



WUST at the NTCIR-14 STC-3 CECG Subtask

Zhanzhao Zhou, Maofu Liu, Zhenlian Zhang

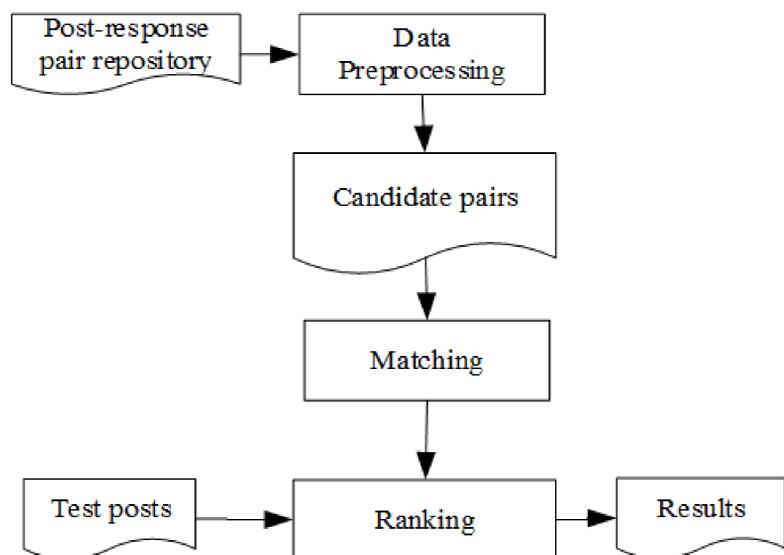
School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430065, China

liumaofu@wust.edu.cn

Introduction

- ◆ Emotions in dialogue can significantly enhance the effect of human-computer interaction. An increasingly growing number of researchers pay attention to the generated responses that is not only appropriate in content but also adequate in emotion.
- ◆ In NTCIR-14, the STC-3 CECG Subtask is defined as follows: Given a Chinese post with a user-specified emotion category of the response to be generated, the goal is to generate a response Y that is coherent with the emotion category.
- ◆ We regard this task as an information retrieval problem. For the given post, we choose the best response by computing the text similarities between the given post and candidates.

System Architecture



1. Data Preprocessing

- ◆ Firstly, we remove the English and Cantonese dialogues in the original dataset. And we construct the inverted index table for the training dataset.
- ◆ Secondly, According to the emotion categories of responses, we divide the training dataset into six subsets and mark the corresponding emotion categories respectively, so the six training subsets are obtained.
- ◆ Finally, for each testing instance, we search for appropriate responses to construct the candidate dataset in the preceding corresponding training subset whose emotion class is the same as the testing instance emotion class by the inverted index table.

2. Matching

- ◆ TF-IDF (Term Frequency-Inverse Documentation Frequency) is an important indicator when extracting keywords. TF-IDF is used to assess the importance of a term to a document.
- ◆ This paper uses the Jieba, a word segmentation tool, to extract keywords by calculating TF-IDF weights.

- ◆ Firstly, the extracted keywords of the two sentences are combined into a union set. The size of the union set is the size of the sentence vector. Secondly, if a sentence contains a keyword of the union set, its weight should be added into the sentence vector. Otherwise, zero should be added into it. In this way, the lengths of two sentence vectors are the same.
- ◆ This paper holds the assumption that if the posts are similar, then the response of the candidate post can serve as the response for the given post.

3. Ranking

- ◆ First, the given post and the post of the candidate dialogue are converted to the TF-IDF vectors respectively, and the text similarity is computed between the two vectors by the cosine similarity.
- ◆ Then we also convert the response in the candidate dialog to a TF-IDF vector and take the same aforementioned works to reach the text similarity.
- ◆ Finally, we take a linear function to combine the two kinds of similarities and rank the responses.

Experiments

- ◆ The official evaluation results are showed as follows:

Table1. The official evaluation results

Submissions/ Emotions	Label0	Label1	Label2	Total	Overall Score	Average Score
WUST	601	211	188	1000	587	0.587
Like	117	36	47	200	130	0.65
Sadness	124	31	45	200	121	0.605
Disgust	111	69	20	200	109	0.545
Anger	137	48	15	200	78	0.39
Happiness	112	27	61	200	149	0.745

- ◆ As the evaluation results show, the system can generate appropriate and reliable responses both in content and emotion and rank the fifth of all the participants. Retrieval based models should be competitive in this challenge.
- ◆ On the one hand, the system don't capture semantic of the dialogues. On the other hand, if we can compare more ranking or re-ranking strategies, that would be better.

Conclusions

- ◆ In this paper, we take the information retrieval method to solve this task. Our proposed WUST system can generate appropriate and reliable responses both in content and emotion.
- ◆ In the future, we would like to propose a novel model using the generated-based approach. And the semantic relevance between post and response should be taken into consideration on proposing novel model. And the special mechanism for addressing emotional factor in dialogues should be proposed.