Collaborative Filtering with Social Local Models

Huan Zhao

Joint work with Quanming Yao, James T. Kwok and Dik Lun Lee
Department of CSE, HKUST, Hong Kong
Recommender System

Inspired by your shopping trends

Benchmark Bouquets White Elegance, With Vase
$37.42

Benchmark Bouquets Signature Roses and Alstroemeria, With Vase
$39.44

KaBloom Romantic Red Rose Bouquet: 12 Fresh Cut Red Roses...
$35.99

See more

Products Recommendation
Recommender System

So many people are learning machine learning. What should I do to stand out?

Abhishek Patnia, Applied Scientist at Amazon.com

Updated Sat · Upvoted by Jordan Frank, Datamaker at Facebook and Siraj Memon, MS Computer Science, University of Maryland, Baltimore Coun...

Yes, it is true that many people are trying to learn Machine Learning. However, most people abandon their efforts really fast because: * Writing c Read More

Questions and Answers Recommendation
Recommender System

Restaurants Recommendation
Recommender System

- Recommender systems (RS) are everywhere.
- They are not only useful for people, but also create huge revenues for companies.

- The most popular RS method is collaborative filtering (CF).
  - User-based CF
  - Item-based CF
  - Matrix Factorization (MF)
Matrix Factorization (MF)

• The assumption is that users’ preferences are controlled by a small number of factors.

$$\min_{U,B} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - u_i b_j)^2 + \frac{\lambda_1}{2} ||U||^2_F + \frac{\lambda_2}{2} ||B||^2_F$$

https://buildingrecommenders.wordpress.com/2015/11/18/overview-of-recommender-algorithms-part-2/
Problems of MF

• Sparsity of the rating matrix.
  – More than 99% entries are missing.

• Cold Start.
  – Some users or items have no ratings.

• More importantly, the low-rank assumption may be too restrictive.
  – Lee et.al. relaxed this assumption to propose the local low rank matrix approximation (LLORMA) framework. [Lee et.al. ICML 2013]
LLORMA

• The rating matrix is composed of a number of smaller matrices which are of low rank.
• The rating is then predicted by weighted ensemble of predictions from all submatrices.

$$\mathbf{R}_{ij} = \sum_{t=1}^{q} \frac{\mathbf{w}_{ij}^{t}}{\sum_{s=1}^{q} \mathbf{w}_{ij}^{s}} \left[ \mathbf{u}_{i}^{t}(\mathbf{v}_{j}^{t})^{\top} \right].$$
Problems of LLORMA

• Meaningless of submatrices
  – Randomly select *anchor points* from the rating matrix.
  – Select other points within some distance threshold.

• Inconsistence of submatrices.

\[
d(i_t, k) = \arccos \left( \frac{u_{i_t} u_k}{\|u_{i_t}\| \cdot \|u_k\|} \right)
\]

*Lemma III.1.* Given any matrices \( U \) and \( V \) which are an optimal solution to (2), then \( \hat{U} = UQ \) and \( \hat{V} = VQ^{-1} \) are also an optimal solution, if matrix \( Q \) is invertible.

• Space and computational cost.
  – Need to save pairwise similarities of all points in all submatrices.
Our Work (Social Local Models)

• We are the first to integrate social connections with LLORMA.
  – It can enjoy the advantages of both social recommendation and local low rank assumption.

• Meaningfulness of submatrices.
  – Social communities can explain the local models.

• Social regularization with LLORMA can further improve the recommending performance.
Our Work (Social Local Models)

- Right part is LLORMA.
- Left part motivates our framework.
Framework

• Identify social groups from the social graph and then construct submatrices based on the groups.
• Apply MF to all the submatrices, independently, and obtain multiple groups of user-specific and item-specific latent features.
• Predict the missing ratings using the ensemble of predictions from all submatrices.

\[ R_{ij} = \frac{1}{q} \sum_{t=1}^{q} u_i^t (v_j^t)^\top \]
Submatrices Construction

• Heuristic methods.
  – Build the social groups based on the fact that users’ influences to each other can propagate through the networks.
  – Submatrices are constructed by firstly select influential users, called connectors, and their friends within a fixed number of hops.

• Different methods to select connectors.
  – Hub: *those with the most friends*.
  – Random: *randomly selection*.
  – Random-Hub: *randomly selection from a set of Hub users*.
  – Greedy: each time we select a connector, we select from those not yet covered by existing connectors.
Submatrices Construction

• Systematic methods.
  – Overlapping community detections methods can be exploited.
  – BIGCLAM are utilized for its scalability and good performance. [Yang et.al. ICDM 2013]

• After we create social groups, for each group, we select the users and the items they rate to construct a submatrix.
Submatrices Factorization

- Social LOcal Matrix Approximation (SLOMA).
  - Apply MF to each submatrix independently.

