



WUST System at NTCIR-12 QALab-2 Task

Maofu Liu¹, Limin Wang¹, Xiaoyi Xiao¹, Lei Cai¹, Han Ren²



1. College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430065, P.R. China

2. College of Computer Science, Hubei University of Technology, Wuhan 430068, P.R. China

liumaofu@wust.edu.cn

Introduction

- ◆ In NTCIR-12, the goal of the QALab-2 is to investigate the real-world complex QA (Question Answering) technologies using Japanese university entrance exams and their corresponding English translations on the subject of “world history”.
- ◆ This paper describes our question answering system at NTCIR-12 on QALab-2 task. We first retrieve the documents and sentences related to the question from Wikipedia. Then, the classification model has been constructed based on SVM (Support Vector Machine) in order to solve the question by choosing right or wrong sentence in multiple choice-type questions for the National Center Test, and five kinds of features about questions and choices have been extracted as inputs to the model. Finally, we choose the answer according to the score of each choice.

4. Feature extraction

- ◆ Similarity between question and choice
- ◆ Similarity between choice and related sentence
- ◆ Average similarity between choice and related sentences
- ◆ Distance between choice and negative sentences
- ◆ Word overlap

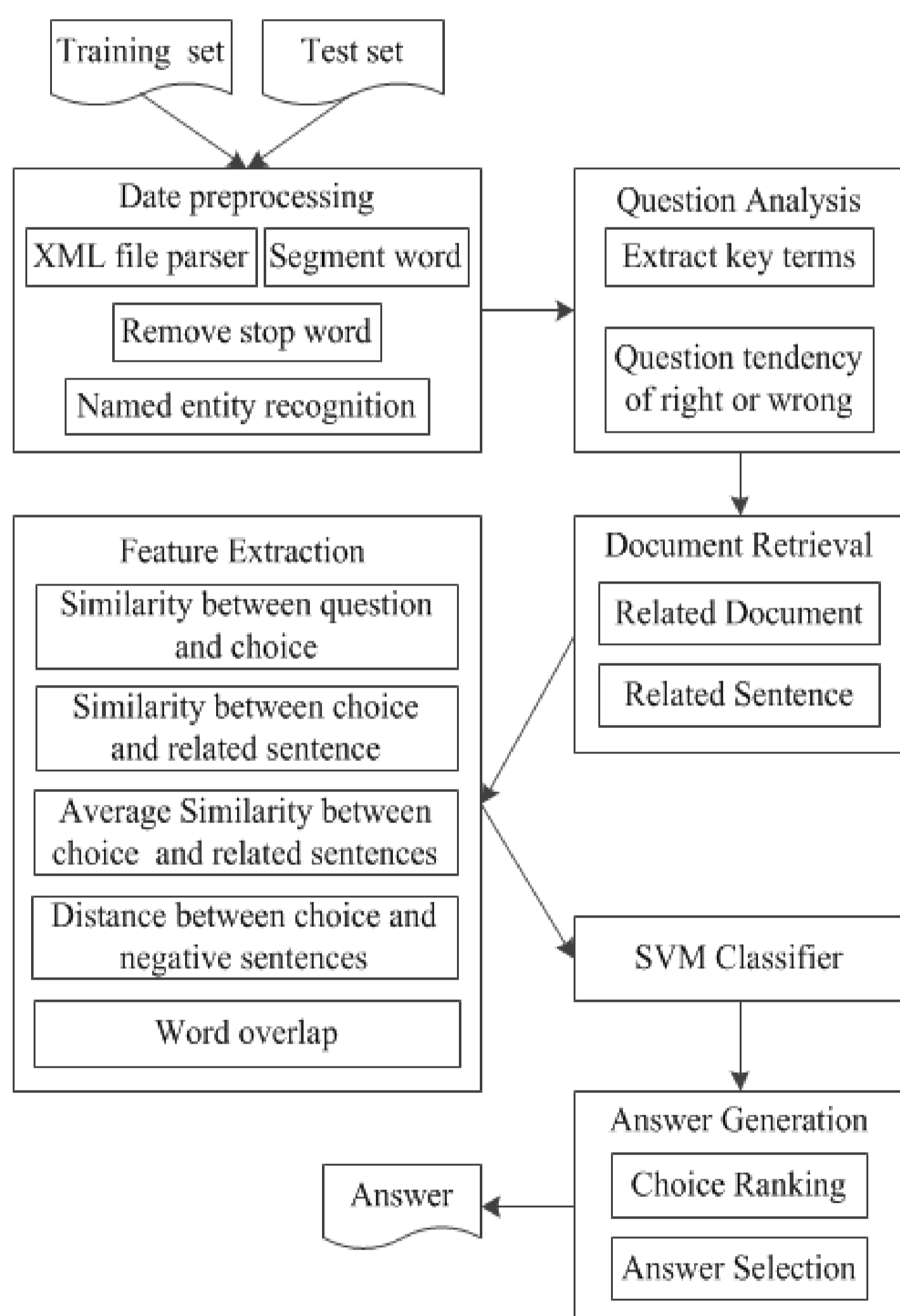
5. SVM classifier

- ◆ We use SVM to solve binary classification problem. The training set are divided into two categories, positive samples and negative samples, among which the correct candidate answer choice constitutes positive sample and marked “1”, the false candidate answer choice constitutes negative sample and marked “0”.

