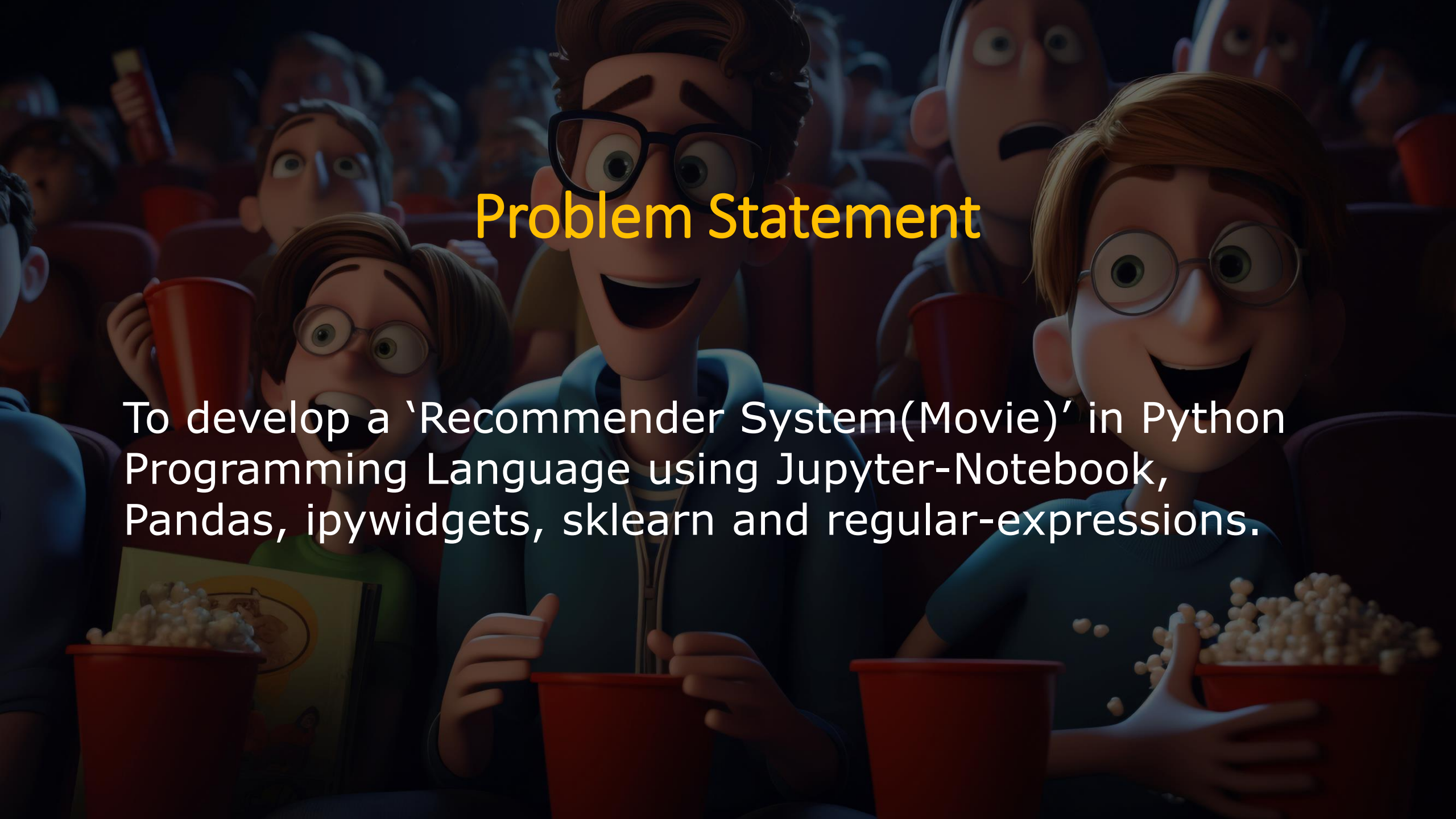
Recommendation System (Movie)

Name : Arsh Arora
University Roll No. : 2019536
Course : B-Tech CE (CS)
Mentor : Ms. Vishu Tyagi



Problem Statement

To develop a ‘Recommender System(Movie)’ in Python Programming Language using Jupyter-Notebook, Pandas, ipywidgets, sklearn and regular-expressions.

