

Project: Movie Recommendation Engine

Karsten Poddig

June 14, 2020

Abstract

The project described in this report is a Movie Recommendation Engine

Github Repo: https://github.com/KarstenPoddig/django_movie_project

Keywords: Data Science, Data Engineering, Statistical Dashboard, Analytics, Movie Recommendation Engine, Hierarchical Clustering, Python, Django, Docker, Kubernetes, Continuous Integration, Continuous Development, Google Cloud Platform, SQL, PostgreSQL, JavaScript, Git, GitHub



Contents

1	Introduction	3
2	App functionalities	3
2.1	Sign up / Sign In	3
2.1.1	Sign In	4
2.2	Search Movies	4
2.3	Rate Movies	5
2.4	View Rated Movies	5
2.4.1	Movies with Details	6
2.4.2	Clustered Movies	6
2.4.3	Statistics	6
2.5	Views Suggestions	7
3	Data Science Part	7
3.1	Similarity	7
3.2	Suggestions	7

4	Data Engineering Part	7
4.1	Source of Data	7
4.2	Google Cloud SQL	7

1 Introduction

2 App functionalities

This section provides a quick overview of the functionalities of this app. The following functionalities can be accessed without an account:

- Search Movies
- Search for similar movies for a specific movie

The main functionalities require an account. These functionalities include

- Rate Movies
- Get Suggestions based on your ratings

Since the main purpose of this project is to provide movie suggestions I recommend to sign in if you want to explore this app. You can do this by creating an account and create your own ratings. But rating a sufficient number of movies can be time consuming, so if you want to go on fast and simply want to explore what how this app works you can simply use the following account with existing ratings:

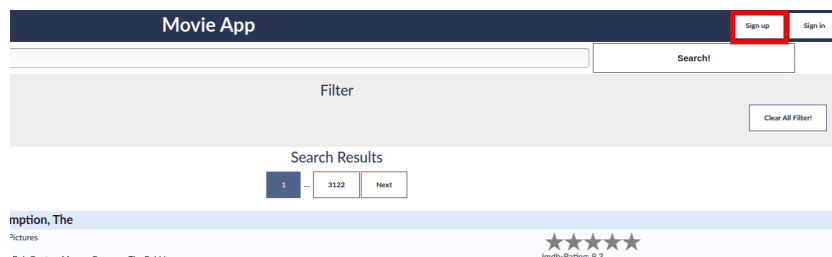
username	testuser
password	password

The following subsections guide you through the app. But since the app should be somehow intuitive you can also skip these subsections.

2.1 Sign up / Sign In

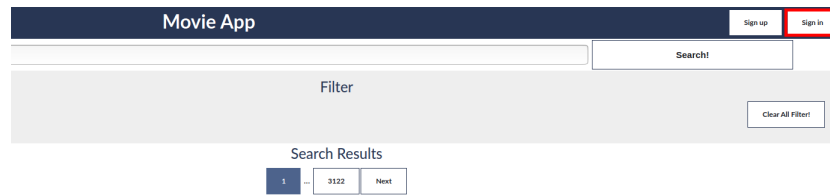
Sign Up

Click on the button in the right top corner and enter all necessary data.



2.1.1 Sign In

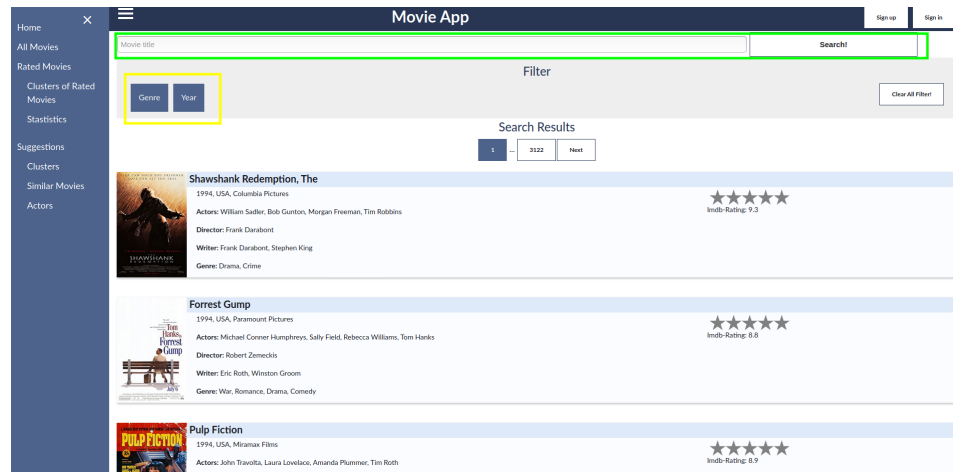
Click on the button in the right top corner and all necessary data.



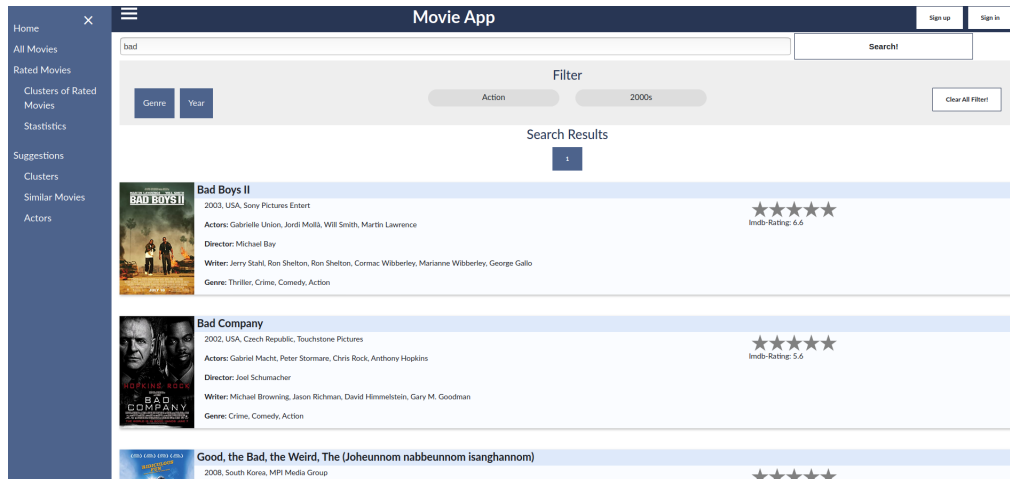
Also you are redirected to the Sign In - page if you are accessing areas which require a login (like Rated Movies, etc.)

