Chinese Sentence Semantic Matching Based on Multi-Granularity Fusion Model

Xu Zhang, Wenpeng Lu, Guoqiang Zhang, Fangfang Li, and Shoujin Wang

Qilu University of Technology
Sydney University of Technology
oOh! Media
Macquarie University

xuzhang.p@foxmail.com

Outline

>> Introduction

- The semantic matching task measures the relationship between two sentences.

Sentence Pairs	Semantic Match?
S1:哲学的物质范畴和自然科学的物质范畴是什么关系 EN: What is the relation between the physical category of philosophy and that of natural science? S2:简述哲学物质范畴与自然科学的物质概念关系 EN: Describe the relation between physical category of philosophy and the meaning cf material in natural science	YES
S3:我对这些并不感冒是什么意思 EN: If someone says that I'm not interested at these things, what does he means? S4:我对你并不感冒是什么意思? EN: If someone says that I'm not interested at you, what does he means?	NO

Outline

>> Model

- Our work focuses on Chinese sentence semantic matching.
- Traditional approaches
  - Traditional work deals with sentences either from a word perspective or from a character perspective.

Word Granularity

S3: 我 / 对 / 这些 / 并 / 不 / 感冒 / 是 / 什么 / 意思

S4: 我 / 对 / 你 / 并 / 不 / 感冒 / 是 / 什么 / 意思 / ?

Character Granularity

S3: 我 / 对 / 这 / 些 / 并 / 不 / 感 / 冒 / 是 / 什 / 么 / 意 / 思

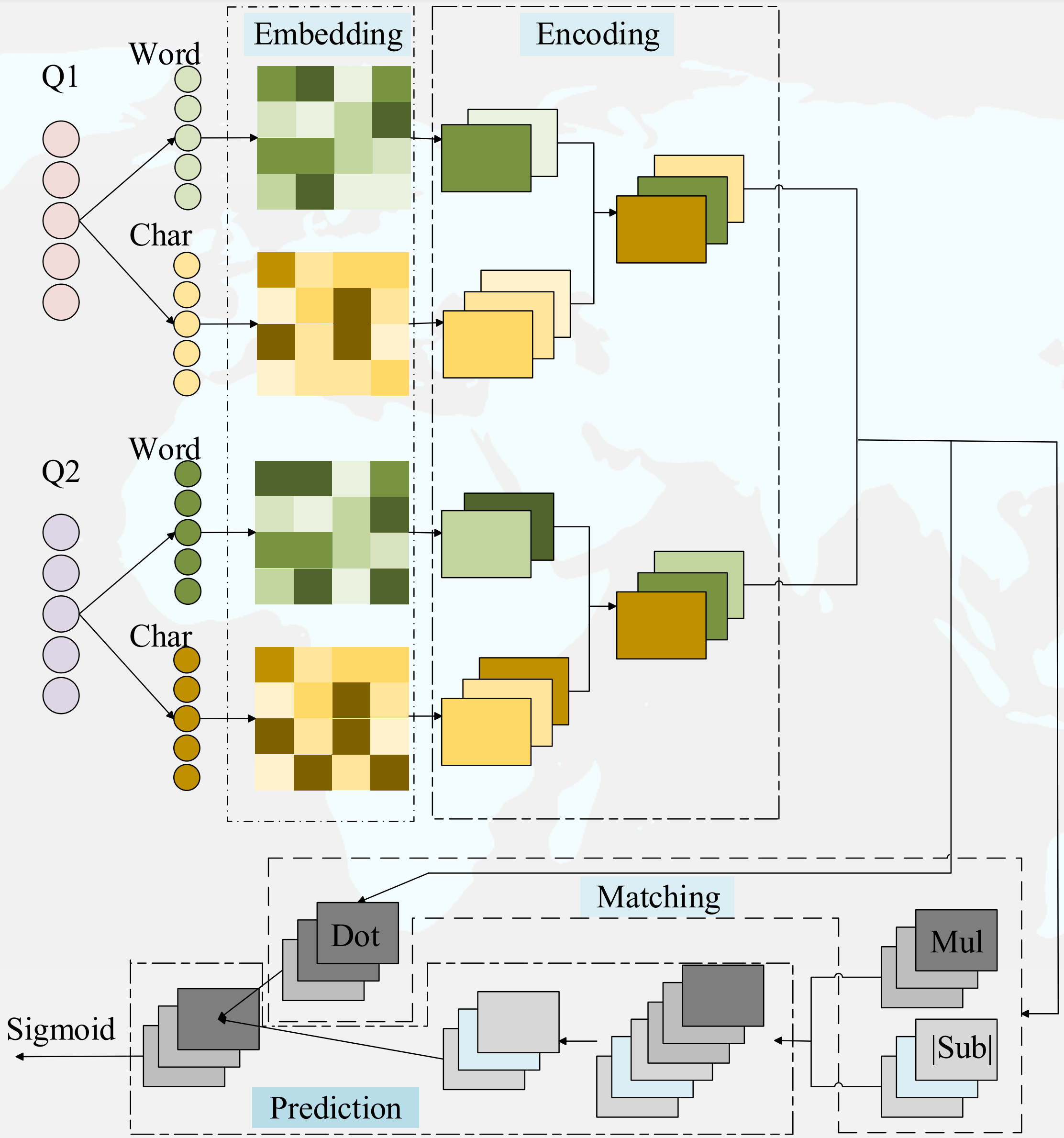
S4: 我 / 对 / 你 / 并 / 不 / 感 / 冒 / 是 / 什 / 么 /意 / 思 / ?

