Prior-BERT and Multi-Task Learning for Target-Aspect-Sentiment Joint Detection

Cai Ke, Qingyu Xiong*, Chao Wu, Zikai Liao, Hualing Yi

CHONGQING UNIVERSITY

icassp 2022 Singapore, China, Online

ID: 3638

Abstract

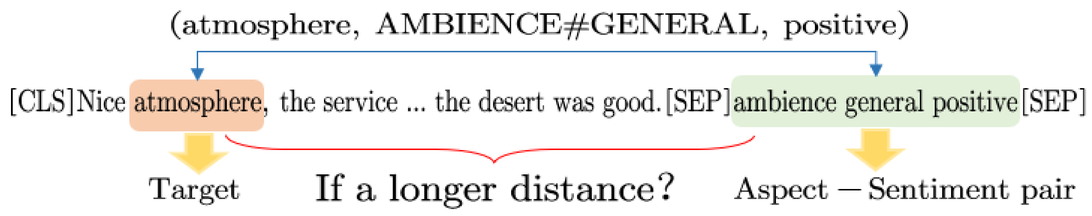
Aspect-Based Sentiment Analysis (ABSA) is a fine-grained sentiment analysis task and has become a significant task with real-world scenario value. The challenge of this task is how to generate an effective text representation and construct an end-to-end model that can simultaneously detect (target, aspect, sentiment) triples from a sentence. Besides, the existing models do not take the heavily unbalanced distribution of labels into account and also do not give enough consideration to long-distance dependence of targets and aspect-sentiment pairs. To overcome these challenges, we propose a novel end-to-end model named Prior-BERT and Multi-Task Learning (PBERT-MTL), which can detect all triples more efficiently. We evaluate our model on SemEval-2015 and SemEval-2016 datasets. Extensive results show the validity of our work in this paper. In addition, our model also achieves higher performance on a series of subtasks of target-aspect-sentiment detection. Code is available at https://github.com/CQUPTCaiKe/PBERT-MTL.

Motivation

◆There are **heavily unbalanced distribution** of labels after they reformulate datasets, which causes the model learning biased towards dominant labels.

Datasets		Aspects	Original			Reformulation			Implicit Targets
			sentences	yes	no	sentences	yes	no	
Res15	Train	13	1315	1	2	43642	1	38	375 (22.67%)
	Test	13	685	1	2	22660	1	38	248 (29.35%)
Res16	Train	12	2000	1	2	61453	1	35	627 (25.01%)
	Test	12	676	1	2	21097	1	35	208 (24.21%)

◆The **dependence of targets and aspect-sentiment pairs** cannot be resolved in a longer distance.



Methods

◆We propose **Prior-BERT (PBERT) — a simple but universal method** combining prior distribution knowledge of datasets with BERT for heavily unbalanced datasets. In detail, the probability distribution vector $g \in R^2$ on the "yes/no" label is defined below:

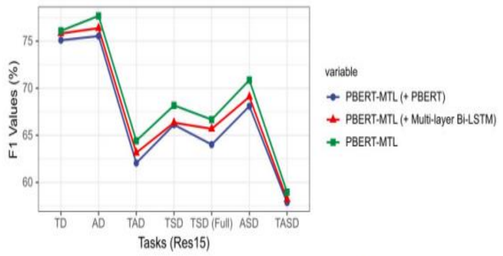
$$g = softmax(P_{[CLS]} + \tau \cdot \log P(y_i))$$

where τ is a tuning parameter to calibrate $P_{[CLS]}$ and y_i is the i^{th} element of $y \in \{yes, no\}$.

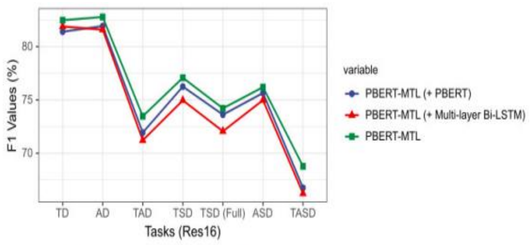
◆We propose **a novel end-to-end multi-task joint detection model (PBERT-MTL)** to usefully address the challenges of the TASD task.

