1. The Task

- We target at document-level sentiment classification
  ‒ Given a document as input, the task is to predict the overall sentiment polarity

2. Basic Idea

- The basic idea
  ‒ Users who write the review and products which are evaluated are important for inferring the sentiment of a text
  ‒ We develop a deep learning method to use semantics of texts, users and products for sentiment classification

3. The Approach

- Softmax
- Pooling
- Convolution
- Tanh
- Linear
- Lookup

4. Experiment

- We derive 3 datasets from IMDB and Yelp Dataset Challenge 2013 and 2014.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#users</th>
<th>#products</th>
<th>#reviews</th>
<th>#docs/user</th>
<th>#docs/product</th>
<th>#sents/doc</th>
<th>#words/doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMDB</td>
<td>1,310</td>
<td>1,635</td>
<td>84,919</td>
<td>64.82</td>
<td>51.94</td>
<td>16.08</td>
<td>394.6</td>
</tr>
<tr>
<td>Yelp 2014</td>
<td>4,818</td>
<td>4,194</td>
<td>231,163</td>
<td>47.97</td>
<td>55.11</td>
<td>11.41</td>
<td>196.9</td>
</tr>
<tr>
<td>Yelp 2013</td>
<td>1,631</td>
<td>1,633</td>
<td>78,966</td>
<td>48.42</td>
<td>48.36</td>
<td>10.89</td>
<td>189.3</td>
</tr>
</tbody>
</table>

- Semantic representations of users and products can improve classification accuracy.
- User/product vectors are more powerful than user/product matrices.
- User representations are more useful than product representations.

Datasets and codes are available at [http://ir.hit.edu.cn/~dytang/](http://ir.hit.edu.cn/~dytang/)