S3-4: 我 / 对 / 你 / 并 / 不 / 感冒/ 感 / 冒 / 是 / 什 / 么 /意 / 思 / ?

>> Model

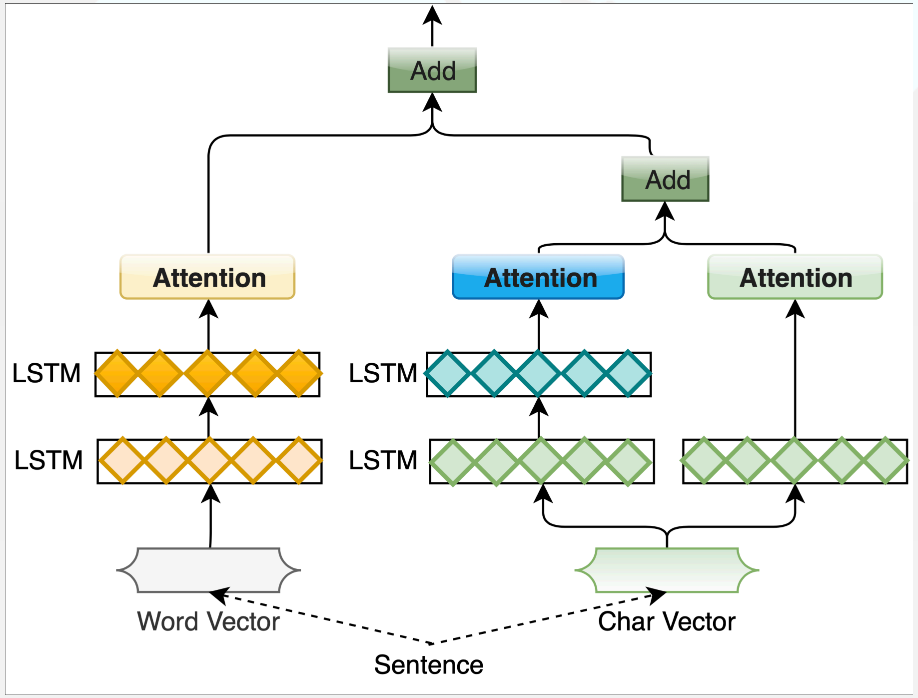
- Ideas
  - For better performance, it is necessary to consider and capture all-side features and complex relationships among the multi-granularity perspective, e.g., characters and words, which requires us to design and implement more sophisticated neural networks.
  - To better optimize the model, we propose the equilibrium cross-entropy, a novel loss function, by setting mean square error (MSE) as an equilibrium factor of cross-entropy.