Experimental Results

Tasks	Methods	Res15	Res16
TD	MTNA [21]	67.73	72.95
	DE-CNN [22]	-	74.37
	THA-STN [4]	71.46	73.61
	BERT-PT [23]	73.15	77.97
	TAS-BERT [10]	75.00	81.37
	PBERT-MTL	**75.66**±0.76	**82.16**±0.23
AD	BERT-pair-NLI-B [24]	70.78	80.25
	MTNA [21]	65.97	76.42
	TAN [6]	-	78.38
	Sentic LSTM+TA+SA [5]	73.82	-
	TAS-BERT [10]	76.34	81.57
	PBERT-MTL	**77.14**±0.55	**82.34**±0.47
TAD	TAS-BERT [10]	63.37	71.64
	PBERT-MTL	**64.21**±0.57	**72.97**±0.24
TSD	E2E-TBSA [25]	53.00	63.10
	DOER [26]	56.33	65.91
	TAS-BERT [10]	66.11 (64.29)	75.68 (72.92)
	PBERT-MTL	**67.53**±0.71(**66.12**)±0.67	**76.44**±0.14(**74.01**)±0.35
ASD	Baseline-1-f_lex [9]	-	63.50
	BERT-pair-NLI-B [24]	63.67	72.70
	TAS-BERT [10]	68.50	74.12
	PBERT-MTL	**70.43**±0.29	**75.88**±0.12
TASD	Baseline-1-f_lex [9]	-	38.10
	TAS-BERT [10]	57.51	65.89
	PBERT-MTL	**58.52**±0.23	**67.65**±0.34



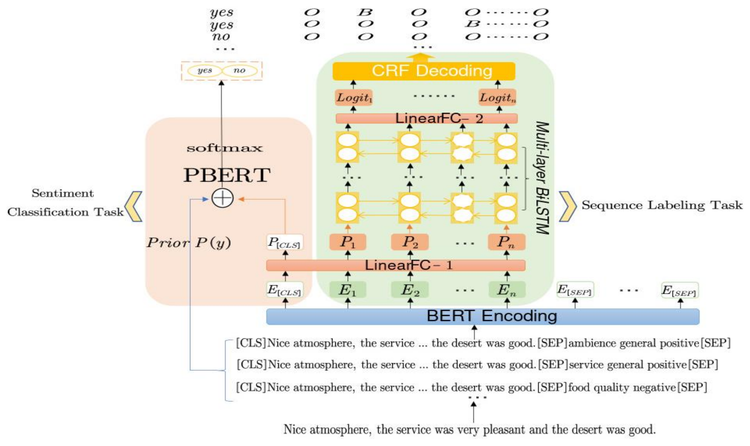
(a) Ablation experimental results on Res15 for each task.



(b) Ablation experimental results on Res16 for each task.

Text	Gold	Method	Prediction	Type
it was romantic - and even nice even with my sister, reminded me of italy, and had **artwork** and **music** that kept up the feeling of being in a Mediterrean villa.	{artwork, AMBIENCE#GENERAL, positive} {music, AMBIENCE#GENERAL, positive}	TAS-BERT	{music,AMBIENCE#GENERAL,positive}	Error-1
		Our method	{artwork,AMBIENCE#GENERAL,positive}	Correct
			{music,AMBIENCE#GENERAL,positive}	Correct
the best **place** for a leisure sunday breakfast amidst yachts, then take a stroll through the nearby farmer's market.	{place, RESTAURANT #MISCELLANEOUS, positive}	TAS-BERT	{NULL,NULL,NULL}	Error-2
		Our method	{place,RESTAURANT#MISCELLANEOUS,positive}	Correct

Model Architecture



Conclusions

◆We propose **a novel end-to-end multi-task model named PBERT-MTL** for TASD task which utilizes **the proposed PBERT method** to alleviate heavily unbalanced labels distribution and **the multi-layer Bi-LSTM** to capture the long distance dependence.

◆Experiments on Res15 and Res16 demonstrate that our model can detect (target, aspect, sentiment) triples efficiently and **achieve higher performance on the TASD task and its subtasks**.