Online forums

- Discussions online
  - About technical matters (e.g., Stackoverflow)
  - In response to an article
- Can be found
  - On technical sites
  - On social media sites like Facebook, Twitter
  - On newspaper sites, in particular the Guardian
Technical forums

Newspaper article comment threads
Rubbish? Bury council votes to collect wheelie bins just once every three weeks

Locals fear the new move will lead to an increase in fly tipping and attract foxes and vermin, but the council insists it will make the borough more environmentally friendly. Is it just a desperate cost cutting measure? Lyndsay Warner on a controversial decision

A council in Greater Manchester is to be the first in England to start collecting wheelie bins only once every three weeks, scrapping the current fortnightly collection.

The controversial decision was unanimously passed by councillors in Bury on Wednesday night, despite fears fly tipping would increase. One councillor who voted for the motion accused her opponents of “scaremongering” after they warned rubbish would pile up and attract vermin. Another argued the money saved could be spent on more social workers.

More than 5,100 people signed a petition opposing the new initiative, which will be introduced on 6 October.

It affects the grey bins used for general household waste which can’t be recycled. The green and blue coloured bins for paper, glass and plastic will be collected more often than they are currently - every three weeks instead of monthly. The emptying of all communal bins serving flats will not change.

Text analytics and online forums

• Extracting information

• Summarizing

• Predicting

Summarizing online forums
Summarizing online forums: what to summarize?

The Online Forums Summarization (OnForumS) Shared Task (2015, 2017)

We carried out a number of tests where each participant tried to make a summary of a news article and reader’s comments.

Three aspects were highlighted:

1. **Summary of the news article** (e.g., key article sentences (KAS))
2. **(Labelled) Links between comments and key article sentences**
3. **Summary of comments** (e.g., key comment sentences (KCS))
The OnForumS task

• Systems have to extract a simplified form of Argument Graph from the sentences in a text and its comments
• For each sentence, they have to tell whether it’s related by an agreement/disagreement link to a previous sentence; and if so, if the link is AGREE or DISAGREE

The OnForumS data

• Large data dumps have been created by Websays for English and Italian
  • For English: a corpus of ten news articles from the Guardian newspaper was extracted from data dumps:
    • Around 40K words
    • Pre-tokenised and sentence-split
    • Top 50 comments chosen based on number of likes
• For Italian: a corpus of six blog articles from the La Repubblica newspaper was extracted:
  • Around 40K words
  • Pre-tokenised and Sentence split
  • Top 50 comments
OnForumS 2015: Submissions

- Four groups participated, submitting two runs each
  - USFD & UNITN
  - EC’s Joint Research Centre (JRC)
  - Beijing University of Posts and Telecommunication (CIST)
  - University of West Bohemia (UWB)

- Two baselines:
  - First: links every comment sentence to first sentence of article or parent comment in case of comment replies
  - Overlap: links a comment sentence to the article (or parent comment) sentence with the most common words (stopwords are removed)
  - Both use the same approach, lexicon-based, for argument (Inquirer) and sentiment (MPQA lexicon)

---

Evaluation

- Judge system’s output on the relevance between reader’s comments and news article sentences
  - Snippet: "While it is almost all circumstances, dredging reveals more problems than it solves, and though, as valve..."
  - Comment: "…broader picture. But in the end, the data in the..."
  - Would you agree that the highlighted sentence in the comment (orange) is related to the highlighted sentence from the article snippet (blue)?
    - Yes
    - No
  - What is the comment’s stance (orange) with respect to the sentence in blue in the snippet? (Use ‘Not Applicable’ if you answered ‘no’ to the first question)
    - In favour
    - Against
    - Not applicable
  - What is the comment’s sentiment (orange) with respect to the sentence in blue in the snippet?
    - Positive
    - Neutral
Online Forum Summarization (OnForumS): Evaluation (en)

• Just like with IR systems, only Precision can be computed directly from the crowdsourcing evaluation (by counting yes’s and no’s for each dimension, linking, argument and sentiment)