>> Model


Embedding
Encoding
Q1
Word
Char
Q2
Word
Char
Matching
Dot
Mul
|Sub|
Sigmoid
Prediction

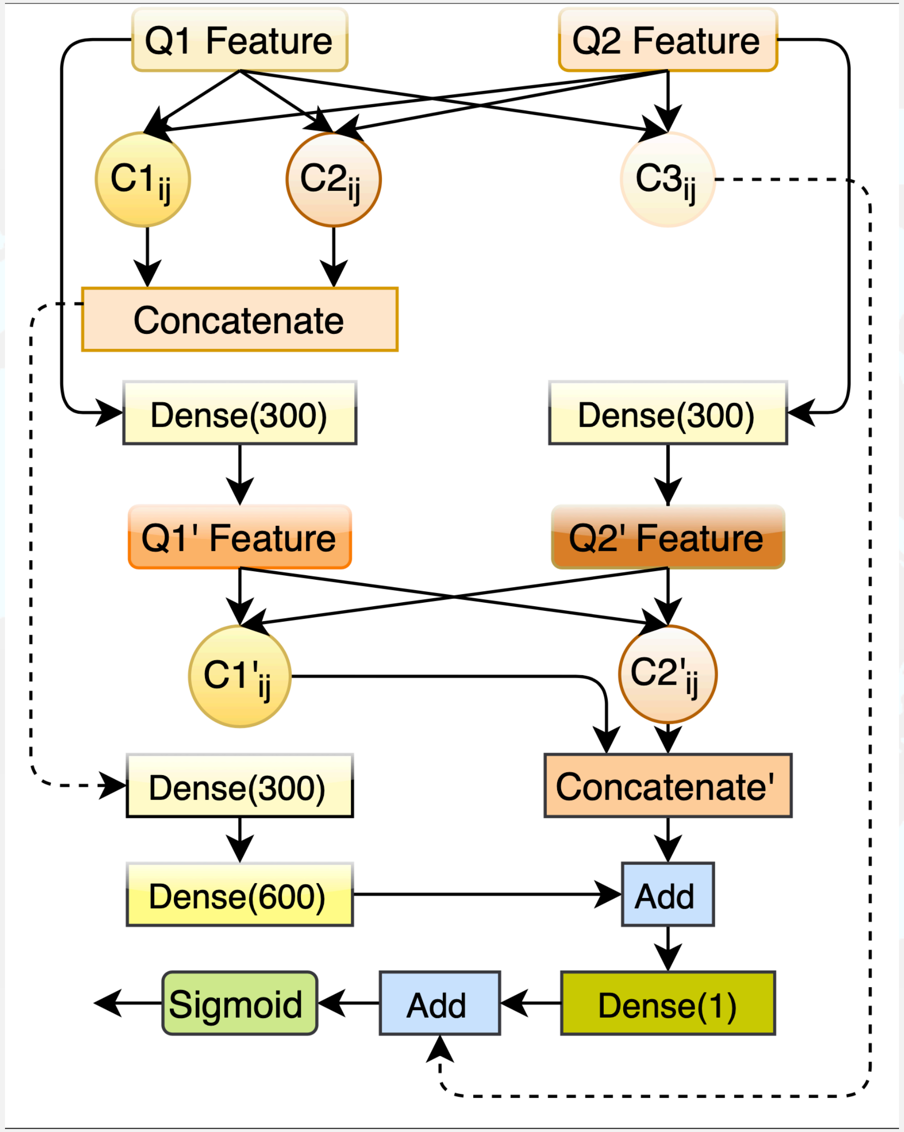
>> Models

- Multi-Granularity Fusion Encoding Layer
  - In the multi-granularity fusion encoder module, we build two different encoders to capture relationships and features among embedding-based representation sentences from character and word perspective respectively.