Introduction

The Movie Recommendation System (MRS) was designed and implemented to enhance user experience by providing personalized movie suggestions based on their preferences.

Steps involved in solving the problem:

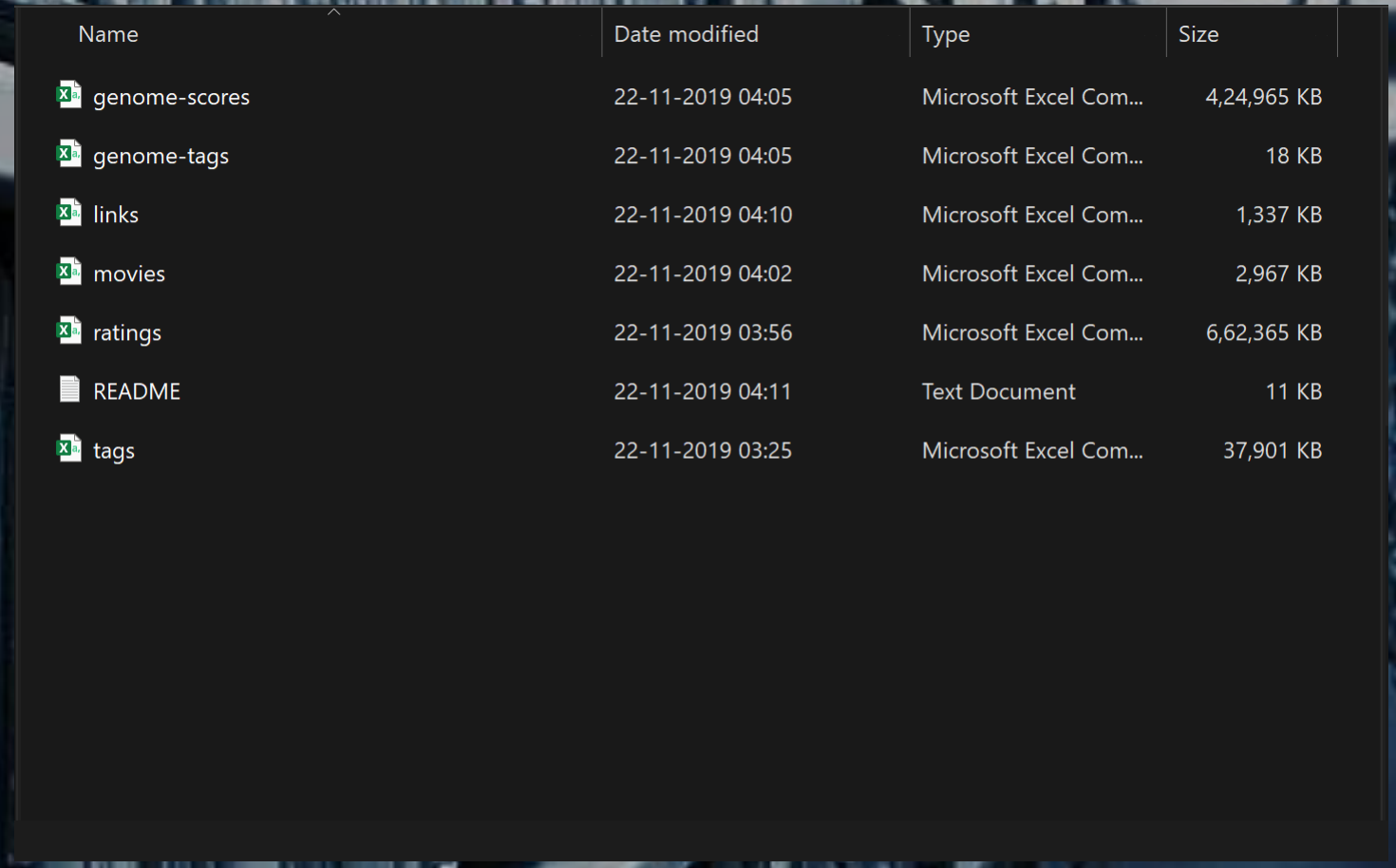
1. **Find** a metric to rate the movies.
2. **Calculate** the metric for all the movies in the list.
3. **Sort** in descending order to showcase top results.

