1. Introduction

Goal
• Finding a good way to reuse Weibo comments to respond to new posts

Approach
• Retrieval-based method

Results
• Our run is statistically indistinguishable from the lowest performing run

2. System

Tools
• Jieba Segmentation
• Baidu Stopwords
• Word2Vec

3. Method

• Word2Vec representation
• Select the most similar Post (cosine sim)
• Through all pair comments, rank them with their score $S(c)$

$$w(c) = \frac{l_1(c) + l_2(c) + l_3(c)}{3} + 1.$$ (label score)

$$S(c) = w(c) \text{sim}(q, c) + \ln \sum_{t \in c} \text{eft}(t).$$ (term frequency)

4. Conclusion

• The differences between our run and the lowest performer are not statistically significant
• Comment ranking function may need to be improved
• Our run was not good enough

<table>
<thead>
<tr>
<th>Run</th>
<th>Mean nG@1</th>
<th>Mean P+</th>
<th>Mean nERR@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG01-C-G1</td>
<td>0.5867†</td>
<td>0.6670†</td>
<td>0.7095†</td>
</tr>
<tr>
<td>SLSTC-C-R1</td>
<td>0.0750</td>
<td>0.1171</td>
<td>0.1148</td>
</tr>
<tr>
<td>ckip-C-G1</td>
<td>0.0017</td>
<td>0.0029</td>
<td>0.0015</td>
</tr>
</tbody>
</table>

5. References

[1] Z. Chen, R. Song, and X. Xie. at NTCIR-12 STC task