An Attention Based Multi-view Model for Sarcasm Cause Detection (Student Abstract)

Hejing Liu[1,2], Qiudan Li[2,3], Zaichuan Tang[1,2] and **Jie Bai[2,3]**
[1]School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China
[2]Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[3]Shenzhen Artificial Intelligence and Data Science Institute (Longhua)
{liuhejing2018, qiudan.li, tangzaichuan2019, baijie2013}@ia.ac.cn

Abstract

Sarcasm often relates to people's implicit discontent with certain products and policies. Existing research mainly focus on sarcasm detection, while the deep causal relationships in the full conversation remained unexplored. This paper formulates a novel research question of **sarcasm cause detection**, and proposes an attention based model that simultaneously captures different semantic associations as well as the inner causal logics in multi-view manner. Experiments on public Reddit dataset prove the efficacy of the proposed model.

Introduction

Sarcasm are often used on social media, especially when people are disappointed with certain products, policies, and events. Detecting the cause of sarcasm could help the business companies and management departments better understand people's attitudes, obtain effective feedbacks, and make timely adjustments. Here is an example of sarcastic conversation from the online forum Reddit, where people are discussing their use experience of headphone devices:

> Context (User A): … *1.5 hours per charge.* ***Listen for an hour, the battery is running low.*** *Wait 15-30min to charge them, rinse and repeat. Even with fast charging it gets annoying.*
>
> Sarcasm response (User B): *Just listen one ear at a time!*

It can be seen that the poor battery performance leads to the discontent of both users, and triggers the ironic comment of user B. The cause of sarcasm uncovers the weak point of the headphone, and provides directions for the manufactory to improve their product. Similar situations also exist in policy and event-related discussions, where sarcasm cause detection benefits the comprehension of people's opinions and responses. Although the practical value is significant, there is few work related to identifying sarcasm causes from the context at present. Existing research mainly regards the context as an external information resource to detect sarcasm, while the potential causal relationships of the conversations remained unexplored. To mine the sarcasm cause, there are several challenges need to be addressed. First, the topic in one conversation is usually very specific, it's difficult to distinguish the most relevant part to the sarcasm. In the above example, although the word *charge* appears repeatedly, the fundamental cause of sarcasm is the battery, not charging. Second, the sarcasm usually has deep semantics, and the causal relationships are hard to be identified. For example, the part listen for an hour and listen one ear at a time form a causal logic that using only one ear of the head-phone might last the listening time, such semantic association takes much effort to capture if without the full view of the scene.

In this paper, we formulate a novel research question of sarcasm cause detection, and propose an attention based multi-view model that simultaneously captures different semantic associations, as well as the inner causal logics between the sarcasm and the context. Specifically, the model leverages the advantages of four different measurements to learn the semantic relevance in both word-wise and sentence-wise. The importance of each measure is automatically identified through attention mechanism (Bahdanau et al. 2014). In addition, the long-term semantics are captured in pair-level to discover the deep causal relationships from the full view. Experimental results on public Reddit datasets show the efficacy of this method.

Proposed Model

The sarcasm cause detection aims to identify the specific part of the context that provokes the user's sarcasm. Given a sarcastic post S and its context clause set $\{c_1, c_2, \dots, c_m\}$,

for each possible sentence pair in $\{(S, c_1), \ldots, (S, c_m)\}$, the question is to determine whether the c_i is the cause of S.

Semantic Relevance Measure on Sentence-level. Sarcasm and its cause may exist different kinds of association patterns. Previous research has proved the utility of explicitly modeling the association patterns from multiple information resources (Li et al. 2020). Therefore, we propose a hybrid measurement to obtain the semantic relevance feature from multi-view manner. After encoding through an attentive Bi-LSTM layer (Hochreiter et al. 1997), the representation vectors of sarcasm $\boldsymbol{S}$ and clause $\boldsymbol{C}$ are compared using the following four measures, i.e. taking the absolute value of $\boldsymbol{S} - \boldsymbol{C}$ (Sub-abs measure), the square value of $\boldsymbol{S} - \boldsymbol{C}$ (Sub-square measure), the Hadamard product, and the average value of $\boldsymbol{S}$ and $\boldsymbol{C}$. The first two measures are used to show the distance vector of the sarcasm and the clause in semantic space, which captures the conflict and similar expressions in sentence-wise. The Hadamard product and average operation are used to acquire the word-to-word semantics. Considering the measures might have different importance, the model adopts attention mechanism to learn their weights, then generates the final relevance representation $\boldsymbol{R}$.