<table>
<thead>
<tr>
<th>System run</th>
<th>Prec (Link)</th>
<th>System run</th>
<th>Prec (Argument)</th>
<th>System run</th>
<th>Prec (Sentiment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE-overlap</td>
<td>0.931039481</td>
<td>CIST-run2</td>
<td>0.9929205654</td>
<td>CIST-run1</td>
<td>0.950810672</td>
</tr>
<tr>
<td>USFD_UNITN-run2</td>
<td>0.887096774</td>
<td>CIST-run1</td>
<td>0.99207429</td>
<td>CIST-run2</td>
<td>0.99285714</td>
</tr>
<tr>
<td>UWB-run1</td>
<td>0.864583333</td>
<td>UWB-run1</td>
<td>0.975 BASE-overlap</td>
<td>BASE-run1</td>
<td>0.9375</td>
</tr>
<tr>
<td>UWB-run2</td>
<td>0.864583333</td>
<td>UWB-run2</td>
<td>0.975 BASE-first</td>
<td>CIST-run1</td>
<td>0.914895833</td>
</tr>
<tr>
<td>IRC-run1</td>
<td>0.862068896</td>
<td>BASE-first</td>
<td>0.926701571</td>
<td>USFD_UNITN-run2</td>
<td>0.925925926</td>
</tr>
<tr>
<td>IRC-run2</td>
<td>0.83125301</td>
<td>IRC-run1</td>
<td>0.907063197</td>
<td>IRC-run2</td>
<td>0.902985075</td>
</tr>
<tr>
<td>USFD_UNITN-run1</td>
<td>0.818731118</td>
<td>USFD_UNITN-run1</td>
<td>0.893939394</td>
<td>USFD_UNITN-run1</td>
<td>0.89113208</td>
</tr>
<tr>
<td>USFD_UNITN-run2</td>
<td>0.708547254</td>
<td>USFD_UNITN-run2</td>
<td>0.862385321</td>
<td>IRC-run1</td>
<td>0.878612717</td>
</tr>
</tbody>
</table>

Online Forum Summarization (OnForumS): Evaluation (it)

• And for Italian...
Online Forum Summarization (OnForumS): Creating Gold Data

- From the validated links by crowdsourcing, there are two ways for creating golden data (non-exhaustive):

  - A) By validation: any link and labels that have validation="yes" (aggregated) effectively are gold labels.
  - B) By exclusion: if the same link (i.e., pair of sentences) have had three out of four labels validated with "no", then the remaining fourth label is taken as gold label.

Using XML input files and the aggregated labels from CrowdFlower, gold XML files were created and released at:


Online Forum Summarization (OnForumS): P/R/F1 Argument

- Having a gold standard enables computing P/R/F1 measures per label:

<table>
<thead>
<tr>
<th>IN_FAVOUR</th>
<th>AGAINST</th>
<th>IMPARTIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>Recall</td>
<td>Recall</td>
</tr>
<tr>
<td>GroupAndRun</td>
<td>GroupAndRun</td>
<td>GroupAndRun</td>
</tr>
<tr>
<td>CST-run1</td>
<td>83.08823529 UWB-run2</td>
<td>2.941176471 BASE-overlap</td>
</tr>
<tr>
<td>CST-run2</td>
<td>55.14705882 BASE-first</td>
<td>0 JRC-run1</td>
</tr>
<tr>
<td>UWB-run1</td>
<td>21.52294118 BASE-overlap</td>
<td>0 JRC-run2</td>
</tr>
<tr>
<td>BASE-overlap</td>
<td>10.29411765 CST-run2</td>
<td>0 CST-run1</td>
</tr>
<tr>
<td>JRC-run1</td>
<td>1.470588235 IRC-run1</td>
<td>0 CST-run2</td>
</tr>
<tr>
<td>JRC-run2</td>
<td>1.470588235 IRC-run2</td>
<td>0 USFD_UNITN-run1</td>
</tr>
<tr>
<td>BASE-first</td>
<td>0 USFD_UNITN-run1</td>
<td>0 USFD_UNITN-run2</td>
</tr>
<tr>
<td>UWB-run1</td>
<td>0 USFD_UNITN-run1</td>
<td>0 UWB-run1</td>
</tr>
<tr>
<td>USFD_UNITN-run2</td>
<td>0 UWB-run1</td>
<td>0 UWB-run2</td>
</tr>
</tbody>
</table>
Online Forum Summarization (OnForumS): P/R/F1 Sentiment

