Hybrid Product Recommender System

Team Members:
Ankush Sachdeva
Khagesh Patel
Motivation

- Widely used in many e-commerce companies like Amazon, Flipkart.
- Netflix challenge
Dataset Used

- Netflix
  - 100 Million ratings
  - 480 thousand customers
  - 17 thousand movies
- Movielens
  - 10 Million ratings
  - 71 thousand customers
  - 11 thousand movies
Analysis of MovieLens Data
General Approach

- User-User Collaborative filtering
  - K nearest neighbor using different similarity metric: Manhattan, Euclidean, Pearson correlation coefficient, Cosine similarity.

- Item-Item Collaborative filtering
  - Above approach.
  - Slope one.

- Graph based method
  - Spanning tree.
Singular Value Decomposition

- Regularized Singular Value Decomposition
  \[
  \hat{r}_{ui} = p_u^T q_i + b_u + b_i + \mu
  \]

- Asymmetric Singular Value decomposition
  - Train feature vector for only items
  - \[
  p_u^T = |R(u)|^{-0.5} + \sum_{j \in R(u)} (r_{uj} - b_{uj}) x_j + |N(u)|^{-0.5} \sum_{j \in N(u)} y_j
  \]

- Modified Singular Value Decomposition with feedback from implicit rating.

- Integrating above models for Singular Value Decomposition.
Slope-one algorithm (item-item collaborative filtering)

Uses simple regression model of form $f(x) = x + b$ for different items.

**Example:**

User A gave a 1 to Item I and an 1.5 to Item J.
User B gave a 2 to Item I.

How do you think User B rated Item J?
The Slope One answer is to say 2.5 ($1.5 - 1 + 2 = 2.5$).

- Take average of all similar users.
- It was shown to be much more accurate than linear regression in many cases.
- Linear regression has greater tendency for over fitting.
• Root mean square error observed for Movielens dataset by slope one algorithm is 1.03136.
Singular Value Decomposition

- Decompose rating matrix $M \times N$ to $M \times k$ and $k \times N$ such that root mean square error is minimum.

- Our approach:
  - Perform gradient descent until no further improvement can be achieved.
  - This approach does not require missing values so no need to fill arbitrary values in our matrix.
  - Exact SVD if all entries are filled otherwise can be taken as approximate SVD.

```c
#define MAX_RATINGS 100001
#define MAX_MOVIES 1683
#define MAX_CUSTOMERS 944
#define MIN_EPOCH 120
#define MAX_EPOCH 200
#define MAX_FEATURES 50
#define MIN_IMPROVEMENT 0.0001 // Minimum improvement required to continue current feature
#define INIT 0.1 // Initialization value for features
#define LRATE 0.001 // Learning rate parameter
#define K 0.015 // Regularization parameter used to minimize over-fitting
#define PseudoCount 25.0
```
- Root mean square error observed for Movielens dataset by SVD algorithm is 0.471307.
Temporal effects (TODO)

There are two main temporal effects in the data

1. Movie biases – Certain movies may become more or less popular/liked over time. We use the item bias to capture this effect.

2. User biases – Users tend to change their baseline rating over time, mainly because the users give ratings relative to the previous movies they had seen. We use the user bias to capture this effect.

Both the biases are time dependent function.

Item bias changes slowly over time compared to user bias.
Thank You

Questions?