2.2 Search Movies

Click on the link 'All Movies' in the left bar. You will see a list of all movies in the database with the regarding details. To search a movie you can enter parts of title (green area) and optionally set certain filter criteria (yellow area). The results are ordered decreasingly by popularity (total number of ratings).



For example a search for movies with the term 'Bad' in the title with the Genre Action in the 2000s will look like the following:



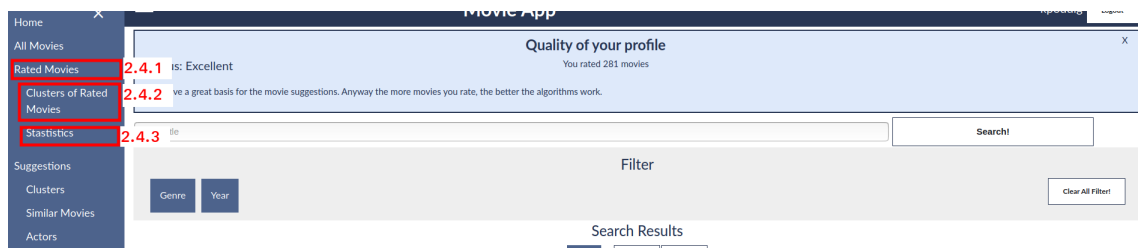
2.3 Rate Movies

Once you are logged in you can rate movies. Just enter the page 'All Movies', search the movie you want to rate and click on the regarding rating (red area in the following picture). If you choose a rating of 0 stars an existing rating will be deleted.



2.4 View Rated Movies

At some point you maybe want to take a look at the movies you already rated. The links for the sections 2.4.1-2.4.3 are displayed in the following picture.



2.4.1 Movies with Details

This is basically the same page as 'All Movies' (section 2.2) with the exception that just rated movies are listed here. You can search and navigate through your rated movies in the same way as in section 2.2.

2.4.2 Clustered Movies

To determine better suggestions your rated movies are clustered. Take a look at section. Basically each cluster (each line) contains movies which are similar to each other. Factors which imply a similarity could be the genre, the production year, the atmosphere, actors, etc. For a better understanding what similarity in this context means I again refer to Section If you want to update the clustering hit the button 'Refresh Cluster' (marked red).

Quality of your profile ×

Status: Excellent
You rated 281 movies
You have a great basis for the movie suggestions. Anyway the more movies you rate, the better the algorithms work.

Refresh Cluster

Your clustered movies

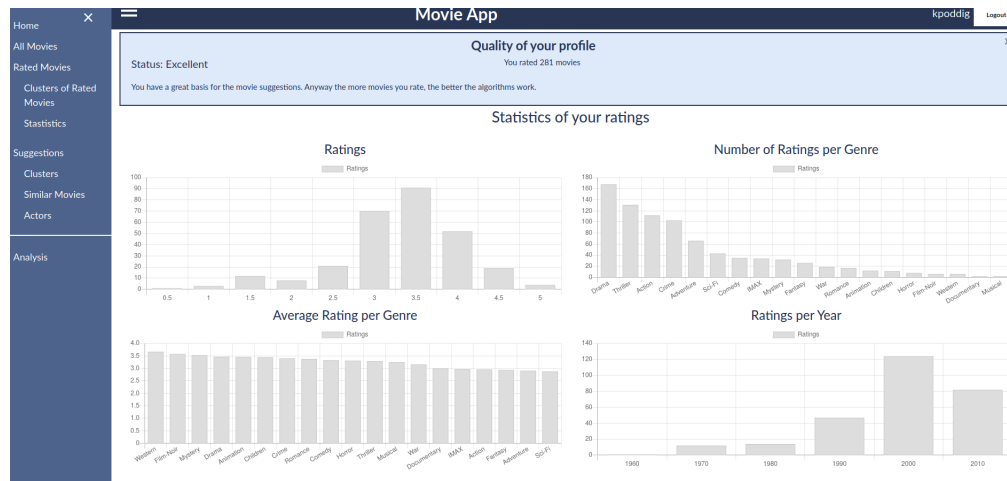
Cluster 347
original great ending storytelling mentor dialogue

Cluster 348
original action chase mentor pg-13

Cluster 349
original mentor great ending great acting good

2.4.3 Statistics

This page provides a statistical dashboard for your ratings.



2.5 View Suggestions

Now it comes to the core functionality of the app. The following picture displays the links to the sections 2.5.1 - 2.5.3.

2.5.1 Suggestions based on clusters

2.5.2 Search Similar Movies

2.5.3 Suggestions based on Actors

3 Data Science Part

3.1 Similarity

3.2 Suggestions

4 Data Engineering Part

4.1 Source of Data

MovieLen, Imdb-API Own ratings

4.2 Google Cloud SQL