Opinion Retrieval in Twitter
Zhunchen Luo (zhunchenluo@nudt.edu.cn)
Miles Osborne (miles@inf.ed.ac.uk)
Ting Wang (tingwang@nudt.edu.cn)

Introduction
• We consider the problem of finding opinionated tweets about a
given topic
E.g., Given a query: ICWSM

Relevant tweet:
We had some pretty imaginative papers titles at ICWSM-12 - what
was your favourite? bit.ly/GVQjuQ "OMG, I Have to Tweet That!"

Irrelevant tweet:

• We automatically construct opinionated lexica from sets of tweets
matching specific patterns indicative of opinionated messages
• We use a machine learning approach to learn a ranking function for
tweets that uses the available social features and the opinionated
feature
• Dataset: 50 queries and 5000 judged tweets

Opinionated Tweets
• “Pseudo” Subjective Tweet (PST): a tweet of the form “RT
@username” with text before the retweet
E.g., Amber Lyon @AmberLyon
Tragically not much has changed RT "@SAIwadee: BBC report on
Bahrain from 90s. History repeating itself. youtube.com/watch?v=4pzs..."

• “Pseudo” Objective Tweet (POT): If a tweet satisfies two criteria: (1)
 it contains links and (2) the user of this tweet posted many tweets
before and has many followers
E.g., BBC Breaking News @BBCBreaking
Minimum price of alcohol in Scotland proposed at 50p per unit -
Scottish Health Secretary Nicola Sturgeon bbc.in/1lmnV0

• A term can be measured how dependent with PST set and POT set

Examples of Opinion Terms

<table>
<thead>
<tr>
<th>Topic Independent</th>
<th>Sub</th>
<th>Obj</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i, lol, .., :) *, u, my, :d, me, morn</td>
<td>new, via, ..., video, tip, social, 2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Breaking Dawn</th>
<th>Sub</th>
<th>Obj</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i, go, me, lol, !!!, excit, im, :), so, too</td>
<td>video, premier, kristen, robert</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UK Strike</th>
<th>Sub</th>
<th>Obj</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i, you, my, lol, :, u, me, so, !, good</td>
<td>follow, fridai, week, bbc, #ows, #jobs</td>
</tr>
</tbody>
</table>

Features
• BM25: the Okapi BM25 score
• Mention: whether the tweet contains “@username”
• URL: whether the tweet contains a link
• Statuses: the number of tweets the author has ever written
• Followers: the number of followers
• Opinionatedness: the opinionatedness score of a tweet

Experiment

<table>
<thead>
<tr>
<th></th>
<th>MAP</th>
<th>MAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM25</td>
<td>0.2509</td>
<td>BM25+Followers 0.2532</td>
</tr>
<tr>
<td>BM25+Mention</td>
<td>0.2814</td>
<td>BM25+Opinionatedness 0.3667</td>
</tr>
<tr>
<td>BM25+URL</td>
<td>0.3380</td>
<td>Best (all features) 0.4020</td>
</tr>
<tr>
<td>BM25+Statuses</td>
<td>0.2726</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion
• Using the social and opinionatedness information of tweets can help
opinion retrieval in Twitter
• Our Dataset URL: http://sourceforge.net/projects/ortwitter/