Relation Extraction using Neural Network

Ankur Debnath
M.Tech (EE)
ankurdebnath

Danish Shaikh
M.Tech (ECE)
shaikhm

Rajat Nagpal
M.Tech (SE)
rajatnagpal

Abstract

Relation extraction plays an important role in extracting structured information from unstructured sources such as raw text. One may want to find interactions between drugs to build a medical database, understand the scenes in images, or extract relationships among people to build an easily searchable knowledge base.

1 Introduction and Problem Statement

Many tasks in natural language understanding require an understanding of the semantic relations between entities. Relation extraction (RE) is defined as the task of extracting semantic relations between given entity pairs from plain text. For example, in the sentence "**People** have been moving back into **downtown**", the entities **[People]** and **[downtown]** are in a Entity-Destination relationship. There is a considerable interest in automatic relation classification, both as an end in itself and as an intermediate step for many NLP applications such as Information extraction, Document summarization and automatic enrichment of existing knowledge bases.

Supervised methods based on neural networks have been found successful for relation extraction tasks. But such methods require largescale manually-constructed corpus, which is expensive to create. Recently, distant supervision has gained a lot of attention which automatically produce training corpus. In distant supervision, we simply label all the sentences containing entity e1 and e2 by their corresponding relation present in the knowledge base.

Given an entity pair (e',e") from a knowledge base such as Freebase, assuming that the predefined semantic relation on the KB is r, we simply label all sentences containing the two entities by label r. However, it has a major shortcoming that the distant supervision assumption is too strong and may cause the wrong label problem. It is possible that these two entities may simply share the same topic or maybe representing some other relation than that present in knowledge-base. In order to address this problem, proposed to model distant supervision as Multi-Instance Multi-Label (MIML) problem. It assumes that in a set of sentences corresponding to an entity pair, at least one sentence in that set should express the true relation assigned to the set. In a more recent work, proposed sentence attention model to deal with the challenges associated with distant supervision.

Conventionally, the task of relation extraction is divided into two distinct parts. First, is to extract candidate sentences and entity pairs from a plain text between whom we want to identify the relations and the second task is the relation classification task. In this work, our main focus is to perform relation extraction under supervised settings.

2 Related Work

As an important and fundamental task in NLP, relation extraction has been studied extensively. Initially, handcrafted features were extracted using pre-existing tools but that lead to propagation of errors in the existing tools and hinders the performance of the system. Further, kernal-based methods were also introduced which automatically generate features.

Recently, the neural network models have dominated the work of Relation Extraction because of higher performances. Zeng et al., 2014 used a convolutional deep neural network (DNN) to extract lexical and sentence level features. This work also proposed the inclusion of position features for the task of relation extraction. These two levels of features are concatenated to form the final extracted feature vector. Softmax classier is used to predict the relationship between two marked entity. In an extension to this, Yatian Shen, Xuanjing Huang, 2016 proposed a attentionbased convolutional neural network architecture for this task that makes full use of word embedding, part-of-speech tag embedding and position embedding information. Further to incorporate distant supervision, Zeng et al., 2015 combined the multi-instance learning with piecewise convolutional neural networks to learn more relevant features. Lin et al., 2016 employed CNN with sentence-level attention over multiple instances to encode the semantics of sentences. Miwa and Bansal, 2016 used a syntax-tree-based long short-term memory networks (LSTMs) on the sentence sequences. Further, Sharmistha Jat et al., 2017 proposed a weighted ensemble model of a BiGRU-based word attention model and Entity Attention (EA), an entity-centric attention model. The current stateof- art model for relation extraction is proposed in Zhengqiu He et al., 2018 centered on the ideas of using tree-GRU based syntax aware embeddings.

3 Solution Sketch

- Word Embedding.
- Position Embedding.
- Divide the sentence in n-parts (hyperparameter) and perform pooling in each parts and later concatenate.
- Calculate attention vector for each word by learning attention weights.
- Calculate the output after passing through the non-linear function (say softmax).

4 Datasets and metrics identified

For this project, SemEval-2010 Task 8 benchmark dataset "Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid O. S eaghdha, Sebastian Pad o, Marco Pennacchiotti, Lorenza Romano, and Stan Szpakowicz. 2010, "Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals" In Proceedings of the 5th International Workshop on Semantic Evaluation, SemEval 10, pages 3338.' is proposed to be used for experimentation. On this dataset, "Yatian Shen, Xuanjing Huang, "Attention-Based Convolutional Neural Network for Semantic Relation Extraction", in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics." is the present state-of-the-art with F1 score of 85.9.

5 Baseline

Our first attempt is to make an ensemble model. Here, both the models (PCNN and Bi-GRU) independently compute the scores for each class which we get by passing the respective sentence representations through a dense layer, and we combine the two models by taking a weighted sum of the output scores. Now the softmax layer is applied to these ensembelled output scores to determine the relation in the sentence.

6 Experiment Hypothesis

1. Concatenate Model
2. Sequential Model