Methodology

1. Data Collection:

Gather relevant data sources, including user preferences, movie attributes, and historical interactions.

Ensure data quality and handle any missing or noisy data.

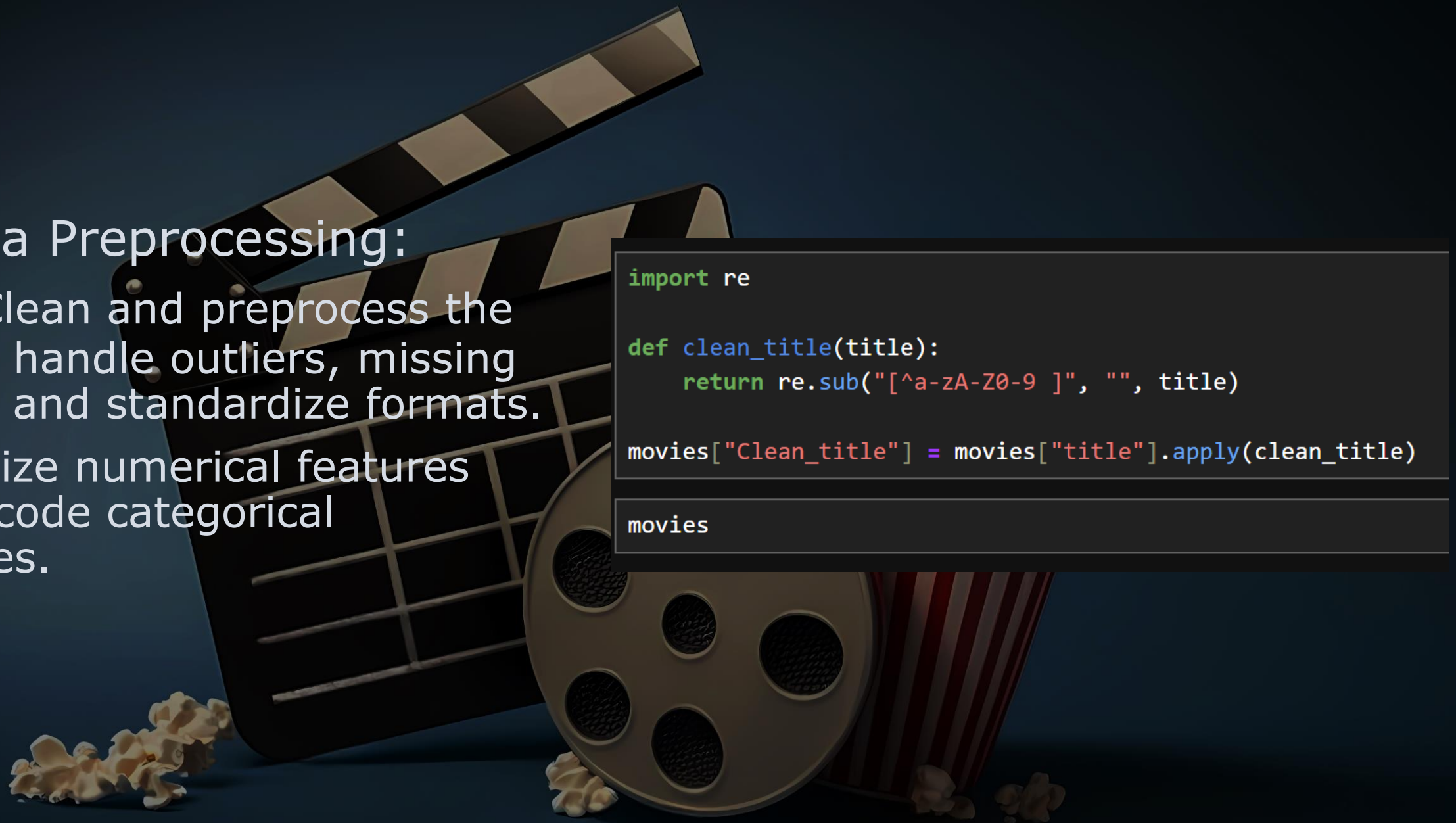


Name	Date modified	Type	Size
genome-scores	22-11-2019 04:05	Microsoft Excel Com...	4,24,965 KB
genome-tags	22-11-2019 04:05	Microsoft Excel Com...	18 KB
links	22-11-2019 04:10	Microsoft Excel Com...	1,337 KB
movies	22-11-2019 04:02	Microsoft Excel Com...	2,967 KB
