Lyrics-based Music Tagging and Recommendations

Heidi Chen, Wenxuan Dai, Yijun Zhou, Xindi Zhao
Partners: Joe Cauteruccio, Aparna Kumar  Instructor: Pavlos Protopapas  TF: Isaac Slavitt

Overview

Spotify provides a great variety of songs to users all over the world on multiple platforms. Its key mission is to help people find the right music at every moment. In this project, we analyze the language component of songs to associate song lyrics with moods and themes. We seek to enhance the users' personalized music experience further through this project.

We incorporate both unsupervised methods, including Latent Dirichlet Allocation (LDA) and Word2Vec, and supervised models, including Long Short-Term Memory (LSTM), for language processing, sentiment analysis and predictive modeling. The final model allows us to develop an automatic tagging system which returns the suggested tags for a specific song, and a song recommender that generates a list of associated songs with user inputs.

Approach

Support Vector Machines
Random Forest
Gaussian Naive Bayes
Long Short-Term Memory

Bag of Words
Raw Lyrics

Feature Engineering

Word2Vec
Latent Dirichlet Allocation

Data

Song Lyrics
• Bag of words from MusiXmatch (MMX) for 5000 top frequency words.
• Raw lyrics from azlyrics.com for over 80,000 songs available, with approximately 14,000 matched with selected tags from Last.fm.

Tags
• Over 30,000 distinct tags generated by users from Last.fm API.
• 17 selected tags: sad, dark, love, energetic, happy, rain, emotional, funny, religious, political, rain, party, memory, Halloween, Christmas, grunge, and freedom.

Analysis and Results

1. Topic Exploration
• Word clustering from Latent Dirichlet Analysis (LDA) suggests some natural topic separations for themes like party, social issues, or emotions in general.
• Use the topics as a guide to select 17 target tags for our predictive models.

2. Language Processing and Feature Extraction
• Remove stop words and use Hashing Vectorizer to reduce dimension to 150.
• Use pre-trained Word2Vec embedding based on Google News dataset (~100 billion words) to transform to 300-dimension word vectors.

3. Predictive Modeling
• Baseline model:
  Feature Selection + Hashing Vectorizer + Random Forest
• Advanced model:
  Word2Vec + Long Short-Term Memory (LSTM)

4. Overall Prediction Accuracy
• Prediction Methodology:
  Predict one tag only or top 3 tags with highest probabilities.
• Loss Function:
  • Single Tag: 1 if prediction matches true tag, 0 otherwise
  • Top 3: 1 if any of the top 3 matches true tag, 0 otherwise

<table>
<thead>
<tr>
<th></th>
<th>Random Forest</th>
<th>Long Short-Term Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Train</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single Tag</td>
<td>38.45%</td>
<td>61.12%</td>
</tr>
<tr>
<td>Top 3</td>
<td>60.53%</td>
<td>82.95%</td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single Tag</td>
<td>19.13%</td>
<td>38.55%</td>
</tr>
<tr>
<td>Top 3</td>
<td>40.15%</td>
<td>60.35%</td>
</tr>
</tbody>
</table>

5. Prediction results (Advanced Model)

Song Recommendations:
With user input keywords, we look for the most related tag and return a list of song recommendations that best fit the tag.

Conclusions

• Song tags are imbalanced where there are lots of love and emotional themes, adding difficulties in making accurate prediction for minor classes.
• Limited vocabulary and non-mutually exclusive tags are the major problems we faced when trying to make precise predictions.
• Word processing techniques like Word2Vec and specialized language models like LSTM are effective in improving classification accuracy.
• Lyrics are useful in labeling songs with the most relevant themes.

Citation and Links

Colah, Christopher. "Understanding LSTM Networks." http://colah.github.io/posts/2015-08-Understanding-LSTMs/