Complementary Venue Recommendation Model for Yelp
Hyun Sik Kim (hsik@stanford.edu), Ryan Wong (rawong@stanford.edu)

Overview
We developed a novel machine learning model that recommends nearby complementary venues (e.g., café) that the user is likely to enjoy based on a restaurant searched using Yelp.

- To simplify the search process for selecting multiple venues for a single outing using Yelp
- Our model recommends complementary venues using binary classifiers that predict the venue score using review text:
  - Great venue (1): 4-5 star rating
  - Mediocre/poor venue (0): 1-3 star rating
- Optimal recommendation/s based on:
  - Highest predicted score
  - Proximity (within 1 mile of restaurant)
- Model trained based on review text written by a common reviewer of the restaurant and complementary venue

Data and features
Feature vector comprised the count of unigrams and bi-grams in Yelp reviews of complementary venues from the Yelp academic dataset.

Feature selection using MI
- Document frequency
- MI slightly better for larger feature lengths and hence was adopted

Feature vector length:
- ~27k tokens (American)
- ~16k tokens (Japanese)

Naive Bayes comparable to SVM – linear was optimal SVM kernel

American Restaurants

<table>
<thead>
<tr>
<th>Model</th>
<th>Average Error Rate</th>
<th>Training Error</th>
<th>Test Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes (TF-IDF)</td>
<td>2.2%</td>
<td>1.4%</td>
<td>0.8%</td>
</tr>
<tr>
<td>SVM with linear kernel</td>
<td>2.0%</td>
<td>0.8%</td>
<td>0.2%</td>
</tr>
<tr>
<td>SVM with RBF kernel</td>
<td>3.0%</td>
<td>2.0%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Japanese Restaurants

<table>
<thead>
<tr>
<th>Model</th>
<th>Average Error Rate</th>
<th>Training Error</th>
<th>Test Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes (TF-IDF)</td>
<td>1.3%</td>
<td>1.2%</td>
<td>1.1%</td>
</tr>
<tr>
<td>SVM with linear kernel</td>
<td>2.0%</td>
<td>1.0%</td>
<td>0.5%</td>
</tr>
<tr>
<td>SVM with RBF kernel</td>
<td>3.0%</td>
<td>2.0%</td>
<td>0.7%</td>
</tr>
</tbody>
</table>

Naive Bayes was a faster learner

- Japanese Restaurants
  - NBM took about half the number of training examples

TF-IDF and normalization key to Naive Bayes accuracy

- Japanese Restaurants (Naive Bayes)
  - 4% reduction in test error

Bigrams improved performance

- Japanese Restaurants (Naive Bayes)
  - 2% reduction in test error

Document frequency was a simple and effective filter but MI was better

- MI slightly better for larger feature-lengths and hence was adopted

Future work
- Second classifier that distinguishes between 4 and 5 stars to provide finer granularity of venue scores
- Vector representations of words (word2vec)
- Exploration of n-grams to explore predictive power of phrases and idioms