

Movie Recommendation System Using SVD

1 Finding the Problem - Application

Based on User Ratings
Identifying Similar Movies

2 Collecting Dataset

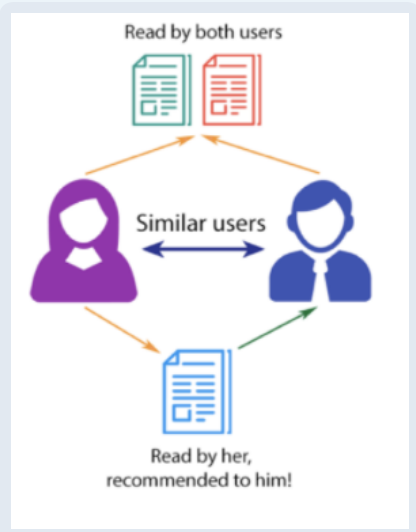
Input: 'ratings.dat' and 'movies.dat'
MovieLens IM Dataset

3 Recommendation System



Content-based Filtering

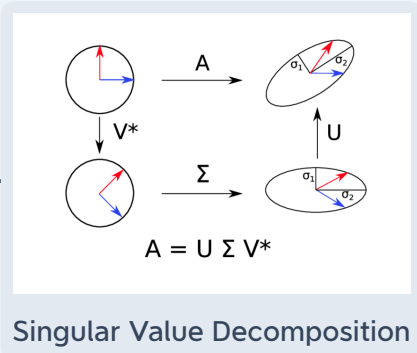
It is based on a description of the item



Collaborative Filtering

The people who have liked an item in the past will also like the same in future.
This approach builds a model based on the past behaviour of users.
The user behaviour may include previously watched videos, purchased items, given ratings on items

4 Algorithm



Singular Value Decomposition

$A = U D V^T$
Left singular vectors Singular values Right singular vectors
Definition

It decomposes a matrix of any shape into a product of 3 matrices

Lucid Analogy

57=1x3x19 - Prime Number in Middle

Elements

Matrix U - m*r

Singular matrix of (user*latent factors)

Matrix S - r*r

Diagonal matrix (shows the strength of each latent factor)

Matrix V - r*n

Singular matrix of (item*latent factors)

Latent factors - characteristics of the item

Steps

1

A - m x n utility matrix | It Represents relationships between users and items

$A = U S V^T$

Decomposing

Decomposing a matrix into three other matrices

2

Factorization

$\hat{r}_{ui} = x_i^T y_u$

Expected rating by a user on an item

Each item be represented by a vector x_i
Each user is represented by a vector y_u

Determining x_i & y_u

$$Min(x, y) = \sum_{(u, i) \in K} (r_{ui} - x_i^T y_u)^2$$

x_i & y_u can be determined by Square error difference between their dot product and the expected rating in the user-item matrix is minimum

Adding Regularization Term

$$Min(x, y) = \sum_{(u, i) \in K} (r_{ui} - x_i^T y_u)^2 + \lambda(\|x_i\|^2 + \|y_u\|^2)$$

Let the model generalise well and not overfit the training data

Using Bias term - To reduce error b/w predicted value and actual value

$$Min(x, y, b_i, b_u) = \sum_{(u, i) \in K} (r_{ui} - x_i^T y_u - \mu - b_i - b_u)^2 + \lambda(\|x_i\|^2 + \|y_u\|^2 + b_i^2 + b_u^2)$$

Final Equation of SVD

μ - average rating of all items
 b_i = average rating of item i - μ
 b_u = average rating given by user u - μ

5 Parsing Data from .Dat File

Ratings Data
Movies Data

6 Creating Matrix

Rows: movies
Columns: ratings

7 Data Normalization - Rescaling

8 Computing Singular Value Decomposition

Compute SVD