Generative Models for Mining Latent Aspects and Their Ratings from Short Reviews

Huayu Li\(^+\), Rongcheng Lin\(^+\), Richang Hong\(^*\), and Yong Ge\(^+\)

\(^+\)Computer Science Department, UNC Charlotte
\(^*\)Hefei University of Technology
Outline

- Introduction
- Related Work
- Methods
- Experiments
- Conclusion
Introduction

This is coooooool by Georgia Veres
Sep 04, 2015
I like this app a lot!! When I grow up I'm going to be a secret agent with 5 of my friends. This app is awesome!

GETTING SHARP
By Francofille on November 1, 2015
Verified Purchase
Some games very challenging but with practice I definitely improve. Have to work hard not to feel I have a low IQ! Takes awhile when you haven't some these in many years. But good brain training.

“great family place!”
Rated October 28, 2015
Great place to go with family! It is a small, quaint place but the staff are very friendly and the menu is laid out well. My girlfriend and I enjoyed lunch there...will definitely be back

Really good food, always fresh and tasty. The 1/4 chicken combo is what I usually get. Try the sauces!!! Love love them, specially the yellow. Also the Sangria mix drink is really good too.
Introduction

“Very Nice Place”
Reviewed October 3, 2015

It was near the hotel where I was staying in Charlotte for a convention, and had been to other restaurants and was looking for some place new to try. The staff was friendly and the food was very good. The pulled pork was some of the best I’ve had, and they had a nice selection of craft beer to order from. It is a short walk from Time Warner Arena and the hotels near there. A place to go for a nice relaxing meal and drink.

Visited September 2015

This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.
Introduction

1. What aspects is a review taking about?
2. How much does the user like each aspect?

“Very Nice Place”
Reviewed October 3, 2015

It was near the hotel where I was staying in Charlotte for a convention, and had been to other restaurants and was looking for some place new to try. The staff was friendly and the food was very good. The pulled pork was some of the best I've had, and they had a
Outline

- Introduction
- Related Work
- Methods
- Experiments
- Conclusion
Related Work

• Aspect (topic) extraction
  • Summarization based technique (AAAI’04, SIGKDD’04 and HLT’05)
  • Latent dirichlet allocation (LDA) (03’JMLR)
  • Global topic and local topic (WWW’08, ACL’08)
  • Opinion phrase based models (SIGIR’11, CIKM’12 and WWW’13)

• Aspect rating prediction
  • Supervised regression based approach (KDD’10)
  • Unsupervised regression based approach (LARAM)(KDD’11)
Outline

- Introduction
- Related Work
- Methods
- Experiments
- Conclusion
Problem Definition:

Given a collection of review texts and their overall ratings, how to extract the aspects they talk about and predict ratings on each aspect.

Methods: Task

- Extract Aspect
- Predict Aspect Rating
Methods : AIR – Observations

• **Intuition:**

When users have a latent high (or low) rating on one aspect of a product, they are more likely to comment on this aspect with positive (or negative) words.
Methods: AIR – Observations

• Intuition:

When users have a latent high (or low) rating on one aspect of a product, they are more likely to comment on this aspect with positive (or negative) words.

A Google User  September 6, 2012
★ ★ ★ ★ ★

Fantastic must have game is This game is great and everyone should have it. Smooth graphics, smooth control and its free. Everyone should have it.
Methods: AIR – Observations

- Intuition:
  When users have a latent high (or low) rating on one aspect of a product, they are more likely to comment on this aspect with positive (or negative) words.

A Google User September 6, 2012

★★★★★

Fantastic must have game is. This game is great and everyone should have it. Smooth graphics, smooth control and it's free. Everyone should have it.

