Twitter Opinion Topic Model:
Extracting Product Opinions from Tweets by Leveraging Hashtags and Sentiment Lexicon

Presenter: Peter Christen
Authors: Kar Wai Lim, Wray Buntine

6 November 2014
Aspect-based Opinion Aggregation

• Opinion Aggregation for reviews.
  – A process to collect reviews of products and services to analyze in aggregate.

• Aspect-based.
  – Groups reviews based on “aspects”.
  – Example:
    • Product types
      – Game consoles
      – Mobile phones
    • Product specs
      – Computer specs
      – Flight quality

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game console</td>
<td>PS4, Xbox One, Wii U...</td>
</tr>
<tr>
<td>Mobile phone</td>
<td>iPhone, Samsung Note...</td>
</tr>
<tr>
<td>Computer spec</td>
<td>CPU, RAM, GPU...</td>
</tr>
<tr>
<td>Flight quality</td>
<td>Food, customer service...</td>
</tr>
</tbody>
</table>
Existing Method

• Independent Latent Dirichlet Allocation (ILDA).
  – Current state-of-the-art for aspect-based opinion aggregation (Moghaddam, 2012).
  – ILDA is a type of topic model.
  – Perform analysis on target-opinion pairs.

• Target-opinion pairs are extracted during preprocessing using Stanford dependency parser.
  – Examples:

<table>
<thead>
<tr>
<th>Target</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPhone</td>
<td>Awesome</td>
</tr>
<tr>
<td>Service</td>
<td>Good</td>
</tr>
<tr>
<td>Weather</td>
<td>Hot</td>
</tr>
</tbody>
</table>
Aspects and sentiments are discrete labels learned by the model.
ILDA

- Graphical model:

Note:
Shaded = observed
Unshaded = latent

They capture the interaction between the variables, tell us about the corpus.
ILDA

• What ILDA does:
  – Automatically groups target-opinion pairs into various aspects.
  – Learns the opinions corresponding to various sentiments.

• Limitation of ILDA:
  – Sentiment labels are arbitrary.
    • Need to manually inspect the associated opinions to know whether they are positive, neutral or negative.
  – Targets and opinions are related only via latent variables.
    • Interaction between targets and opinions are not considered.
    • The pair ‘friendly dumpling” is perfectly reasonable under ILDA.

* Picture stolen via Google search.
Opinion Aggregation on Tweets

**Why?**
- More opinions lying around, but less structured.
  - Easier to create than proper review.
- Less targeted by fake review companies.
  - Tweets are usually written for friends and family.

**How?**
- Design Twitter Opinion Topic Model (TOTM) for Tweets.
- Extension of ILDA but address its limitation.
- Make use of emoticons (common in Tweets).
- Use hashtags to aggregate Tweets.
- Incorporate existing sentiment lexicons.
TOTM

- Graphical model:
TOTM

- Graphical model:

Emotion indicator – determined by seen emoticons or strong sentiment words (such as ‘happy’, ‘sad’).

Positive opinions tend to associate with positive emotions.
Model target-opinion interaction directly. This improves opinion prediction significantly.
• Graphical model:

Incorporate sentiment lexicon as prior.
Incorporating Sentiment Lexicon

• Existing approach (He, 2012):
  – Rule-based system for topic models.
  – Modify the Dirichlet prior for opinion word distributions.
    • Note we have 3 opinion word distributions:
      – Positive-opinion distribution.
      – Neutral-opinion distribution.
      – Negative-opinion distribution.
    • The prior parameter is initialised as 0.33 for each opinion word of any sentiment-opinion distribution (uniform Dirichlet).
    • The prior parameter is then adjusted to 0.9 or 0.05 depending on the sentiment of a given opinion word (according to lexicon).
Incorporating Sentiment Lexicon

• Our approach:
  – Introduce a tunable parameter $b$ to control the strength of sentiment prior.
  – The prior for the sentiment-opinion distribution is given by:
    $$\phi_{rv}^* \propto (1 + b)^{X_{rv}}$$
    – $X_{rv}$ is the sentiment score of an opinion word determined from lexicon.
    – $b$ is strictly positive, so positive $X_{rv}$ enhances the prior while negative $X_{rv}$ lowers the prior. (see details in paper)
    – Why exponential in the formula?
      • Ensures positivity of the priors.
      • Gives a simple learning algorithm for $b$. 
Experiments

• Dataset:
  – Subset of Twitter 7 dataset (Yang & Leskovec, 2011).
    • 9 millions tweets on Electronic Products.
  – And 2 smaller corpus.