$$w_i = \frac{exp(v_w^T Measure_i)}{\sum_i exp(v_w^T Measure_i)} \quad (1)$$

$$R = \sum_1^4 w_i Measure_i \quad (2)$$

Causal Relationship Learning on Pair-level. To identify the causal relationships between two sentences, the model further learns the pair-level representations for each sarcasm and clause. This component first concatenates the embeddings of the sarcasm and the clause, then encodes the pair embedding through attentive Bi-LSTM. The final vector $\boldsymbol{V}$ is obtained by integrating the sentence representations $\boldsymbol{S}$ and $\boldsymbol{C}$, the relevance feature $\boldsymbol{R}$, as well as the pair representation $\boldsymbol{P}$. We use the standard cross-entropy as the objective loss.

$$V = S \oplus C \oplus R \oplus P \quad (3)$$

Experiments and Results

Dataset. We constructed our experimental data from the public Reddit sarcasm dataset SARC (Khodak et.al, 2017). We extracted the sarcasm responses and their contexts of 7 subreddits, then invited three annotators to label the cause clauses, the Fleiss' kappa coefficient was 0.76. The constructed dataset includes 4987 clause pairs, with 1269 positive samples and 3663 negative samples. The training set, validation set, and test set are random split as 7:2:1.

Baseline Methods. The proposed model is compared with: (1) TF-IDF+SVM, (2) CNN (Kim 2014), (3) LSTM. The baseline methods use the concatenation of sarcasm and clause vector as the final representation. We also compared the results of sentence-level component with single measure. The F1 score and AUC score are used as evaluations.

Methods		F1	AUC
(1) TF-IDF+SVM		0.5906	0.4533
(2) CNN		0.6774	0.5456
(3) LSTM		0.6633	0.5312
Proposed Model* (sentence-level only)	Sub-abs	0.6755	0.5523
	Sub-square	0.6809	0.5631
	Hadamard	0.6843	0.5626
	Average	0.6762	0.5818
	Merge by attention	**0.7048**	**0.5937**
Proposed Model (sentence and pair-level)		**0.7114**	**0.6292**

Table 1. The Experimental Results of Each Meth

Results. Table 1 shows the experimental results. It can be seen that our proposed model that uses both the sentence and pair-level components achieves the best performance. On the sentence-level component with single measure, the Hadamard product shows the best results, followed by Sub-square measure, average operation and Sub-abs measure. One possible reason is the word-wise operation provides more detailed information. The Sub-square measure performs better than average operation and Sub-abs measure, this might because the conflict semantics of sarcasm and its cause have more weights with Sub-square operation, which is a helpful feature on sarcasm cause detection. The results demonstrate that merging the measurements through attention mechanism is the most effective way to capture the different semantic associations on sentence-level.

Acknowledgements

Qiudan Li is the corresponding author. This work was partially supported by the National Key Research and Development Program of China (Grant No.2020AAA0103405), the National Natural Science Foundation of China (Grant No.62071467, 71902179, 71621002).

References

Bahdanau, D., Cho, K., and Bengio, Y. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Khodak, M., Saunshi, N., and Vodrahalli, K. 2017. A large self-annotated corpus for sarcasm. *arXiv preprint arXiv:1704.05579*.

Kim, Y. 2014. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.

Li, Q., Zeng, D. D., Xu, D. J., Liu, R., and Yao, R. 2020. Understanding and Predicting Users' Rating Behavior: A Cognitive Perspective. *INFORMS Journal on Computing*.