ratings	22-11-2019 03:56	Microsoft Excel Com...	6,62,365 KB
README	22-11-2019 04:11	Text Document	11 KB
tags	22-11-2019 03:25	Microsoft Excel Com...	37,901 KB

2. Data Preprocessing:

Clean and preprocess the data to handle outliers, missing values, and standardize formats.

Normalize numerical features and encode categorical variables.



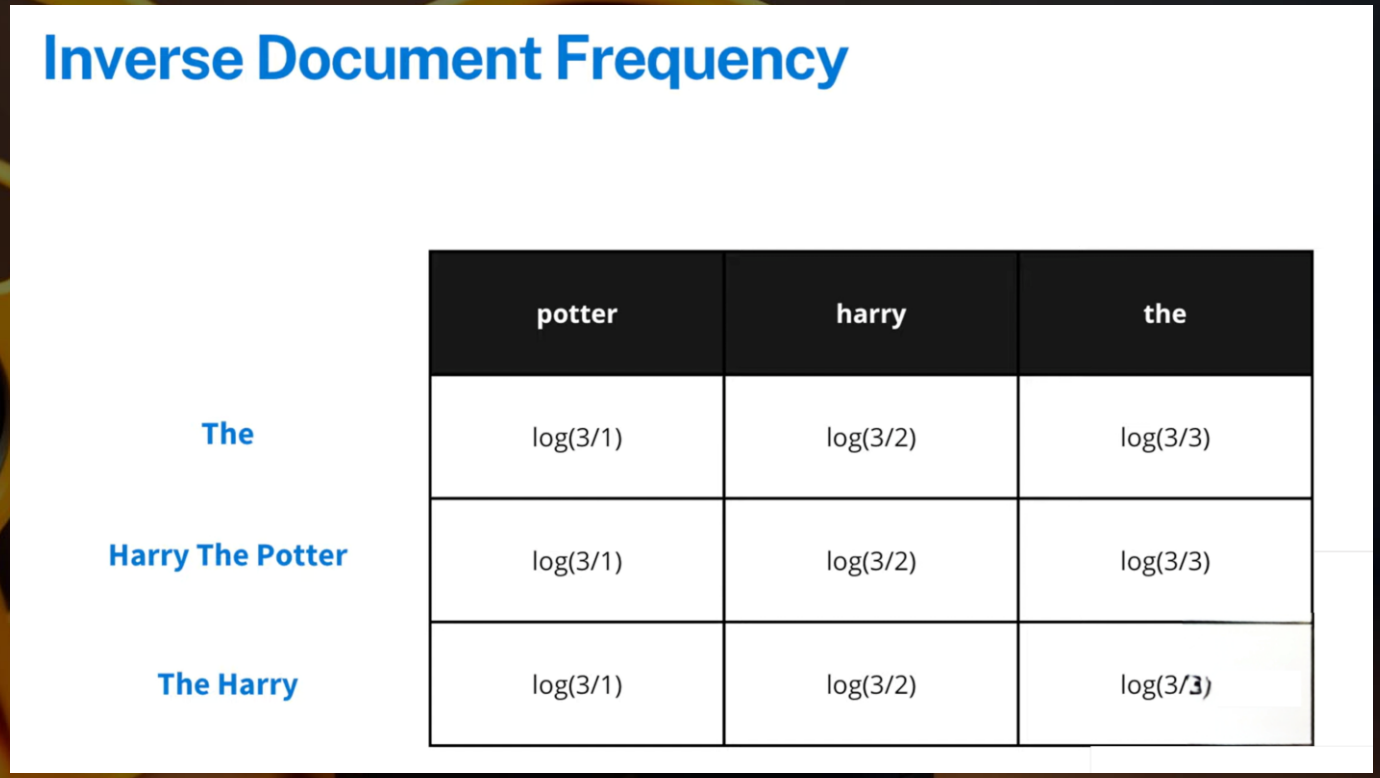
```python
import re

def clean_title(title):
    return re.sub("[^a-zA-Z0-9 ]", "", title)

movies["Clean_title"] = movies["title"].apply(clean_title)
```

```python
movies
```

