Supplementary Materials

June 23, 2021

This file gives extra details of experiments in the paper.

> **Error correction** In the submitted paper, we modified the name of a compared method. But there are some places where the name has not been changed. The method named "Bert" should be "Ditto". Sorry about that!

1   Data sets

We crawled 1742 papers from Bing, DBLP and IEEE according to the keywords "Xindong Wu", "Fei Jie" and "Yanxiang Huang'. Then the paper was labeled manually. Two papers constitute an example whose label indicates the two papers refer to the same one or not. To avoid imbalance in examples, we sampled in negative examples and than split the extended examples into training set and test set. The ratio between positive and negative examples is 1:4. The Tab. 1 gives some statistics about datasets.

Table 1: Some statistics of datasets.

Dataset	Size	Pos.	Attributes
All	3212	586	5
Train	2248	405	5
Test	964	181	5

where "Pos." means the number of positive examples.

2   The experiment settings for different serialization strategies

The parameter settings in the experiments of serialization strategies are given in Tab. 2. The model using different serialization strategies have the same parameter settings.

Table 2: The parameters in the experiments.

Epoch	Batch Size	Learning Rate
10	16	3e-5

3 The experiments with different models

Here are some experimental results of the two-phase, where the range following the model name is the confidence interval from SVM classification results. We can see that $[0.3, 1]$ is the best setting for the confidence interval. The SVM method has the fastest processing speed ("QPS" column), while its accuracy is the lowest. If a user care more about the speed of a system, a traditional method – SVM is a better choice.

Table 3: The experimental results of different models.

Model	F_1	QPS
SVM	0.9547	5379.03
Ditto	0.9972	16.45
SVM+Ditto([0.3, 0.7])	0.9547	663.95
SVM+Ditto([0.2, 0.8])	0.9730	442.37
SVM+Ditto([0.1, 0.9])	0.9837	269.80
SVM+Ditto([0.05, 0.95])	0.9864	162.38
SVM+Ditto([0.3, 1.0])	0.9944	79.55
SVM+Ditto([0.4, 1.0])	0.9862	82.43
SVM+Ditto([0.5, 1.0])	0.9835	84.63