

Recommendation System

Objective

This project aims to train a Machine learning model to learn the user's preferences from the Movies rating dataset and then recommend a movie for any user based on its learning. As this project deals with the huge dataset, big data tools like Spark framework has been leveraged.

Dataset

The dataset used in this project is fetched from [MovieLens 20M Dataset](#)

The datasets describe ratings and free-text tagging activities from MovieLens, a movie recommendation service. It contains 20000263 ratings and 465564 tag applications across 27278 movies. These data were created by 138493 users between January 09, 1995 and March 31, 2015. This dataset was generated on October 17, 2016. Users were selected at random for inclusion. All selected users had rated at least 20 movies.

User to user Collaborative Filtering

$$s(i, j) = \frac{1}{2}$$