3. Algorithm Selection:

Choose recommendation algorithms based on project requirements (collaborative filtering, content-based filtering, hybrid models).

Using ‘TfidfVectorizer’ from ‘sklearn’ module.

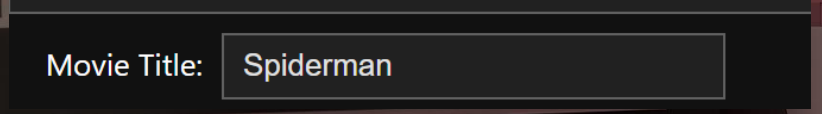


4. User Interface Design:

Develop an intuitive user interface to present recommendations to users.

Using ‘widgets’ from ‘ipywidgets’ library, to create an interactive search menu box.

```
import ipywidgets as widgets
from IPython.display import display

movie_input = widgets.Text(
    value = "Toy Story",
    description = "Movie Title: ",
    disabled = False
)

movie_list = widgets.Output()

def on_type(data):
    with movie_list:
        movie_list.clear_output()
        #display(data)
        title = data["new"]
        if len(title) > 5:
            display(search(title))

movie_input.observe(on_type, names = 'value')
display(movie_input, movie_list)
```



Result

The Movie Recommendation System (MRS) was designed and implemented to enhance user experience by providing personalized movie suggestions based on their preferences. The system utilized collaborative filtering and content-based filtering techniques to analyze user behavior and movie attributes.

Movie Title: Iron man 2010

	score	title	genres
14628	64.533909	Iron Man 2 (2010)	Action|Adventure|Sci-Fi|Thriller|IMAX
20513	46.091422	Thor: The Dark World (2013)	Action|Adventure|Fantasy|IMAX
19678	45.392430	Iron Man 3 (2013)	Action|Sci-Fi|Thriller|IMAX
12425	40.789039	Incredible Hulk, The (2008)	Action|Sci-Fi
20018	39.905463	Wolverine, The (2013)	Action|Adventure|Fantasy|Sci-Fi
16312	36.784484	Thor (2011)	Action|Adventure|Drama|Fantasy|IMAX
13460	36.372598	Transformers: Revenge of the Fallen (2009)	Action|Adventure|Sci-Fi|IMAX
16725	35.668612	Captain America: The First Avenger (2011)	Action|Adventure|Sci-Fi|Thriller|War
13226	32.498069	Fast & Furious (Fast and the Furious 4, The) (...	Action|Crime|Drama|Thriller
13277	32.169398	X-Men Origins: Wolverine (2009)	Action|Sci-Fi|Thriller

Conclusion

In conclusion, the development and implementation of the Recommendation System projects have produced important discoveries and developments in the field of customized content distribution. The main objective of the project was to improve user experience by making personalized movie recommendations, and the outcomes demonstrate how effective the approach used was.