



Capstone Vlog4

Team DoubleX
Xufeng Liu Yixu Zhou



Business need

A movie Recommendation System

People always waste too much time selecting movies, our app can recommend movies for our users according to their viewing history or search history. Users can also recommend their favorite movie to their friends.

Our wish is that our users are all over the world, so they can even recommend some movies from different countries.



status

yellow

specific activity

Collect dataset between few websites

Movie Recommender Systems

[Notebook](#) [Data](#) [Logs](#) [Comments \(72\)](#)

794 Copy & Edit 4298

270,000 users. Includes tag genome data with 12 million relevance scores across 1,100 tags.

- **The Small Dataset:** Comprises of 100,000 ratings and 1,300 tag applications applied to 9,000 movies by 700 users.

I will build a Simple Recommender using movies from the *Full Dataset* whereas all personalised recommender systems will make use of the small dataset (due to the computing power I possess being very limited). As a first step, I will build my simple recommender system.

```
In [1]: %matplotlib inline
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
from ast import literal_eval
from sklearn.feature_extraction.text import TfidfVectorizer, CountVectorizer
from sklearn.metrics.pairwise import linear_kernel, cosine_similarity
from nltk.stem.snowball import SnowballStemmer
from nltk.stem.wordnet import WordNetLemmatizer
from nltk.corpus import wordnet
from surprise import Reader, Dataset, SVD, evaluate

import warnings; warnings.simplefilter('ignore')
```

Table of Contents

1 Movies Recommender System

Movie Industry

[Data Card](#) [Code \(80\)](#) [Discussion \(5\)](#)

321 New Notebook Download (434 kB)

Detail	Compact	Column	10 of 15 columns		
▲ name	▲ rating	▲ genre	▲ year	▲ released	▲ score
name of the movie	rating of the movie (R, PG-13, etc.)	main genre of the film	release year	release date	IMDb sco
7512 unique values	R PG-13 Other (1859)	48% 28% 24%	Comedy Action Other (3718)	29% 22% 48%	3415 unique values
Any Which Way You Can	PG	Action	1988	December 17, 1988 (United States)	6.1
The Gods Must Be Crazy	PG	Adventure	1988	October 26, 1984 (United States)	7.3
Popeye	PG	Adventure	1988	December 12, 1988 (United States)	5.3
Ordinary People	R	Drama	1988	September 19, 1988 (United States)	7.7
Dressed to Kill	R	Crime	1988	July 25, 1988 (United States)	7.1
Somewhere in Time	PG	Drama	1988	October 3, 1988 (United States)	7.2
Fame	R	Drama	1988	May 16, 1988 (United States)	6.6
9 to 5	PG	Comedy	1988	December 19, 1988 (United States)	6.9
The Fog	R	Horror	1988	February 8, 1988 (United States)	6.8
Star Crazy	R	Comedy	1988	December 12, 1988 (United States)	6.8



Feedback & Comment

- Need a good dataset for AI system
- Vlog slide should be more specific
- Github should be orgnized
- Add some special functions to our app that can make it difference (like platform, more detailed info, share function etc.)



Next UP

share function

AI model training



Thanks