A Regularized Recommendation Algorithm with Probabilistic Sentiment-Ratings

Filipa Peleja, Pedro Dias and João Magalhães
Department of Computer Science
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
OBJECTIVE: HOW TO IMPROVE RECOMMENDATIONS WITH USER COMMENTS AND REVIEWS?

Recommendation algorithm

Users

Products

Users

Products

Rating: ★★★★★
Review: Love it or hate it!
...

Rating: ★★
Review: This is a miserable film.
PROPOSED SOLUTION

Unrated Reviews

Sentiment analysis
1. Corpus representation
2. Orientation and intensity of words
3. Review classification

SentiWordNet
Semantic Orientation

Ratings

Rated Reviews

Sentiment-based recommendation
1. User/movie biases
2. Matrix factorization
3. Recommendations inference

User-movie recommendations
Sentiment Analysis Challenges

• Opinions are written in natural language which implies:
  - subjectivity;
  - sarcasm;
  - irony;
  - idiomatic expressions;
  - misspelling; etc.

• The same opinion word may be used in a positive or negative context

• Negative, Conditional and Comparative expressions
OPINION WORD ORIENTATION AND INTENSITY

• Semantic Orientation$^1$: 

$$SO(word) = \log_2 \left( \frac{\text{hits}(word, "excellent") \times \text{hits}("poor")}{\text{hits}(word, "poor") \times \text{hits}("excellent")} \right)$$

• How positive or negative is an opinion word?

  – SentiWordNet$^2$

(1) TURNEY, P. 2002, THUMBS UP OR THUMBS DOWN? SEMANTIC ORIENTATION APPLIED TO UNSUPERVISED CLASSIFICATION OF REVIEWS

(2) ESULI, A. AND SEBASTIANI, F., 2006, SENTIWORDNET: A PUBLICLY AVAILABLE LEXICAL RESOURCE FOR OPINION MINING
“Love it or hate it.”

“However, can someone tell me what on earth the last page...”

<table>
<thead>
<tr>
<th>word</th>
<th>family</th>
<th>SO (Google)</th>
<th>+ SentiWordNet</th>
<th>- SentiWordNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>love</td>
<td>n</td>
<td>-0.0824</td>
<td>1.375</td>
<td>0.0</td>
</tr>
<tr>
<td>it</td>
<td>nointerest</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>or</td>
<td>nointerest</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>hate</td>
<td>v</td>
<td>-0.8399</td>
<td>0.0</td>
<td>0.75</td>
</tr>
<tr>
<td>it</td>
<td>nointerest</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>however</td>
<td>r</td>
<td>-0.34153</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>someone</td>
<td>N</td>
<td>-0.65935</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>tell</td>
<td>V</td>
<td>-0.3956</td>
<td>0.875</td>
<td>0.625</td>
</tr>
<tr>
<td>me</td>
<td>nointerest</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>what</td>
<td>nointerest</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>on</td>
<td>nointerest</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>earth</td>
<td>n</td>
<td>-0.4041</td>
<td>0.0</td>
<td>0.625</td>
</tr>
</tbody>
</table>
**Multiple Bernoulli Classification**

\[
p(r_{a_i} = r \mid r_{e_i}) = \frac{f^r(r_{a_i}, r_{e_i})}{\sum_{L=1}^{10} f^L(r_{a_i}, r_{e_i})}
\]

- **The prediction is normalized accordingly to the predictions of all ratings.**
- **A classifier is learned for every rating value.** Thus, for each rating value there is a prediction for each review.
- **Rating range** IMDb dataset is 1 to 10
**Dataset Reviews from IMDB**

A total of 1,729,293 reviews were collected

<table>
<thead>
<tr>
<th>Split</th>
<th>#Reviews</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>335,975</td>
<td>Only to train SA</td>
</tr>
<tr>
<td>B</td>
<td>335,975</td>
<td>Test SA/Train RS</td>
</tr>
<tr>
<td>C</td>
<td>417,147</td>
<td>Train RS (no explicit ratings)</td>
</tr>
<tr>
<td>D</td>
<td>335,976</td>
<td>Train RS</td>
</tr>
<tr>
<td>E</td>
<td>201,586</td>
<td>Test RS</td>
</tr>
<tr>
<td>F</td>
<td>102,634</td>
<td>Validate RS</td>
</tr>
</tbody>
</table>
**Performance - Sentiment Analysis**