• Compare TOTM against
  – ILDA;
  – LDA-DP [Vanilla LDA but modify prior according to He (2012)].

• Evaluations:
  – Perplexity;
  – Sentiment prior evaluation;
  – Sentiment classification.
Perplexity

• Commonly used to evaluate topic models.
• Measure topic model’s goodness of fit.
  – Negatively related to log likelihood so lower perplexity is better.

<table>
<thead>
<tr>
<th></th>
<th>Target</th>
<th>Opinion</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDA-DP</td>
<td>N/A</td>
<td>510.15 ± 0.08</td>
<td>N/A</td>
</tr>
<tr>
<td>ILDA</td>
<td>594.81 ± 13.61</td>
<td>519.84 ± 0.43</td>
<td>556.03 ± 6.22</td>
</tr>
<tr>
<td>TOTM</td>
<td>592.91 ± 13.86</td>
<td><strong>137.42 ± 0.28</strong></td>
<td><strong>285.42 ± 3.23</strong></td>
</tr>
</tbody>
</table>

Better fit for opinion words by modelling the target-opinion interaction directly.
Qualitative Analysis

• Inspect the top words from target word distributions.

<table>
<thead>
<tr>
<th>Aspects (a)</th>
<th>Target Words (t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera</td>
<td>camera, pictures, video camera, shots</td>
</tr>
<tr>
<td>Apple iPod</td>
<td>ipod, ipod touch, songs, song, music</td>
</tr>
<tr>
<td>Android phone</td>
<td>android, apps, app, phones, keyboard</td>
</tr>
<tr>
<td>Macbook</td>
<td>macbook, macbook pro, macbook air</td>
</tr>
<tr>
<td>Nintendo games</td>
<td>nintendo, games, game, game, gameboy</td>
</tr>
</tbody>
</table>

• Inspect the top words from opinion word distributions.

<table>
<thead>
<tr>
<th>Target (t)</th>
<th>+/-</th>
<th>Opinions (o)</th>
</tr>
</thead>
<tbody>
<tr>
<td>phone</td>
<td>-</td>
<td>dead damn stupid bad crazy</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>mobile smart good great f***ing</td>
</tr>
<tr>
<td>battery life</td>
<td>-</td>
<td>terrible poor bad horrible non-existence</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>good long great 7hr ultralong</td>
</tr>
<tr>
<td>game</td>
<td>-</td>
<td>addictive stupid free full addicting</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>great good awesome favorite cat-and-mouse</td>
</tr>
<tr>
<td>sausage</td>
<td>-</td>
<td>silly argentinian cold huge stupid</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>hot grilled good sweet awesome</td>
</tr>
</tbody>
</table>

* Words in **bold** are more specific and can only describe certain targets.
Qualitative Analysis

• For comparison purpose, we can analyze hashtags that correspond to electronic companies such as #sony, #canon, #samsung...

<table>
<thead>
<tr>
<th>Brands</th>
<th>Sentiment</th>
<th>Aspects / Targets’ Opinions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Camera</td>
</tr>
<tr>
<td>Canon</td>
<td>–</td>
<td>camera → expensive small bad lens → prime cheap broken</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>camera → great compact amazing pictures → great nice creative</td>
</tr>
<tr>
<td>Sony</td>
<td>–</td>
<td>camera → big crappy defective lens → vertical cheap wide</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>photos → great lovely amazing camera → good great nice</td>
</tr>
<tr>
<td>Samsung</td>
<td>–</td>
<td>camera → digital free crazy shots → quick wide</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>camera → gorgeous great cool pics → nice great perfect</td>
</tr>
</tbody>
</table>
Major Contributions

• Introduce TOTM for aspect-based opinion aggregation on Tweets.
  – Makes use of auxiliary information on Tweets.

• Novel way of incorporating sentiment prior information into topic models.
  – Simple to implement and allow automatic learning of the hyperparameter (b).

Please email Kar Wai (karwai.lim@gmail.com) if you have any questions, thank you.