Prior-BERT and Multi-Task Learning for Target-Aspect-Sentiment Joint Detection
Cai Ke, Qingyu Xiong*, Chao Wu, Zikai Liao, Hualing Yi

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.

<table>
<thead>
<tr>
<th>Original</th>
<th>Reformulation</th>
<th>Impact Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>sentences yes no</td>
<td>sentences yes no</td>
<td>375 (22.67%)</td>
</tr>
<tr>
<td>Train 13</td>
<td>1315 1 2</td>
<td>43642 1 38</td>
</tr>
<tr>
<td>Test 13</td>
<td>683 1 2</td>
<td>22060 1 38</td>
</tr>
<tr>
<td>Res15</td>
<td>248 (29.35%)</td>
<td></td>
</tr>
<tr>
<td>Train 12</td>
<td>2000 1 2</td>
<td>61453 1 35</td>
</tr>
<tr>
<td>Test 12</td>
<td>676 1 2</td>
<td>21097 1 35</td>
</tr>
<tr>
<td>Res16</td>
<td>208 (24.21%)</td>
<td></td>
</tr>
</tbody>
</table>

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

(atmosphere, AMBIENCE#GENERAL, positive)

[CLS]Nice atmosphere, the service ... the desert was good [SEP]ambient general positive[SEP]

Target: If a longer distance? Aspect-Sentiment pair

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 \( g \in \mathbb{R}^2 \) on the "yes/no" label is defined below:

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

where \( \tau \) is a tuning parameter to calibrate \( P_{\text{CLS}} \) and \( y_i \) is the \( i \)-th element of \( y \in \{\text{yes, no}\} \).

◆ 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

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

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

Conclusions
◆ We propose a novel end-to-end multi-task model named PBERT-MTL for TASD task which utilizes the proposed PBE RT 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.