- SLOMA++.
  - Add social regularization to each submatrix factorization.

\[
\min_{U,V} \frac{1}{2} \sum_{(i,j) \in \Omega} (O_{ij} - u_i v_j^T)^2 + \frac{\lambda}{2} (\|U\|_F^2 + \|V\|_F^2) \\
+ \frac{\beta}{2} \sum_{i=1}^{m} \sum_{j \in \mathcal{F}(i)} S_{ij} \|u_i - u_j\|_2^2.
\]
Experimental Results

• Datasets.

![Table of Datasets](image)

TABLE I

|       | Users  | Items  | Ratings   | R_density   | S_edges  | S_density   |
|-------|--------|--------|-----------|-------------|----------|-------------|
| Yelp  | 76,220 | 79,257 | 1,352,762 | 0.022%      | 647,451  | 0.022%      |
| Douban| 103,054| 57,908 | 15,129,113| 0.254%      | 753,358  | 0.028%      |

• Evaluation metrics.

  – The smaller, the better.

\[
\text{MAE} = \frac{1}{|\Omega|} \sum_{(i,j) \in \Omega} |O_{ij} - R_{ij}|,
\]

\[
\text{RMSE} = \sqrt{\frac{1}{|\Omega|} \sum_{(i,j) \in \Omega} (O_{ij} - R_{ij})^2},
\]
## Experimental Results

| Datasets | K  | Metrics | RegSVD | LLORMA | SocReg | SLOMA  | SLOMA++ |
|----------|----|---------|--------|--------|--------|--------|---------|
|          |    | MAE     | 0.9478 | 0.9459 | 0.9228 | 0.9362 | 0.9301  |
|          |    | Improve | +1.87% | +1.67% | -0.79% | +0.65% | +0.04%  |
|          |    | RMSE    | 1.1908 | 1.1843 | 1.1802 | 1.1760 | 1.1755  |
|          |    | Improve | +1.28% | +0.74% | +0.40% | +0.04% | +0.04%  |
| Yelp     | 10 | MAE     | 0.9499 | 0.9477 | 0.9190 | 0.9389 | 0.9240  |
|          |    | Improve | +2.73% | +2.50% | -0.54% | +1.59% | +0.76%  |
|          |    | RMSE    | 1.1918 | 1.1862 | 1.1754 | 1.1788 | 1.1698  |
|          |    | Improve | +1.85% | +1.38% | +0.48% | +0.76% | +0.76%  |
|          | 20 | MAE     | 0.5828 | 0.5811 | 0.5662 | 0.5744 | 0.5603  |
|          |    | Improve | +3.86% | +3.58% | +1.04% | +2.45% | +0.07%  |
|          |    | RMSE    | 0.7347 | 0.7310 | 0.7165 | 0.7255 | 0.7105  |
|          |    | Improve | +3.29% | +2.80% | +0.84% | +2.07% | +2.07%  |
| Douban   | 10 | MAE     | 0.5803 | 0.5779 | 0.5638 | 0.5715 | 0.5573  |
|          |    | Improve | +3.96% | +3.56% | +1.15% | +2.48% | +2.48%  |
|          |    | RMSE    | 0.7320 | 0.7278 | 0.7142 | 0.7225 | 0.7080  |
|          |    | Improve | +3.28% | +2.72% | +0.87% | +2.01% | +2.01%  |

- Our proposed methods outperform baselines consistently, which demonstrates the effectiveness of our framework.
- Comparing SLOMA and LLORAM, the performance gain is obtained by incorporating the social connections.
- Comparing SLOMA++ and SocReg, the performance gain is obtained by the local low rank assumption.
Experimental Results

- Impact of number of local models.
  - When the number is smaller than 10, the performance is weak.
  - When the number is larger, e.g., 50, the performance is consistently good.
Experimental Results

• Impact of number of hops.
  – When the number is smaller than 3, the performance is weak.
  – When the number is larger than 3, the performance is consistently good.
Experimental Results

- Impact of Connector Selection methods.
  - Hub and Greedy are the best two methods.
  - Community-based one is the worst.
Conclusion

• We propose social local models to enhance LLORMA.

• Social recommendation and local low rank assumption can both benefit the recommending performance.

• Experimental results have demonstrates the effectiveness of our SLOMA and SLOMA++. 