Movie Recommendation

Team Valak
Manpreet Kaur
Hao Su
Overview

- Introduction
- Collaborative Filtering
- Sequential Algorithm
- Map Reduce
- Parallel Algorithm
- References
Introduction

**Recommendation System:** A technique which predicts what the user may like based on his history of interaction with the system. They are applied in fields like:

- Online Shopping
- Browsing Portals
- Streaming Portals

Techniques used to implement a recommendation system are:

- Content Based
  - Only contents in the same categories will be recommended
- Collaborative Filtering
  - Based on past behavior
Collaborative Filtering

Based on the similarity in preferences, tastes and choices of users to analyses how similar users are and makes recommendation.

- Memory Based
  - User - Item: users who are similar to you also liked ...
  - Item - Item: users who liked this also liked ...
- Model Based

<table>
<thead>
<tr>
<th>Item</th>
<th>Item1</th>
<th>Item2</th>
<th>……</th>
<th>Itemn</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User1</td>
<td>R11</td>
<td>R12</td>
<td>……</td>
<td>R1n</td>
</tr>
<tr>
<td>User2</td>
<td>R21</td>
<td>R22</td>
<td>……</td>
<td>R2n</td>
</tr>
<tr>
<td>……</td>
<td>……</td>
<td>……</td>
<td>……</td>
<td>……</td>
</tr>
<tr>
<td>Usermn</td>
<td>Rmn1</td>
<td>Rmn2</td>
<td>……</td>
<td>Rmn</td>
</tr>
</tbody>
</table>
Example
Collaborative Filtering (cont.)

Similarity matrix

- Metrics that measure similarity between two users.

- Pearson correlation

\[ \text{sim}(i, j) = \frac{\sum_{c \in I_y} (R_{i,c} - \bar{A}_i)(R_{j,c} - \bar{A}_j)}{\sqrt{\sum_{c \in I_y} (R_{i,c} - \bar{A}_i)^2} \cdot \sqrt{\sum_{c \in I_y} (R_{j,c} - \bar{A}_j)^2}} \]

- Cosine vector similarity

\[ \text{sim}(i, j) = \frac{\sum_{k=1}^{n} R_{ik} R_{jk}}{\sqrt{\sum_{k=1}^{n} R_{ik}^2} \cdot \sqrt{\sum_{k=1}^{n} R_{jk}^2}} \]
Sequential Algorithm

- Using Similarity matrix to calculate how similar two users are for every pair of users.
- Select the closest K neighbors.
- Predict the rating of movies by calculating the weighted average of neighbor's rating.
- Recommend the top N movies by prediction.
Map Reduce

Top 3 movie recommendations for user K
Parallel Algorithm

- For every users, each core would run smaller datasets to calculate the similarity.
- The outcome of each core needs to be reduced.

For each user:

For every other user:

Mapper

Reducer
References

- A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item Clustering By SongJie Gong
- https://medium.com/@james_aka_yale/the-4-recommendation-engines-that-can-predict-your-movie-tastes-bbec857b8223
- https://towardsdatascience.com/various-implementations-of-collaborative-filtering-100385c6dfe0
Any Question?