

Project name- ML based Item Recommender (for movies)

Club- Programming Club

Mentor- Varun Khare

Team:

1. Pushkar Dixit
2. Gagandeep Singh
3. Himanshu Arora
4. Mayank Lunayach
5. Niket Agarwal

Documentation of the work done upto now (as of 7 June):

The project work was split up into front-end and back-end tasks and so was the team, to take care of each.

Team Decisions

We decided to deploy the **item-recommender** as a web app running on a local server. So, it was decided to use **django** as the web-framework for our project. The algorithm to be used was discussed upon in great detail and finally we considered '**collaborative filtering**' to be the best for the purpose.

Discussions

Collaborative filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating) while content based recommendation method relies on similarity between products to recommend new products for users. The pros and cons of either were discussed and it was established that the former can give better and less obvious results than the latter most of the times. The method to establish similarity between two users and also how to quantify it was discussed upon.

Primary Tasks undertaken

The part of the team responsible for back-end tasks looked into the working of django and related stuff. The part of the team responsible for front-end tasks looked into some advanced bootstrap stuff to fetch ideas for the design of the web app. The team extracted and sorted out the relevant data from the complete 100-K dataset downloaded from movielens.

Issues that surfaced

- Sqlite the database engine that comes with django itself cannot handle more than a certain amount of data due to which it cannot fulfill the needs of our project.
- The django framework proved to be a bit difficult to understand at first,so much and so that the team doubted if it was really required for the project,anyways the team decided to stick to it.
- Extracting the relevant data from the complete dataset and putting it into a form that could be fed directly in our functions was difficult