Recommender system

movie recommendation

Team Member:
张陬，陈睦尧，夏鸣飞

background information

We have lots of movies and users, and we hope that the system can recommend movies to users according to their preferences

Data

What we have?

1. Movie information:
    1. Number of votes
    2. Average rating
    3. Other detail information about the content
    4. Movie Cast, Crew, Keywords and Genre

2. User information:
    1. Rating to certain movie at certain timestamp

Hybrid recommendation

Combining Collaborative Filtering and Content Filtering

Collaborative filtering has the following problems:

1. Cannot recommend something without history record
2. The timeliness was not taken into consideration
3. performance decrease when we have sparse data

We found an open-source code on github which use Content-Based and Collaborative Filtering with Deep learning Model

(https://github.com/khanhnamle1994/movielens)

Introduction and Disadvantage

The open-source code is base on Matrix Factorization, implemented using Keras

The core idea is to map users and items to a low-dimensional latent factor space and predict the score by calculating the dot of the user vector and the item vector

Disadvantage:

Linear limitations: Dot product operations only capture linear relationships and are difficult to model complex interactions, such as nonlinear feature combinations

Improvements

Other possible implements:

1. Collaborative Filtering Based on Spark (Matrix Factorization Optimization)
2. Introduce the time difference factor and the weights of popular/unpopular items (Weighted similarity calculation)
3. Introduce the timestamp weight in the similarity calculation