Lexicon Guided Attentive Neural Network Model for Argument Mining

Jian-Fu Lin[1], Kuo Yu Huang[1], Hen-Hsen Huang[2,3] and Hsin-Hsi Chen[1,3]

[1] Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan
[2] Department of Computer Science, National Chengchi University, Taipei, Taiwan
[3] MOST Joint Research Center for AI Technology and All Vista Healthcare, Taipei, Taiwan

Introduction

- Lexicon information is one of the most frequently used features in the argument mining research.
- Most of the previous researches incorporating lexicon use lexicons composed by human beings or derived by hand-crafted rules, resulting in domain-specificity.
- In comparison to scarcity of lexicon resources in AM, the resources in other domains like sentiment analysis are relatively abundant.
- We propose a attentive NN model that can incorporate lexicon information and experiment on lexicons form various domains

Data

- We conduct the experiments on the dataset released by *Stab et al. (2018)* which contains **25,492** labeled sentences in total. The label of each sentence can be either **supporting**, **attacking**, or **non-argumentative**.

All	Supporting	Attacking	Non-Arg
25,492	4,944	6,195	14,353

Lexicon Resource

- **Claim Lexicon.** The claim lexicon is a lexicon containing words with argumentative characteristics. *Levy et al. (2017)* use the appearance of the term ”that” as a weak signal of sentences containing argumentative components. The lexicon contains around 600 claim words.

- **Sentiment Lexicon.** We leverage the sentiment lexicon built by *Hu and Liu (2004)* that contains around 6,800 words. Each word is labeled as negative and/or positive.

- **Emotion Lexicon.** The emotion lexicon built by *Mohammad and Turney (2013)* contains around 14,200 words. Each word in the lexicon is given eight emotion labels. In the experiment, we select the words that have at least one emotion labeled as positive, resulting a list of 6,468 words.

- **WordNet.** To expand a topic T composed of words *w1, w2, ..., wk,* we expand each of the words in it. For each word *wi*, we use WordNet *(Miller, 1995)* to find its corresponding synonyms. The size of WordNet is around 155,300.

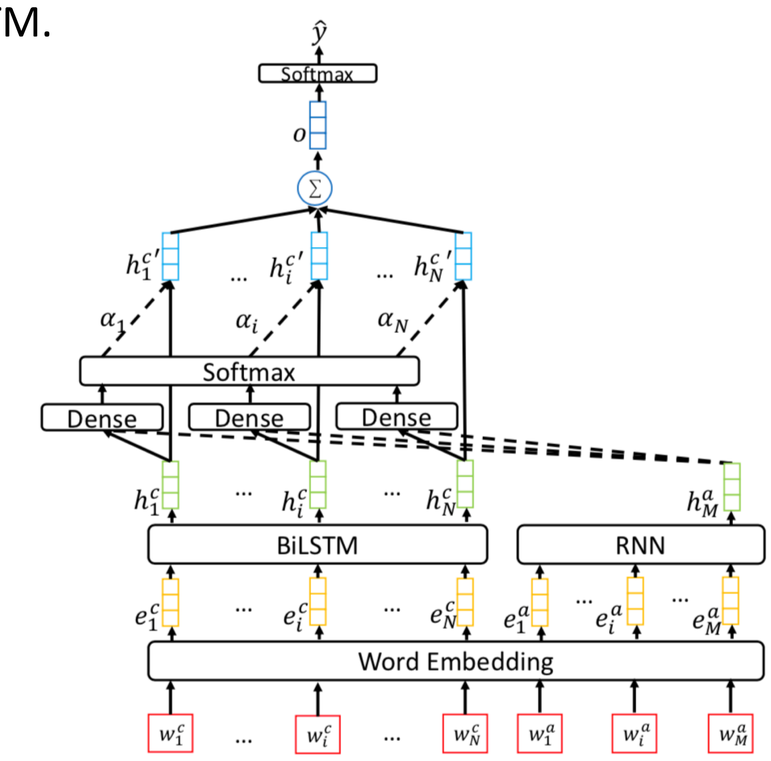
Model

Baseline Model

- In this experiment we implement a baseline model with the vanilla bidirectional LSTM (BiLSTM).

Lexicon-Guided Attentive Neural Network Model

- The proposed model integrates an outside lexicon resource into attention mechanism. For each input sentence, we compose an additional sentence $\boldsymbol{A}$, which contains words $\boldsymbol{w_1^a, w_2^a, \ldots, w_M^a}$ based on the given lexicon. The sentence A is then fed into a simple RNN, and we use the output of the RNN as a query of attention mechanism, weighting the outputs of the BiLSTM.



The architecture of Lexicon Guided Attentive Neural Network Model

Result

Model	Macro-F1	Lexicon size (#words)
BiLSTM	$.5337 \pm .0123$	n/a
ClaimLex*	$.5684 \pm .0222$	~600
SentimentLex*	$.5718 \pm .0165$	~6,800
EmotionLex*	$.5695 \pm .0129$	~6,500
WordNet*	$\mathbf{.5788 \pm .0142}$	~155,300

Conclusion

- We propose a novel approach to leverage the lexicon, and conduct our experiment on various lexicons that are from both in-domain and out-of-domain sources.
- Experimental results confirm the effectiveness of the integration of lexicon information.
- We observe that the result of integrating claim lexicon (Levy et al., 2017) is out of our expectation. It reflect that leveraging out-domain but high-coverage resources can outperform leveraging in-domain but low-coverage resource.

National Taiwan University
科技部人工智慧技術暨全幅健康照護聯合研究中心
Most Joint Research Center for AI Technology and All Vista Healthcare
NLP Lab