6. Answer generation

- ◆ If the question requires finding the wrong choice, that is, the question belongs to the wrong tendency question, we choose the choice with lowest score as the final answer. Otherwise, if the question belongs to the right question, we choose the choice with highest score as the final answer.

System Architecture



1. Data Preprocessing

- ◆ In the data preprocessing, the main work of the system is to parse the question XML file and process the question sentences.

2. Question analysis

- ◆ Question analysis mainly includes extracting key terms in the question and judging the question tendency of right or wrong.

3. Document retrieval

- ◆ Figure 1 can show the retrieval phase.

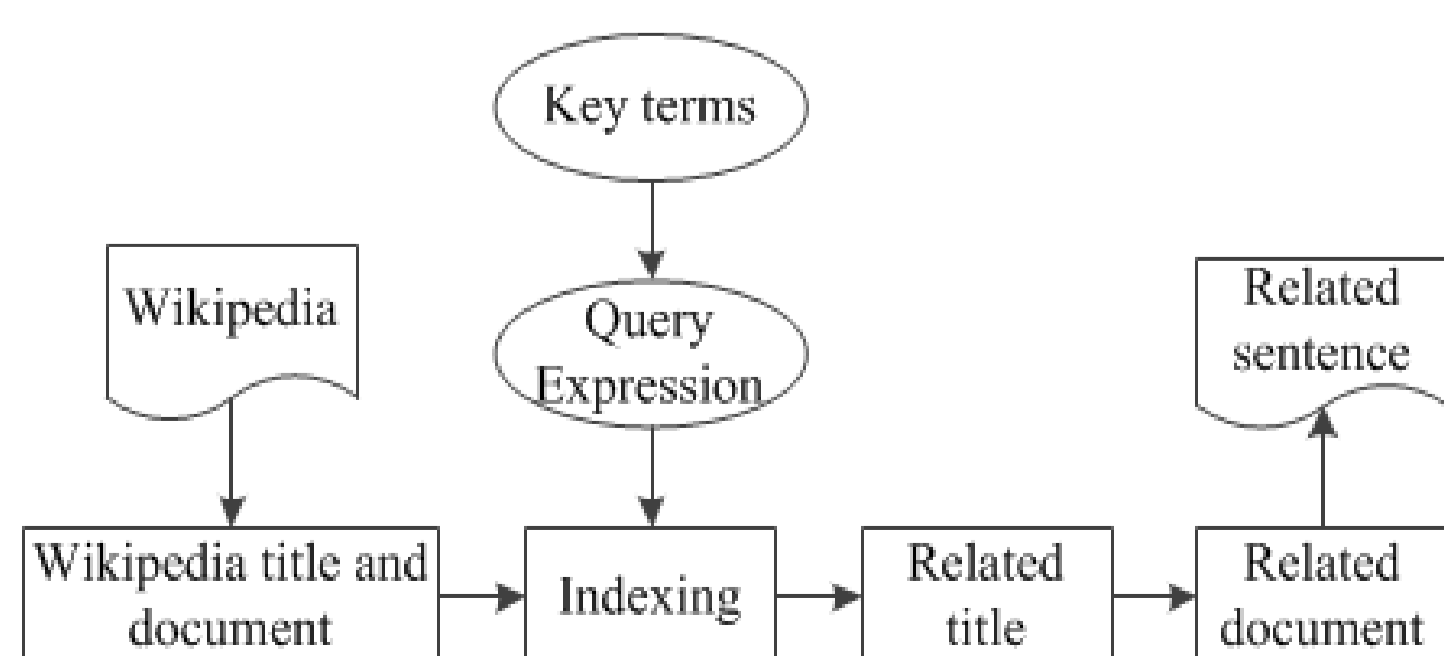


Figure 1. Retrieval phase

Experiments

- ◆ We submitted one formal runs on the phase 3 test data (Center-2011-Main-World History B) for QALab-2 to the NTCIR-12 task organization office and the official evaluation results are listed in the Table 1.

Table 1. The official evaluation result of WUQA system

Run	Center-2011-Main-World History B
Question	36
Correct	6
Total score	17
Correct answer rate	17%

- ◆ The test data contains 36 questions, of which the type of choosing right or wrong sentence has 23. Our system achieved 6 correct answers in 23 questions and the accuracy of the type of choosing right or wrong sentence is 26%.

Table 2. Evaluation result for particular question

Type of questions	question with choosing right or wrong sentence (without image)
Number	23
Correct	6
Accuracy	26%

Conclusions

- ◆ For this result, we find out that there are many deficiencies for our system, such as in the term extraction, question analysis, and feature extraction. In the future, we will improve these deficiencies in order to obtain higher accuracy. For example, we will introduce more features into the system, such as semantic features and so on. In addition, our system is only for the type of choosing right or wrong sentence, we will study other types of questions in the future.