Levi Wicks December 31, 2014

★★★★★

WORST GAME EVER!!!!!!!!!!!!!! This game is horrible!!!!!! It didn't work, and when you turn a corner, you didn't do anything wrong, and the monkeys are on your back, and then next time you turn a corner, you get eaten. I am uninstalling right now. HATE HATE HATE HATE HATE HATE HATE HATE it. DO NOT INSTALL THIS GAME!!!!!!!!!!!!!!!!!!!!!
Methods: AIR – Observations

- Relationship between overall rating and aspect rating

Figure 1: X-axis denotes different values of aspect rating and Y-axis is the frequency of them. (a)–(e) are the histograms with respect to different values of overall ratings. (f)–(j) are the histograms of ratings on Service aspect with respect to different values of overall ratings.
Methods: AIR – Observations

- Relationship between overall rating and aspect rating
  - The histograms of aspect ratings are very close to the Beta distribution.
  - The value of aspect rating with maximum frequency is always equal to the value of overall rating.

![Graphs showing aspect rating distributions for different overall ratings](image-url)

*Figure 1: X-axis denotes different values of aspect rating and Y-axis is the frequency of them. (a)~(e) are the histograms with respect to different values of overall ratings. (f)~(j) are the histograms of ratings on Service aspect with respect to different values of overall ratings.*
Methods: AIR – Observations

- Relationship between overall rating and aspect rating
  - The histograms of aspect ratings are very close to the Beta distribution.
  - The value of aspect rating with maximum frequency is always equal to the value of overall rating.

\[ \Omega_{i,k} \sim Beta(\lambda R_i, \lambda(1 - R_i)) \]

Figure 1: X-axis denotes different values of aspect rating and Y-axis is the frequency of them. (a)~(e) are the histograms with respect to different values of overall ratings. (f)~(j) are the histograms of ratings on Service aspect with respect to different values of overall ratings.
Methods : AIR

The graphical model

The generative process

Gibbs Sampling is used for parameter estimation.

1. For each review $i$,
   a. Draw the latent aspects $\theta_i \sim \text{Dir}(\alpha)$.
   b. Draw the neutral ratio $t_i \sim \text{Beta}(\gamma)$.
   c. For each aspect $k$, draw the predicted rating $\Omega_{ik} \sim \text{Beta}(\lambda R_i, \lambda(1 - R_i))$.

2. For each aspect $k$,
   a. Draw the word distribution $\phi_{ik}^0 \sim \text{Dir}(\beta)$ under neutral sentiment.
   b. Draw the word distribution $\phi_{ik}^1 \sim \text{Dir}(\beta)$ under positive sentiment.
   c. Draw the word distribution $\phi_{ik}^2 \sim \text{Dir}(\beta)$ under negative sentiment.

3. For each word $w_{ij}$ in review $d_i$,
   a. Draw a topic $z_{ij} \sim \text{Multinomial}(\theta_i)$.
   b. Draw a sentiment index $s_{ij} \sim \text{Multinomial}(t_i, (1 - t_i)\Omega_{iz_{ij}}, (1 - t_i)(1 - \Omega_{iz_{ij}}))$.
   c. Draw a word $w_{ij} \sim \text{Multinomial}(\phi_{z_{ij}}^{s_{ij}})$.
Methods: AIRS

- Characteristics of short reviews
  - The aspects mentioned in the review are quite imbalanced.
- Relationship between aspect probability and overall rating
  - If a reviewer talks more about a particular aspect, the rating on this aspect would be affected more by the overall rating.

\[ \Omega_{ik} \sim \text{Beta}(\lambda \theta_{ik} R_i, \lambda \theta_{ik} (1 - R_i)) \]

Maximum A Posterior is used for parameter estimation.
Outline

- Introduction
- Related Work
- Methods
- Experiments
- Conclusion
Experiments: Datasets

- TripAdvisor hotel review (observed aspect rating available)
  - Hotel (Users provide ratings on all predefined 7 aspects.)
  - Incomplete Hotel (Users provide ratings on several aspects.)

- RateBeer review

- Applause app review
  - Applause 1 (Reviews on similar apps.)
  - Applause 2 (Reviews on Temple Run and Temple Run II.)