F-score on the IMDb corpus
Performance - Sentiment Analysis
Inferred Ratings in Recommendation Algorithm

Original ratings

Review-enhanced ratings

Computed by the sentiment analysis algorithm

New recommendations

Review-based recommendation
**Rating Matrix**

\[ R_{ra} = \begin{bmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nm} \end{bmatrix} \]

HIGHLY INCOMPLETE SINCE MOST ELEMENTS ARE EMPTY

**Predict an Unknown Rating:**

\[ \hat{r}_{ui} = p_u \cdot q_i \]

USERS AND PRODUCTS REPRESENTED IN THE SAME LATENT FACTOR SPACE

**With a SVD Decomposition the Rating Matrix**

\[ R_{ra} = \begin{bmatrix} u_{11} & \cdots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{n1} & \cdots & u_{nm} \end{bmatrix} \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nm} \end{bmatrix}^T = P \cdot Q^T \]

MATRIX FACTORIZATION ENABLES THE ASSESSMENT OF USERS PREFERENCES REGARDING THE PRODUCTS BY CALCULATING THEIR FACTOR REPRESENTATIONS
Ratings matrix $R_{ra}$

Factorization with Biases

Goal: Minimize the prediction error

$$[P, Q] = \arg \min_{p_u, q_i} \sum_{r_{ui} \in R_{ra}} (r_{ui} - \hat{r}_{ui})^2 + \lambda(||p_u||^2 + ||q_i||^2 + b_u^2 + b_i^2)$$
Ratings matrix $R_{ra}$

Factorization with Biases

$$[P, Q] = \underset{p_u, q_i}{\arg\min} \sum_{r_{ui} \in R_{ra}} (r_{ui} - \hat{r}_{ui})^2 + \sum_{r_{ui} \in \hat{R}_{rev}} \lambda(||p_u||^2 + ||q_i||^2 + b_u^2 + b_i^2)$$

 Ratings inferred from the Sentiment Analysis Framework are given to the RS. The Reviews Actual rating are known.
RATINGS MATRIX $R_{ra}$

FACTORORIZATION WITH SENTIMENT-BASED REGULARIZATION

$$R = R_{ra}, R_{rev}$$

$$\hat{R}_{ra} = \text{arg min} \sum_{\hat{r}_{ui} \in R_{ra}} (r_{ui} - \hat{r}_{ui})^2 + \sum_{c_{ui} \in \hat{R}_{rev}} \theta_{ui} \cdot (\hat{c}_{ui} - \hat{r}_{ui})^2$$

$$+ \lambda (\| p_u \|^2 + \| q_i \|^2 + b_u^2 + b_i^2)$$

ENRICH THE MATRIX $R$ WITH RATINGS INFERRED FROM REVIEWS WITH KNOWN AND UNKNOWN EXPLICIT RATINGS $R_{rev}$

The confidence level is given by de Sentiment Analysis framework
Recommendations: IMDb Dataset

Lower RMSE error

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Baseline (671,951 reviews)</td>
<td>2.099</td>
</tr>
<tr>
<td>DB+Ci Sentiment-ratings</td>
<td>2.122</td>
</tr>
<tr>
<td>DB+Ci Probabilistic sentiment-ratings</td>
<td>2.094</td>
</tr>
</tbody>
</table>

- Only Explicit Ratings
- Inferred Ratings
- Probabilistic Inferred Ratings
Recommendations: IMDb dataset

- Baseline (335,976 reviews)
- D+Ci
- D+Bi
- D+BCi

RMSE

Sentiment-ratings
Probabilistic sentiment-ratings

Reviews with Unknown ratings
• **Achievements:**
  – Extraction and sentiment analysis of users reviews
  – Introduced sentiment-based ratings in a recommendation algorithm

• **Next step:**
  – alternatives to SentiWordNet
  – semantic orientation metric
  – improve algorithm with opinion targets information
Thank you for your attention

Questions?