Making Better Use of Training Corpus: Retrieval-based Aspect Sentiment Triplet Extraction via Label Interpolation

Guoxin Yu[1,2], Lemao Liu[3*], Haiyun Jiang[3], Shuming Shi[3], Xiang Ao[1,2*]

1 Key Lab of IIP of Chinese Academy of Sciences, Institute of Computing Technology.

2 University of Chinese Academy of Sciences. 3 Tencent AI Lab, China.

1 Introduction

In addition, the food is very good and the prices are horrible.

(***food***, ***good***, *positive*) (***prices***, ***horrible***, *negative*)

Aspect Sentiment Triplets Extraction (ASTE) aims to extract all triplets containing aspect terms, opinion terms and their sentiment polarities.

2 Motivation

It is promising to introduce inter-sentence information into some challenging cases, where the intra-sentence information is insufficient.

Long distance

The food is spicy and delicious, including the shrimp appetizer, the cod, and the ***scallop roll****.* Uncommon Aspect Terms

Knowledge Source

3 Challenges

ASTE has challenges when adapting retrieval-augmented methods:

- its purpose includes predicting the sentiment polarities and it is usually aspect-dependent;
- triplets with similar semantic similarities may have conflict sentiment polarities.

battery + *long* → ***positive***

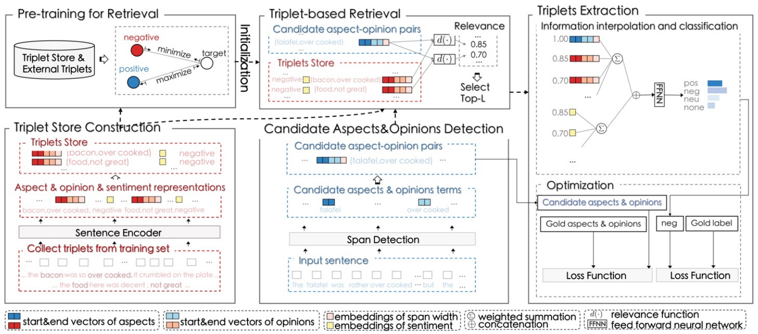
boot-time + *long* → ***negative***

Existing retrieval-augmented methods may fetch triplets with high semantic similarity but opposite sentiment, giving a false guidance.

The ordered cocktail from here was spicy that I could barely drink it.

(***cocktail***, ***spicy***, *negative*) Deduce ⟹ (***scallop roll***, ***spicy***, *negative*) ✗

4 Our Method



- We first construct a knowledge store and detect all candidate aspect-opinion term. $\mathcal{M} = \{(K^i, V^i) | i \in [1, |\mathcal{M}|]\}$

$$E_A = h_{A_1} \oplus h_{A_2} \oplus f_{\text{span}}(A_2 - A_1 + 1), \quad K = E_A \oplus E_O,$$
$$E_O = h_{O_1} \oplus h_{O_2} \oplus f_{\text{span}}(O_2 - O_1 + 1), \quad V = f_{\text{sentiment}}(y),$$

- We pre-train the retriever by contrastive learning, which prompts the retrieved triplets have both semantic and sentiment similarities.

$$\mathcal{L}_{\text{pre}} = d(A, O; A', O')^2 - \left(1 - d(A, O; A'', O'')\right)^2$$

- For each candidate pair, we retrieve semantic similar triplets from the store according to a relevance score.

$$d(A, O; A^i, O^i) = K^\top \mathbf{W} K^i,$$

- We interpolate their label information into the augmented representation of the candidate pair to predict the sentiments.

$$h(A, O) = (K + \sum_{l=1}^{L} \alpha_l K^l) \oplus \sum_{l=1}^{L} \alpha_l V^l,$$

$$\alpha_l = \frac{\exp\left(d(A, O; A^l, O^l)\right)}{\sum_{j=1}^{L} \exp\left(d(A, O; A^j, O^j)\right)},$$

$$P_{\text{ext}}(y|A, O, \mathbf{X}) = \text{softmax}(F(h(A, O)))[y],$$
$$y \in \{\text{positive}, \text{negative}, \text{neutral}, \text{none}\},$$

- We jointly train the retriever and triplets extractor:

$$\mathcal{L}_{\text{det}} = -\sum_{\mathbf{X}} \sum_{S \in \mathcal{S}(\mathbf{X})} \log P_{\text{det}}(c|S, \mathbf{X}),$$

$$\mathcal{L}_{\text{ext}} = -\sum_{\mathbf{X}} \sum_{A, O \in \mathcal{S}(\mathbf{X})} \log P_{\text{ext}}(y|A, O, \mathbf{X}),$$

$$\mathcal{L} = \mathcal{L}_{\text{det}} + \alpha \cdot \mathcal{L}_{\text{ext}},$$

