Cornell Belief and Sentiment System at TAC 2016

Vlad Niculae, Kai Sun, Xilun Chen, Yao Cheng, Xinya Du,
Esin Durmus, Arzoo Katiyar and Claire Cardie
Cornell University
Ithaca, NY
vn66,kai.sun,xc253,yc2258,xd75,ed459,ak979,claire.cardie@cornell.edu

Abstract

We describe the 2016 system of the CornPittMich team for the TAC Belief and Sentimen (BeSt) task for English and Chinese.

The Cornell-Michigan-Pittsburgh team (aka CORNPITTMICH) submitted systems for both English and Chinese. We employ a combination of simple, hand-crafted rules and machine learning-based approaches.

1 English systems

The English system has two stages: the first stage performs link prediction; the second, belief and sentiment classification.

- **System 1.** Our system 1 uses a simple rule-based pruning approach for link prediction and multinomial logistic regression for stage 2.
- **System 2.** The second submission combines the same rule-based link predictor (stage 1) with an LSTM-based encoder (stage 2).
- **System 3.** The third submission uses factorization machines to predict sentiment links and the rule-based predictor for beliefs (stage 1); multinomial logistic regression is used for stage 2.

For all of the above, the rule-based link predictor bases its decisions on information drawn from the text span between the source and target entity/relation/event pairs. All stage 2 systems are trained on gold positive links and spurious NONE/NA links predicted by the rule-based component. The factorization machine + rule-based strategy of System 3 consists of adding the real-valued score from the facorization machine to the binary +1/-1 score of the rule-based decision; we choose a threshold for link prediction based on end-to-end system performance on a validation set.

2 Chinese systems

Our Chinese systems use a separate component for handling sentiment vs. belief.

Sentiment. A hybrid approach to sentiment classification is used for both discussion forum and newswire. The model consists of:

(a) A neural network for sentence-level sentiment analysis. The neural network is composed of a single LSTM layer, an average pooling layer followed by a softmax layer. The neural network is trained with about 4k sentences from Weibo with polarity annotated (positive/negative/none). Features used by the neural network include word vector, part-of-speech tag, and word-level sentiments from dictionaries. The model is used to analyse the sentiments of the mention texts, triggers, and sentences where mention texts or triggers appear.

(b) A rule-based model for finding the source of a sentiment. In particular, given a sentiment with its mention text or trigger, for discussion forum, it looks for the post where the mention text/trigger first appears and uses the post author as the source; for newswire, it searches

for words for reporting speech (such as "say", "mention") around which the mention text/trigger appears, and uses the first entity on the left side of the reporting speech as the source.

(c) A rule-based model that outputs the final results based on the output of model (a), the source output by model (b), as well as a bunch of high level features such as indicators of entity/relation/event, text length, number of entities in the sentence, etc.. The main function of this model is to set different thresholds of accepting the positive/negative predictions from the neural network for different scenarios. For example, the acceptance threshold could be relaxed a little bit if there is only one entity in a sentence, since it would be safer to use the sentence-level sentiment as the sentiment towards an entity. The parameters of this model are automatically tuned on the BeSt training data.

The only difference between our submissions is the metric for tuning the parameters of model (c). In particular, the metrics for submission 1,2,3 are focused on good F-score, recall, precision respectively.

2.1 Belief

Three submissions are identical. For discussion forum, the output was obtained by a combination of a rule-based model and a linear model trained on the BeSt training data. For newswire, the output was obtained by a rule-based model.

Specifically, the rule-based model is a simple model that always outputs type="cb" and polarity="pos" for each relation and event and it uses the same model as the sentiment (b) to find the source of a belief. The linear model[1] takes the text around the relation/event mention and decides whether or not there is a belief. If the answer is no, it removes the corresponding belief output (produced by the rule-based model) from the final output.

[1] We used TextGrocery: https://github.com/2shou/TextGrocery.