<table>
<thead>
<tr>
<th>GroupAndRun</th>
<th>Recall</th>
<th>GroupAndRun</th>
<th>Recall</th>
<th>GroupAndRun</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIST-run1</td>
<td>100</td>
<td>UWB-run2</td>
<td>6.66666667</td>
<td>IRC-run1</td>
<td>14.44444444</td>
</tr>
<tr>
<td>CIST-run2</td>
<td>63.3333333</td>
<td>BASE-first</td>
<td>0</td>
<td>IRC-run2</td>
<td>14.44444444</td>
</tr>
<tr>
<td>BASE-overlap</td>
<td>13.3333333</td>
<td>BASE-overlap</td>
<td>0</td>
<td>BASE-overlap</td>
<td>3.333333333</td>
</tr>
<tr>
<td>UWB-run1</td>
<td>10</td>
<td>CIST-run1</td>
<td>0</td>
<td>UWB-run1</td>
<td>3.333333333</td>
</tr>
<tr>
<td>UWB-run2</td>
<td>6.66666667</td>
<td>CIST-run2</td>
<td>0</td>
<td>UWB-run2</td>
<td>3.333333333</td>
</tr>
<tr>
<td>BASE-first</td>
<td>0</td>
<td>IRC-run1</td>
<td>0</td>
<td>BASE-first</td>
<td>0</td>
</tr>
<tr>
<td>IRC-run1</td>
<td>0</td>
<td>IRC-run2</td>
<td>0</td>
<td>CIST-run1</td>
<td>0</td>
</tr>
<tr>
<td>IRC-run2</td>
<td>0</td>
<td>USFD_UNITN-run1</td>
<td>0</td>
<td>CIST-run2</td>
<td>0</td>
</tr>
<tr>
<td>USFD_UNITN-run1</td>
<td>0</td>
<td>USFD_UNITN-run2</td>
<td>0</td>
<td>USFD_UNITN-run1</td>
<td>0</td>
</tr>
<tr>
<td>USFD_UNITN-run2</td>
<td>0</td>
<td>UWB-run1</td>
<td>0</td>
<td>USFD_UNITN-run2</td>
<td>0</td>
</tr>
</tbody>
</table>

P/R Scatter Plots
OnForumS 2017 (ongoing)

- Using the OnForumS 2015 data for training, new data for testing
- Six systems participating
- Results should be known in a week or two

Voting predictions using online forums
Using text analytics to assess voters’ opinions

• Text Analytics method make it possible to analyse a massive numbers of posts on social media websites such as Twitter or Facebook.
  • (A much larger number of opinions is canvassed than using traditional polling methods)
• The posts about topics of interest are found (the simplest way to do this is by looking for tweets with a particular hashtag, such as #Brexit, although most sites use TEXT CATEGORIZATION techniques).
• What happens next depends on the site. The most popular approach is to use SENTIMENT ANALYSIS to classify the relevant posts according to their positive or negative sentiment towards that topic.

Using agreement structure to correct sentiment analysis bias

• The problem with the methods using only sentiment analysis is that they tend to massively overestimate negative opinion
  • E.g., prior to the Brexit referendum, most sentiment analysis-based sites assessed the rate of negative opinion towards Europe at around 80%
• Our own site, SENSE-EU, uses a different method based on the assumption that posts on social media tend to be part of a CONVERSATION:
  • after classifying these posts as as favouring LEAVE or REMAIN
  • it classifies responses to these posts as being in agreement or disagreement with the statement in the post.
• These methods, of course, are by no means perfect: they often incorrectly assess the sentiment of a post, or whether the poster actually agrees or disagrees (e.g., when the poster is being ironic); nevertheless, the sheer volume of data analysed means that by and large CL-based automatic poll prediction sites tend to be a pretty accurate gauge of opinion towards a topic.
Predicting the outcome of the EU referendum

Extracting information from online forums
Extracting information from online forums

- Cong et al, 2008: Extracting Question/Answer pairs from online forums
  - Huang et al, 2007: Extracting such information for use by a chatbot
- Phillips, 2015: Extracting social structure from Dark Web forums
  - Using social network analysis methods to identify leading individuals in Islamic networks

Towards the exam

- We covered a lot of material; make sure you identify the most important topics and the not-so-important ones
  - A clue towards most important topics: we had labs on them
- Details (eg formula) only matter for key topics
  - E.g., TFIDF etc
References


  