NLPCC 2017 Shared Task Guideline:

Chinese Word Semantic Relation Classification

1. Task

The automatic recognition of word semantic relations is a challenging task for natural language processing. Words with different semantic relations are often mixed together when doing lexical sematic computation, but we need distinguish between them in real-world applications. Our goal is to create a testbed for automatic classification of word semantic relations. Given some pairs of Chinses words, the system is asked to classify them into the four categories: synonym, antonym, hyponym and meronym. The dataset provides the four categories of semantic relations, and each category contains 500 instances of word pairs.

All kinds of strategies are welcome, including the traditional corpus-based distributional method, dictionary-based semantic computation, as well as the recently developed word embedding methods and deep learning strategies. Also, the participating system is encouraged to use any external resources.

2. Data

In this task, we will only provide a test data. Some examples are given in Table 1.

Table 1. Some examples in the test data

Word 1	word 2	relation
花	菊花	hyponym
花	花蕾	meronym
花	消费	synonym
笔	钢笔	hyponym
钢笔	笔帽	meronym
骄傲	自豪	synonym
骄傲	谦虚	antonym
色彩缤纷	姹紫嫣红	synonym
目不转睛	东张西望	antonym

The dataset is constructed by an expert in computational linguistics. The selected words consist of nouns, verbs and adjectives, as well as some functional words like adverbs and conjunctions. The selected words are variable-length, including single-character words, two-character words, three-character words and four-character words (idioms). We also pick up some ambiguous words with multiple senses. All the words are common words, avoiding named entities. This distinguishes our task from much work in information extraction, which tends to focus on specific classes of named entities and on more fine-grained relations.

The four categories of semantic relations are the most general ones in lexical semantics, which can be defined as follows. Give two words A and B:

Synonym: A is similar to B, e.g., 色彩缤纷 and 姹紫嫣红.

Antonym: A is contrast to B, e.g., 骄傲 and 谦虚.

Hyponym: B is a kind of A, e.g., 菊花 and 花.

Meronym：B is a part of A, e.g., 花蕾 and 花.

3. Evaluation

In order to avoid over-fitting on this small test dataset, we will release a large collection of 10,000 word pairs in the testing phase. Our 2,000 word pairs that are constructed by human will be distributed randomly in the large data, and the other word pairs were automatically generated using a large dictionary.

We will compute precision (P), recall (R), and F1-Score for each relation, and will compute the macro averaged P, R and F1. Our official scoring metric is macro-averaged F1-Score for four-way classification.