A Multi-Task Learning Approach for Answer Selection: A Study and a Chinese Law Dataset

Wenyu Du[1,3], Baocheng Li[2,3], Min Yang[3], Qiang Qu[3], Ying Shen[4]

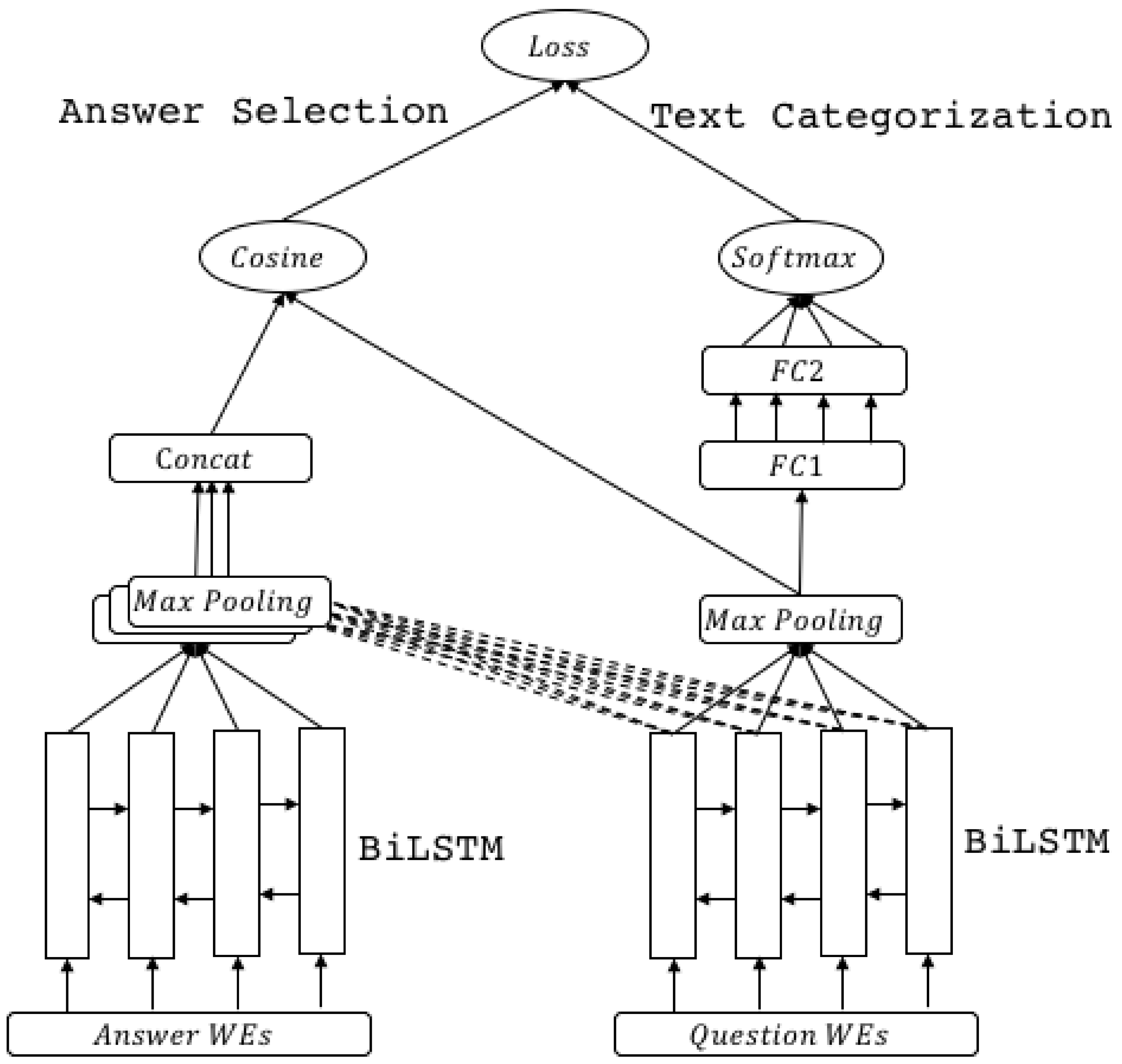
[1]Westlake University [2]Northeast Normal University
[3]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[4]Peking University Shenzhen Graduate School

Introduction

In this paper, we propose a Multi-Task learning approach for Answer Selection (MTAS), motivated by the fact that humans have no difficulty performing such task because they possess capabilities of multiple domains (tasks). Specifically, MTAS consists of two key components: (i) A category classification model that learns rich category-aware document representation; (ii) An answer selection model that provides the matching scores of question-answer pairs. These two tasks work on a shared document encoding layer, and they cooperate to learn a high-quality answer selection system. In addition, a multi-head attention mechanism is proposed to learn important information from different representation subspaces at different positions. We manually annotate the first Chinese question answering dataset in law domain (denoted as LawQA) to evaluate the effectiveness of our model. The experimental results show that our model MTAS consistently outperforms the compared methods.

Model

- **Encoding:** We employ a word embedding layer to convert each word w into a low-dimensional vector e^w, and we use a BiLSTM to learn the hidden states of words in the question and answer.
- **Multihead Attention:** We use multi-head attention mechanism to model the semantics of answers over questions.
- **Answer Selection Task:** The cosine similarities between the final representations of the question and the answer will then be calculated.
- **Text Categorization Task:** Text categorization is an auxiliary task that helps to learn better category-aware text representations.
- **Joint training:** Overall, our model consists of two subtasks, each has a training objective. For the purpose of strengthening the learning of the share document-query representations, we train these two related task simultaneously.



Dataset

- Firstly, we collect a large pool of law related QA pairs with categorical information.
- Then, we remove the redundant QA pairs, and set the minimum length of question and answer to be 14 characters, to avoid the vagueness in the text.
- Our resized QA dataset contains 10 balanced categories with 40,000 questions. Since one question may have multiple answers, we have a clean QA dataset with overall 72,416 positive QA pairs.
- To build the training set for answer selection, we manually collect negative samples by randomly selecting one answer form another category to form the negative sample for each QA pair (positive sample).

Dataset	Train	Dev	Test	Avg len of Question
LawQA	**144,832**	**1000**	**2000**	**45.39**
InsuranceQA	18,540	1000	50	7.16
TrecQA	1162	65	38	11.39
WikiQA	873	126	9	7.18

Experiment

- Experiment result on answer selection task

	Top1 Acc	MAP	MRR
CNN[1]	0.521	0.569	0.640
Bi-LSTM[1]	0.561	0.601	0.674
Bi-LSTM-attention[1]	0.573	0.619	0.688
IARNN-word[2]	0.534	0.584	0.657
AP-LSTM	0.556	0.591	0.669
MTAS w/o multitask	0.577	0.622	0.691
MTAS (Ours)	**0.588**	**0.636**	**0.700**

References

1. Tan, M.; Santos, C. d.; Xiang, B.; and Zhou, B. 2015. *Lstm based deep learning models for non-factoid answer selection*. arXiv preprint arXiv:1511.04108.
2. Wang, B.; Liu, K.; and Zhao, J. 2016. *Inner attention based recurrent neural networks for answer selection* In Proceedings of the 54th ACL, volume 1, 1288–1297.