### Table III: Statistics of Data Sets

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Review Number</th>
<th>Average Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotel</td>
<td>39,586</td>
<td>113.04</td>
</tr>
<tr>
<td>Incomplete Hotel</td>
<td>9,339</td>
<td>31.50</td>
</tr>
<tr>
<td>RateBeer</td>
<td>94,963</td>
<td>23.40</td>
</tr>
<tr>
<td>Applause 1</td>
<td>77,465</td>
<td>14.91</td>
</tr>
<tr>
<td>Applause 2</td>
<td>15,606</td>
<td>17.55</td>
</tr>
</tbody>
</table>

11/16/2015
Experiments : Performance Comparison

Table IV: Performances of Aspect Rating

<table>
<thead>
<tr>
<th></th>
<th>LARAM</th>
<th>AIR</th>
<th>AIRS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hotel Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>1.087</td>
<td>0.782</td>
<td>0.782</td>
</tr>
<tr>
<td>$\rho_{aspect}$</td>
<td>0.457</td>
<td>0.737</td>
<td>0.738</td>
</tr>
<tr>
<td>$Mis_{aspect}$</td>
<td>0.214</td>
<td>0.180</td>
<td>0.178</td>
</tr>
<tr>
<td>nDCG</td>
<td>0.956</td>
<td>0.956</td>
<td>0.956</td>
</tr>
<tr>
<td><strong>Incomplete Hotel Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>1.313</td>
<td>0.774</td>
<td>0.765</td>
</tr>
<tr>
<td>$\rho_{aspect}$</td>
<td>0.259</td>
<td>0.736</td>
<td>0.737</td>
</tr>
<tr>
<td>$Mis_{aspect}$</td>
<td>0.167</td>
<td>0.144</td>
<td>0.130</td>
</tr>
<tr>
<td>nDCG</td>
<td>0.966</td>
<td>0.967</td>
<td>0.969</td>
</tr>
</tbody>
</table>

Table V: KL Divergence Performance

<table>
<thead>
<tr>
<th></th>
<th>LDA</th>
<th>LARAM</th>
<th>AIR</th>
<th>AIRS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hotel Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 topics</td>
<td>9.634</td>
<td>10.683</td>
<td>8.735</td>
<td>8.800</td>
</tr>
<tr>
<td><strong>Incomplete Hotel Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 topics</td>
<td>12.337</td>
<td>24.433</td>
<td>8.912</td>
<td>11.678</td>
</tr>
</tbody>
</table>
Experiments: Aspect Sentiment Words

<table>
<thead>
<tr>
<th>Positive</th>
<th>Usability</th>
<th>Sound &amp; Graphic</th>
<th>Story</th>
<th>Platform</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fighting</td>
<td>Fighting</td>
<td>Fighting</td>
<td>Fighting</td>
<td>Fighting</td>
<td>Fighting</td>
</tr>
<tr>
<td>god</td>
<td>best</td>
<td>amazing graphic</td>
<td>great</td>
<td>amazing</td>
<td>awesome</td>
</tr>
<tr>
<td>king</td>
<td>amazing</td>
<td>excellent</td>
<td>awesome</td>
<td>best</td>
<td>love</td>
</tr>
<tr>
<td>awesome</td>
<td>amazing</td>
<td>adding</td>
<td>great</td>
<td>best</td>
<td>cool</td>
</tr>
<tr>
<td>best</td>
<td>graphic</td>
<td>beautiful</td>
<td>graphic</td>
<td>incredible</td>
<td>enjoy</td>
</tr>
<tr>
<td>beat</td>
<td>buy</td>
<td>perfect</td>
<td>play</td>
<td>fantastic</td>
<td>feel</td>
</tr>
<tr>
<td>epic</td>
<td>awesome</td>
<td>control</td>
<td>love</td>
<td>best</td>
<td>hand</td>
</tr>
<tr>
<td>defeat</td>
<td>best</td>
<td>graphic</td>
<td>wish more</td>
<td>cannot wait</td>
<td>action</td>
</tr>
<tr>
<td>avenge father</td>
<td>play</td>
<td>incredible</td>
<td>fun play</td>
<td>addicting</td>
<td>addicting</td>
</tr>
<tr>
<td>armor weapon</td>
<td>great</td>
<td>gorgeous</td>
<td>play</td>
<td>impressive</td>
<td>sword</td>
</tr>
<tr>
<td>beat bloodline</td>
<td>graphic</td>
<td>fantastic</td>
<td>love</td>
<td>epic</td>
<td>fighting</td>
</tr>
<tr>
<td>die</td>
<td>worth money</td>
<td>beautiful</td>
<td>wish more</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative</th>
<th>Sound &amp; Graphic</th>
<th>Story</th>
<th>Platform</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>fix</td>
<td>sound effect</td>
<td>not</td>
<td>disappointed</td>
<td>dodge button</td>
</tr>
<tr>
<td>save data</td>
<td>no sound</td>
<td>boring</td>
<td>gameloft</td>
<td>fix</td>
</tr>
<tr>
<td>bug</td>
<td>lack</td>
<td>fight</td>
<td>not</td>
<td>frustrating</td>
</tr>
<tr>
<td>fix bug</td>
<td>no</td>
<td>same</td>
<td>not</td>
<td>control</td>
</tr>
<tr>
<td>lost progress</td>
<td>sound</td>
<td>not worth</td>
<td>repeating</td>
<td>problem</td>
</tr>
<tr>
<td>lost</td>
<td>music</td>
<td>repetitive</td>
<td>same</td>
<td>annoying</td>
</tr>
<tr>
<td>play</td>
<td>sound</td>
<td>not</td>
<td>disappointing</td>
<td>bad</td>
</tr>
<tr>
<td>hour</td>
<td>frame rate</td>
<td>terribl</td>
<td>new</td>
<td>issue</td>
</tr>
<tr>
<td>save</td>
<td>decent</td>
<td>horrible</td>
<td>no</td>
<td>unresponsive</td>
</tr>
<tr>
<td>frustrating</td>
<td>missing</td>
<td>sad</td>
<td>no</td>
<td>dodge</td>
</tr>
<tr>
<td>delete save</td>
<td>unfortunately</td>
<td>arena</td>
<td>don’t</td>
<td>attack</td>
</tr>
</tbody>
</table>

