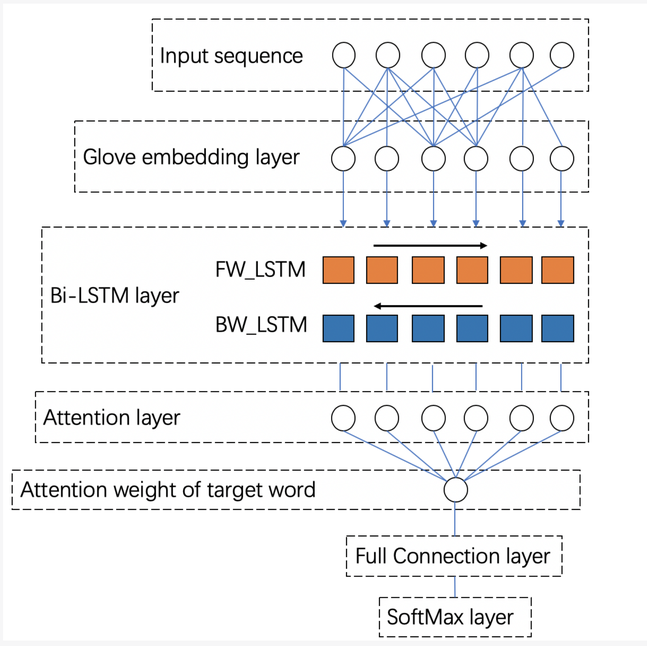

Figure 3. our classification model for metaphor detection.

- We averaged the results of 10 training result and shows that our f1 score reached 75%, though there is still some gap (about 4%) from the state-of-the-art advanced model. Since our model is relatively simple, and we can implement some new approaches or components to improve our model's accuracy in the future.

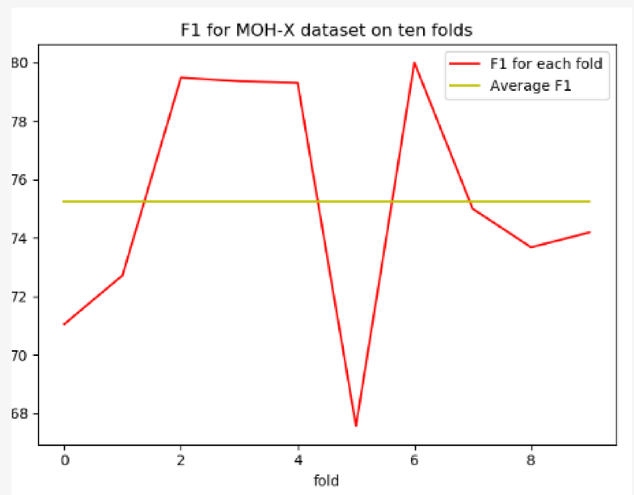

Figure 4. F1 score during 10-fold cross-validation

Related Work

- For computational metaphor processing, enormous methods have been exploited, which provides us with enough ideas and references. Some researchers classified metaphorical and literal words through the word concreteness. Some other researchers use word embedding to classify metaphor. They measured the value of cosine similarity of phrases containing the target domain and their input vector of Skip-gram word embedding. If the value is over a threshold, then the expressions can be considered as metaphorical.

Conclusion

- We developed a Bi-LSTM based model for detecting metaphorical word in context, which can learn from context that contributes much in figurative language. Over model base on the state-of-the-art verb metaphor detection model, give a great performance on the results. Our model still requires more improvement on detection and interpretation on metaphor.