5 Experiment

Comparison on ASTE-DATA-V1

Model	Res14		Lap14		Res15		Res16	
	Pair	*Triplet*	*Pair*	*Triplet*	*Pair*	*Triplet*	*Pair*	*Triplet*
WhatHowWhy◇	56.10	51.89	53.85	43.50	56.23	46.79	60.04	53.62
CMLA+◇	48.95	43.12	44.10	32.90	44.60	35.90	50.00	41.60
RINANTE+◇	46.29	34.03	29.70	20.00	35.40	28.00	30.70	23.30
Unified+◇	55.34	51.68	52.56	42.47	56.85	46.69	53.75	44.51
Dual-MRC◇	74.93	70.32	63.37	55.58	64.97	57.21	75.71	67.40
Generative▽	77.68	72.46	66.11	57.59	67.98	60.11	77.38	69.98
GAS♯	–	70.20	–	54.50	–	59.10	–	65.00
LEGO♯	–	72.60	–	59.50	–	63.20	–	71.50
$\text{JET}^t_{M=6}$ ▽	–	60.41	–	46.65	–	53.68	–	63.41
$\text{JET}^o_{M=6}$ ▽	–	63.92	–	50.00	–	54.67	–	62.98
SPAN*	78.62	73.96	69.48	60.59	71.56	64.50	78.85	70.48
RLI(Ours)	**79.92**	**74.98**	**70.27**	**61.97**	**72.66**	**65.71**	**81.29**	**73.33**

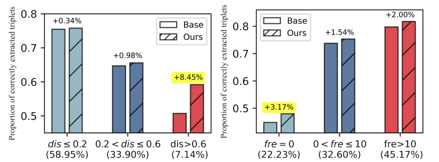
Comparison on ASTE-DATA-V2

Model	Res14			Lap14			Res15			Res16		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
WhatHowWhy◇	43.24	63.66	51.46	37.38	50.38	42.87	48.07	57.51	52.32	46.96	64.24	54.21
TOP♯	63.59	73.44	68.16	57.84	**59.33**	58.58	54.53	63.30	58.59	63.57	71.98	67.52
BMRC♯	72.17	65.43	68.64	**65.91**	52.15	58.18	62.48	55.55	58.79	69.87	65.68	67.35
RL♯	70.60	68.65	69.61	64.80	54.99	59.50	65.45	60.29	62.72	67.21	69.69	68.41
GAS♭	–	–	72.16	–	–	60.78	–	–	62.10	–	–	70.10
OTE-MTL♯	63.07	58.25	60.56	54.26	41.07	46.75	60.88	42.68	50.18	65.65	54.28	59.42
GTS◇	67.76	67.29	67.50	57.82	51.32	54.36	62.59	57.94	60.15	66.08	69.91	67.93
$\text{JET}^t_{M=6}$ ◇	63.44	54.12	59.41	53.53	43.28	47.86	**68.20**	42.89	52.66	65.28	51.95	63.83
$\text{JET}^o_{M=6}$ ◇	70.56	55.94	62.40	55.39	47.33	51.04	64.45	51.96	57.53	70.42	58.37	63.83
SPAN ◇	72.89	70.89	71.85	63.44	55.84	59.38	62.18	64.45	63.27	69.45	71.17	70.26
EMC-GCN ▽	71.21	**72.39**	71.78	61.70	56.26	58.81	61.54	62.47	61.93	65.62	71.30	68.33
RLI (Ours)	**77.46**	71.97	**74.34**	63.32	57.43	**60.96**	60.08	**70.66**	**65.41**	**70.50**	**74.28**	**72.34**

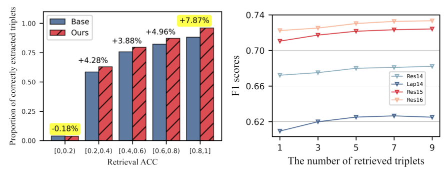
Ablation test

Dataset	Model		Dev F1
Res14	w/o joint		66.85
	joint	w/o sentiment	67.55
		w/o pre-training	67.12
		full model	**68.00**
Lap14	w/o joint		60.03
	joint	w/o sentiment	61.90
		w/o pre-training	61.06
		full model	**62.55**
Res15	w/o joint		70.83
	joint	w/o sentiment	71.54
		w/o pre-training	71.24
		full model	**72.21**
Res16	w/o joint		70.47
	joint	w/o sentiment	71.44
		w/o pre-training	70.75
		full model	**73.04**

Inference Results Analysis



Sensitivity Analysis



Auxiliary Experiment

Models	Base	Base+Aug	Ours
Avg. F1	66.19	66.99(+0.80)	**68.26**(+2.07)