Interactive Attention Transfer Network for Cross-domain Sentiment Classification
Kai Zhang1, Hefu Zhang1, Qi Liu1,2, Hongke Zhao1, Hengshu Zhu1, Enhong Chen1,2
1Anhui Province Key Laboratory of Big Data Analysis and Application, University of Science of Technology of China
2School of Data Science, University of Science and Technology of China

Introduction

Motivation
- Sentiment analysis heavily relies on large-scale corpora for model training and usually perform well on label-rich data.
- The Cross-domain sentiment analysis can reduce the reliance on the massive amounts of labeled data.
- However, most of the previous efforts mainly concentrate on extracting common features and ignore the characteristics which do not express the sentiment directly, such as the individual modeling of the aspect.
- Thus, fully considers the effects of the aspect (e.g., the battery life in reviewing an electronic product) information in the sentences is necessary.

In order to better solve this problem, we propose an Interactive Attention Transfer Network (IATN) for cross-domain sentiment classification.

Challenges
- What is the key information in the sentence that we desire?
- How do we model the most of this critical information? How do we model them together with sentences?
- How to verify that the information we extracted is indeed valid for the task.

IATN Model
- IATN models sentences and aspects independently and conducts an interactive attention learning to them.
- S-net: Learn shared information between different domains.
- Pooling the sentence feature representation vectors to A-net.
- A-net: Mining and modeling the aspect information.
- Use aspect information as the auxiliary task of domain classification.
- S-net and A-net have similar structures.

\[ e_{ia} = \{ e_{ia1}, e_{ia2}, \ldots, e_{ian} \} \quad e_{ia} = \{ e_{ia1}, e_{ia2}, \ldots, e_{ian} \} \]

\[ a_{i} = \frac{\exp(e_{ia})}{\sum_{j=1}^{n} \exp(e_{ij})} \]

\[ s_{i} = \sum_{j=1}^{n} a_{ij} \cdot e_{ij} \]

\[ y_{i} = \text{softmax}(W_{s}S + b_{s}) \]

\[ y_{i} = \text{softmax}(W_{a} \cdot \{ S \cup A \}) + b_{a} \]

\[ \mathbf{m} = \left[ m_{s1}, m_{a1}, \ldots, m_{sn}, m_{an} \right] \]

Framework

Flowchart

Problem Definition: Given two domains, \( D_s \) is the source domain and \( D_t \) is the target domain. Denote that each item, e.g., review, at both domains consists of \( n \) words marked as \( \{ w_1, w_2, \ldots, w_n \} \) and their aspect sequence contains \( m \) words marked as \( \{ s_1, s_2, \ldots, s_m \} \). The goal: To train a robust model based on labeled data in source domain and predict the unlabeled data in target domain.

IATN Model
- IATN models sentences and aspects independently and conducts an interactive attention learning to them.
- S-net: Learn shared information between different domains.
- Pooling the sentence feature representation vectors to A-net.
- A-net: Mining and modeling the aspect information.
- Use aspect information as the auxiliary task of domain classification.
- S-net and A-net have similar structures.

Experiments

Datasets
- We choose the reviews data from four domains: Book (B), DVD (D), Electronics (E) and Kitchen appliances (K).
- Each domain contains 6,000 labeled data, in which there are 3,000 positive reviews and 3,000 negative reviews.
- For each sentence, we extract their aspect sequence and conduct the cross-domain experiments between every two domains, which means we have 12 classification tasks.

Table 1: Sentiment classification accuracy on Amazon dataset.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc</td>
<td>0.765</td>
<td>0.784</td>
<td>0.772</td>
<td>0.704</td>
<td>0.699</td>
<td>0.759</td>
<td>0.678</td>
<td>0.706</td>
<td>0.730</td>
<td>0.748</td>
<td>0.694</td>
<td>0.724</td>
<td>0.755</td>
<td>0.666</td>
</tr>
<tr>
<td>Prec</td>
<td>0.456</td>
<td>0.495</td>
<td>0.479</td>
<td>0.385</td>
<td>0.366</td>
<td>0.428</td>
<td>0.355</td>
<td>0.385</td>
<td>0.397</td>
<td>0.412</td>
<td>0.366</td>
<td>0.385</td>
<td>0.404</td>
<td>0.366</td>
</tr>
<tr>
<td>F1-score</td>
<td>0.553</td>
<td>0.587</td>
<td>0.564</td>
<td>0.506</td>
<td>0.474</td>
<td>0.602</td>
<td>0.530</td>
<td>0.555</td>
<td>0.567</td>
<td>0.586</td>
<td>0.534</td>
<td>0.554</td>
<td>0.576</td>
<td>0.534</td>
</tr>
</tbody>
</table>

Conclusion

- We propose a novel IATN method which utilizes the interactive attention mechanism to get important information from both the sentence and aspect.
- We conduct extensive experiments on two real-world datasets.
- The experimental results validate that our method outperforms other state-of-the-art methods.

Future
- Add aspect information to the hierarchical attention network.
- Do some sentiment transfer learning research at the aspect-level sentiment classification.