Social Media

• Social networking sites allow sharing of opinions and viewpoint
• Express opinion on trending topics – Twitter
• Share your life experiences – wordpress, blogspot
• Community oriented interaction – Facebook
• Online shopping sites – amazon, flipkart
Online Debates

- Online debating sites – convinceme.net, 4forums.net, onlinedebate.net
- Provide users an option to express opinions about their favorite debate topics
- Rich collection of differing opinions
- Multi-party conversation
Debate Structure

• Use of debate topic supporting/opposing opinions
• Rebut another user
• Use of factual statements
• Dynamic nature and large number of debate posts makes summarization essential
• This work is the first in the field of debate summarization
Our Approach

- Extractive summaries created by ranking Dialogue Acts (DAs) in a debate

- Sentence Ranking function is used to calculate the importance of DAs using several features.

\[
\text{score}(s) = \lambda_{\text{topicRel}} \text{topicRel}(s, \text{topics}) + \lambda_{\text{docRel}} \text{docRel}(s, D) \\
+ \lambda_{\text{sentiRel}} \text{sentiRel}(s) + \lambda_{\text{conRel}} \text{conRel}(s, D)
\]
# Summarization Features

<table>
<thead>
<tr>
<th>Feature Category</th>
<th>Feature Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic Relevance</td>
<td>Topic Directed Sentiment Score</td>
</tr>
<tr>
<td></td>
<td>Topic Co-occurrence</td>
</tr>
<tr>
<td>Document Relevance</td>
<td>tf-idf Sentiment Score</td>
</tr>
<tr>
<td>Sentiment Relevance</td>
<td>Number of Sentiment Words</td>
</tr>
<tr>
<td></td>
<td>Sentiment Strength</td>
</tr>
<tr>
<td>Context Relevance</td>
<td>Sentence Position</td>
</tr>
<tr>
<td></td>
<td>Sentence Length</td>
</tr>
</tbody>
</table>
Topic Relevant Features

• Debate posts present users’ opinion on debate topics

• Topic relevant features are most important in the context of debate summarization
Topic Relevant Features

• Topic Co-occurrence – Measures the co-occurrence of the DAs related to the debate topics

\[
Co-occurrence_{DA} = \sum_{w \in DA} \left( \sum_{t \in ET} (HAL(w|t)) \times sentiScore(w) \right)
\]

• Calculates the score taking into account DA words’ co-occurrence with debate topics and its sentiment strength
Topic Relevant Features

• Topic Directed Sentiment Scores – Capturing sentiment directed towards the debate topics in DAs

• Sentiment strength of DA words are captured using SentiWordNet

• Stanford dependency relations are used to identify topic directed subjective words
Document Relevant Features

- tf-idf Sentiment Score – very common feature used in text summarization approaches
- Captures highly document relevant sentences

\[ tf - idf_{DA} = \sum_{w \in DA} (tf - idf(w) \times \text{sentiScore}(w)) \]
Sentiment Relevant Features

• Sentiment carrying words are important in opinion based online debates

• Number of sentiment words

• Sentiment score of the DA – aggregate of sentiment score of each DA word
Context Relevant Features

• Sentence Position – Sentence in the beginning and end are more important than those in the middle

\[ posScore_{DA} = \left| \frac{N}{2} - DA_{position} \right|, \quad N = Total \ DAs \ in \ Post \]

• Sentence Length – Longer sentences tend to contain more information

• count of words in Das are used as an feature to avoid shorter sentences
Evaluation Data

- 10 online debates from convinceme.net

<table>
<thead>
<tr>
<th>Number of users</th>
<th>Number of posts</th>
<th>Number of DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1168</td>
<td>1945</td>
<td>23681</td>
</tr>
</tbody>
</table>

- Extractive summaries of 500, 1000, 1500, 2000 words created by language editors
- 71.7% inter-editor agreement
Results

• Grid search method used to calculate best parameter values of feature weights.

• Best Values with - $\lambda_{\text{topicRel}} = 0.3$, $\lambda_{\text{docRel}} = 0.1$, $\lambda_{\text{sentiRel}} = 0.5$, $\lambda_{\text{conRel}} = 0.1$
Baselines Compared

1. Max-Length – Longest sentences from all users are selected to create summaries
2. Lead – Top sentences from each user are selected
3. pHAL – Sentence scores are calculated using pHAL scores of sentence words and sentences with highest pHAL scores are selected

\[ pHAL(w) = \sum_{w' \in ET} \frac{HAL(w'|w)}{n(w) * K} \]
Baselines Compared

4. Tf-idf – tf-idf scores of sentences words are aggregated and highest scored sentence are selected

5. OpinionSumm: Opinion summarization approach by wang et. al. Score is based on document similarity, topic relevance, sentiment and sentence length
# System Performance Comparison

- ROGUE scores of System summaries (1000 words)

<table>
<thead>
<tr>
<th>System</th>
<th>ROUGE-1</th>
<th>ROUGE-2</th>
<th>ROUGE-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max-Length</td>
<td>0.49892</td>
<td>0.18453</td>
<td>0.48343</td>
</tr>
<tr>
<td>Lead</td>
<td>0.49068</td>
<td>0.14759</td>
<td>0.47839</td>
</tr>
<tr>
<td>pHAL</td>
<td>0.48985</td>
<td>0.16468</td>
<td>0.46955</td>
</tr>
<tr>
<td>tf-idf</td>
<td>0.49922</td>
<td>0.17585</td>
<td>0.48035</td>
</tr>
<tr>
<td>OpinionSumm</td>
<td>0.51631</td>
<td>0.20364</td>
<td>0.49849</td>
</tr>
<tr>
<td>DEBSumm</td>
<td><strong>0.56833</strong></td>
<td><strong>0.27044</strong></td>
<td><strong>0.55326</strong></td>
</tr>
</tbody>
</table>
System Comparison over Summary Size

![Graph showing comparison of summary sizes for different systems.](image)
Conclusion

• Summarization on the basis of topic directed sentiment and topic related information

• Results show DESBSumm outperforms other approach and topic directed sentiment analysis based approach helps in better summarization.

• Generic Text summarization features like Sentence length, tf-idf, sentence position work well for this task.
Future Work

• Word Sense Disambiguation

• User profiling for leveraging users’ DAs

• Domain specific sentiment identification is required. ‘refined’ is good for ‘oil products’ but not in case of ‘agricultural products’
THANK YOU!