Netflix Movie Recommendation System Design – Proposal

Project Source Link: <https://www.kaggle.com/netflix-inc/netflix-prize-data>

Dataset info:

1. Size: 2 GB (.txt file)
2. Raw Data:
   1. Combined\_data: columns (movie ID + user + rating + rating date)
   2. Movie\_titles.csv: columns (movie ID + YearOfRelease, Title)
   3. Probe.txt

Deployment plan:

1. API design plan: create an application which is used for recommending movies from system database based on user history or input movies name. (get idea from gnoosic <https://www.gnoosic.com/faves.php>)
2. We can load Netflix dataset into DB and build a simple system in movie recommendation
   1. User can input some movie title they prefer and system will output some movies options
   2. DB will also save the history for each user movie input
3. Webpage design:
   1. Frontend side
      1. Design a very simple page (user login + user input page + recommend result page)
      2. Framework: Angular + typescript
   2. Backend:
      1. Use python flask build a simple server for handling url request
   3. Database:
      1. PostgreSQL or MySQL
   4. Deployment plan:
      1. Deploy in AWS EC2 or ElasticBean
4. Task Schedule:

|  |  |  |
| --- | --- | --- |
| Task No. | Task Description | Task deadline |
| Data-preprocessing | Load + clean + analysis data, make some general understanding for this dataset | 10/5/2020 – 10/16/2020 |
| Frontend design | Design frontpage of this movie recommendation system | 10/5/2020 – 10/20/2020 |
| Feature Engineering + Model Selection | Generate training data from raw source, find some common collaborative filtering model for training and test performance | 10/20/2020 – 10/31/2020 |
| Performance analysis + model improvement |  |  |
| Flask backend design + Database loading | Process frontend request and extract data from backend | 11/2/2020 – 11/20/2020 |
|  |  |  |
|  |  |  |
|  |  |  |