Recurrent Attention Network on Memory for Aspect Sentiment Analysis

Abstract

1. **Introduction**

- The goal of aspect sentiment analysis
- For example
- Challenging task
- Problem of Previously proposed model(TD-LSTM)
- Introduction of Main framework (Solve the problem)
- datasets

2. **Related Work**

- The task of aspect sentiment classification belongs to entity-level sentiment analysis
- Conventional representative method -> rule-based, statistic-based, 2 tuple of opnion target and opinion word, Probabilistic Soft Logic, SVM -> labo-rious feature engineering work or massive extra-linguistic resources
- Neural Network(NNs) -> Rec-NN(tree for syntactic analysis) -> need dependency parsing
- Convolution NNs -> sentiment of clause -> assumption that an opinion word and its target lie in the same clause
- TD-LSTM -> not work well when the opinion word is far from the target
- Neural Turing Machine(NTM) -> Attention Mechanism -> single or multiple attentions

3. **Our Model**

   A. **Input Embedding**

- Unsupervised method such as Glove or CBOW

   B. **BLSTM for Memory Building**

- MemNet, Deep Bidirectional LSTM (DBLSTM), forward and backward, 2 layers

   C. **Position-Weighted Memory**

- Predicting respective sentiments of targest Problem -> produce a tailor-made input memory for each target -> weighted memory

D. Recurrent Attention on Memory

- To accurately predict the sentiment of a target:
- Correctly distill the related information from its position-weighted memory -> multiple attentions
- Appropriately manufacture such information as the input of sentiment classification -> recurrent network GRUs (non-linearly combines the attention results)
- Calculate

E. Output and Model Training

4. Experiments

A. Experimental Setting

- Four datasets
- 300-dimension word vectors pre-trained by Glove
- Domain-specific training corpus TD-LSTM
- General embedding from ~(paper)
- CBOW corpus for Chinese experiments

B. Compared Methods

- Our framework: Recurrent Attention on Memory(RAM)
- Compare: Average Context(AC), SVM, Rec-NN, TD-LSTM, TM-LSTM-A, MemNet

C. Main Results

- Evaluation metric : Accuracy, Macro-averaged F-measure
- Table

D. Effects of Attention Layers

- Table: 1 to 5 attention layers

E. Effects of Embedding Tuning

- RAM-3AL-T-R: It does not pre-train word embeddings, but initializes embeddings randomly and then tunes them in the supervised training stage
- RAM-3AL-T: using the pre-training embeddings initially, also tuned in the training
- RAM-3AL-NT: the pre-training embeddings are not tuned in the training

F. Case Study

5. Conclusions and Future Work

- Need a mechanism to stop the attention process automatically
- Try other memory weighting strategies to distinguish multiple targets in one comment more clearly

Reference