A Textual Entailment System using Web based Machine Translation System

Partha Pakray*, Snehasis Neogi†, Sivaji Bandyopadhyay‡, Alexander Gelbukh§

*Jadavpur University, Kolkata-700032, India
†Center for Computing Research, National Polytechnic Institute, Mexico City, Mexico

Abstract

The article presents the experiments carried out as part of the participation in Recognizing Inference in Text (NTCIR-9 RITE) for Japanese. NTCIR-9 RITE has four subtasks, Binary-class (BC) subtask, Multi-class (MC) subtask, Entrance Exam and NTCIR-9 RITEQA. We have participated in all the four subtasks.

System Architecture 1 (Run 1): BC Task

The system accepts pairs of text snippets (text and hypothesis) at the input and gives a Boolean value at the output: Y if the pair entails, otherwise N.

In this architecture we calculate the percentage of unigram and bigram match experimented over the NTCIR-9 RITE BC subtask test data. These values are then compared with the predetermined threshold value calculated over the NTCIR-9 RITE BC subtask development data. If the N-gram matching percentage is more than or equal to the threshold value, entailment label “YES” is assigned with the pair. Otherwise, entailment label “NO” is assigned which means no entailment.

System Architecture 2 (Run 2): BC Task

The system architecture is based on N-Gram matching on the input Japanese text pair (t1, t2) to identify whether t1 entails (infer) a hypothesis t2 or not. The system first learns the N-Gram (unigram and bigram) word overlap on the development data. The learned system is then applied on the test data to classify whether the text entails the hypothesis or not.

In this architecture the N-gram percentage is tokenized and the different N-gram percentages (unigram, bigram, trigram) are evaluated. These percentage values are then compared with the percentage values calculated over the development data. If a match is found for the percentage values for each text – hypothesis pair of the test data, the same entailment label (Y,R,B,C,I) of the development data is used to classify the text – hypothesis of the test set. Otherwise it classifies the pair as NO (Independence).

System Architecture 3 (Run 3): BC Task

The system considers the output of the previous two systems (Run 1 and Run 2) as input. The entailment decision is taken based on voting on the decisions in Run 1 and Run 2.

System Architecture 4 (Run 1): MC Task

The RITE MC subtask development data is similar to that of the BC Subtask except the entailment label. Here multi-way entailment label is used to detect entailment or no entailment in the text – hypothesis pair. The 5-way labeling subtask detects (forward / reverse / bidirectional) entailment or no entailment (contradiction / independence) in a text pair.

System Architecture 5: Entrance Exam Task

The Entrance Exam subtask is similar to that of the BC subtask in terms of input and output. All the data are created from actual college-level entrance exams. In this subtask the entailment label is same as the BC subtask (Y, N). The system for the entrance exam Subtask is similar to the system architecture 2 used for the BC subtask with some additions in the matching module.

System Architecture 6: RITEQA

The RITEQA Subtask is also similar to the BC subtask in terms of input and output, but it includes an embedded answer validation component in the Question Answering system. This way, the impact of NTCIR-9 RITE to an overall end-to-end application can be measured.

In this paper, textual entailment systems mainly based on the lexical similarity modules have been developed. For MC subtask, the system can learn using some supervised learning methods (e.g. SVM, CRF).

Conclusions

We acknowledge the support of the CONACyT Mexico – DST, India project “Answer Validation through Textual Entailment”.

Authors

Partha Pakray parthapakray@gmail.com
Snehasis Neogi snehasis9009@gmail.com
Sivaji Bandyopadhyay sivajib@iiti.ac.in
Alexander Gelbukh gelbukh@e邮址.com

Acknowledgements

 prolong