>> Model

- Interaction Matching Layer
  - In the matching module, we utilize multiple calculation methods to hierarchically compare the similarity of the semantic feature vectors for sentences.



>> Model

- Equilibrium cross-entropy Loss
  - We use the following loss to optimize the multi-granularity fusion model .

$$Loss = -\sum_{i=1}^{n}(L_{mse} * y_{true} \log y_{pred} + (1 - L_{mse}) * (1 - y_{true}) \log(1 - y_{pred}))$$

  - L_{mse} is mean square error (MSE) as an equilibrium factor

$$L_{mse} = \frac{1}{2n}\sum_{i=1}^{n}(y_{true} - y_{pred})^2$$

  - Where the true label is y_{true}, and the predicted result is y_{pred}.

Outline

>> Experiments

- Dataset
  - LCQMC: large-scale Chinese question matching corpus.

Data	Total	Positive	Negative
training	238,766	138,574	100,192
validation	8,802	4,402	4,400
test	12,500	6,250	6,250

Table 2: The distribution of different data sets in LCQMC corpus.

1.Lcqmc: A large- scale chinese question matching corpus. COLING (2018)

>> Experiments

- Baselines
  - On LCQMC dataset, Liu et al. [1] and Zhang et al. [2] have realized relevant and representative state-of-the-art methods, which are used as the baselines to evaluate our model.

1.Lcqmc: A large- scale chinese question matching corpus. COLING (2018)
2. Deep feature fusion model for sentence semantic matching. Computers, Materials & Continua (2019)

>> Experiments

- Effect Comparisons

Table 2. Experiments on LCQMC. *char* means embeddings are character-based and *word* means word-based.

Methods	Precision	Recall	F_1-score	Accuracy
WMD_{char}	67.0	81.2	73.4	70.6
WMD_{word}	64.4	78.6	70.8	60.0
C_{wo}	61.1	83.6	70.6	70.7
C_{ngram}	52.3	89.3	66.0	61.2
D_{edt}	46.5	86.4	60.5	52.3
S_{cos}	60.1	88.7	71.6	70.3
CBOW_{char}	66.5	82.8	73.8	70.6
CBOW_{word}	67.9	89.9	77.4	73.7
CNN_{char}	67.1	85.6	75.2	71.8
CNN_{word}	68.4	84.6	75.7	72.8
BiLSTM_{char}	67.4	91.0	77.5	73.5
BiLSTM_{word}	70.6	89.3	78.92	76.1
BiMPM_{char}	77.6	93.9	85.0	83.4
BiMPM_{word}	77.7	93.5	84.9	83.3
DFF^{o}_{char}	78.58	93.88	$\underline{85.51}$	$\underline{84.15}$
DFF^{o}_{word}	77.69	94.08	85.06	83.53
Our Models				
MGF	81.39	92.90	**86.72**	**85.83**

Outline

We argue that it is necessary to consider and capture all-side features and complex relationships among the multi-granularity perspective.

We designed multi-granularity fusion model to model the Chinese sentence semantic matching .

The experiment shows the superiority of Chinese sentence semantic matching.

We carry out experiment on the Chinese intention matching data set BQ, which also achieves good results, shown in the following table.

Model	Precision	Recall	F_1-score	Accuracy
Random	50.43	50.56	50.49	50.43
TF-IDF	64.68	60.94	62.75	63.83
Text-CNN	67.77	70.64	69.17	68.52
BiLSTM	75.04	70.46	72.68	73.51
BiMPM	82.28	81.18	**81.73**	81.85
DIIN	81.58	81.14	81.36	81.41
DFF^o_{char}	85.38	77.51	81.19	<u>82.17</u>
DFF^o_{word}	83.84	78.66	81.09	81.77
Our Models				
MGF	89.24	74.67	<u>81.21</u>	**82.86**

1.BQ:Alarge-scale domain-specific chinese corpus for sentence semantic equivalence identification. (EMNLP 2018)

Thanks for your attention

Q&A

Qilu University of Technology

2020-05-12