11/16/2015

AIR model on Applause 1 dataset
**Experiments: User Behaviors**

### Table VII: Rating Behavior Comparisons with Different Apps.

<table>
<thead>
<tr>
<th>User Name</th>
<th>App</th>
<th>Overall Rating</th>
<th>Sound&amp;Graphic</th>
<th>Story</th>
<th>Usability</th>
<th>Review Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robert Rinehart</td>
<td>Infinity Balde</td>
<td>5</td>
<td>4.978</td>
<td>2.277</td>
<td>5.093</td>
<td>Graphics are very good. Fun to play but it can get tedious. Same monsters same places. But overall as an app and not a Xbox game it’s super.</td>
</tr>
<tr>
<td></td>
<td>Infinity Balde II</td>
<td>5</td>
<td>5.01</td>
<td>5.054</td>
<td>5.336</td>
<td>Such a great game. Amazing graphics. Challenging but not impossible. It never gets old. Keep em coming.</td>
</tr>
<tr>
<td>Wingbirdx</td>
<td>Infinity Balde</td>
<td>5</td>
<td>5.004</td>
<td>5.178</td>
<td>5.182</td>
<td>Great graphics and fun game play can’t wait for more.</td>
</tr>
<tr>
<td></td>
<td>Infinity Balde II</td>
<td>2</td>
<td>0.763</td>
<td>1.992</td>
<td>1.836</td>
<td>The lack of sounds when swords clash and random bits of dialogue are very upsetting in an otherwise very polished game.</td>
</tr>
<tr>
<td>Someguy127</td>
<td>Infinity Balde</td>
<td>5</td>
<td>5.058</td>
<td>5.051</td>
<td>5.085</td>
<td>The game is awesome! Infinity blade 2 better be as good! Simply the most impressive and breathtaking game in the app store!</td>
</tr>
<tr>
<td></td>
<td>Infinity Balde II</td>
<td>4</td>
<td>3.909</td>
<td>4.043</td>
<td>2.046</td>
<td>Crashes at title screen still after entering a new rebirth. Still has bugs.</td>
</tr>
</tbody>
</table>

**AIR model on Applause 2 dataset**
Conclusion

• Propose AIR model to extract aspects and predict their ratings for the general reviews

• Propose AIRS model to address the imbalance issue of aspects on short review

• Evaluate the proposed models in real-world data with various metrics
Thank you! Any Questions?