**AI Mini Project**

**Title:** Movie recommendation system using content-based filtering

**Aim:**  
Develop a movie recommendation system using content-based filtering based on movie descriptions and genres.

**Requirement:**

* python 3.x
* jupyter notebook
* pandas
* scikit-learn

**Theory:**  
**What is a recommendation system?**  
a recommendation system is an artificial intelligence technique that provides suggestions to users based on their preferences, behavior, or similarity to other users. in this project, we use *content-based filtering*, which compares the description of movies using natural language processing.

**Why recommendation system?**

* **entertainment:** helps users find relevant content from large collections like netflix or imdb.
* **e-commerce:** improves user engagement by suggesting products.
* **personalization:** offers custom experiences for each user.
* **data utilization:** makes use of unstructured data like reviews and descriptions to deliver smart choices.

**Libraries used:**

**1. pandas:**  
used for loading and managing structured data like movie titles and descriptions.

**2. sklearn (scikit-learn):**  
used for feature extraction with tfidfvectorizer and calculating similarity scores with cosine\_similarity.

**3. numpy:**  
used internally by pandas/sklearn for matrix operations and numerical computation.

**Code:**

import pandas as pd

from sklearn.feature\_extraction.text import TfidfVectorizer

from sklearn.metrics.pairwise import cosine\_similarity

# sample dataset

data = {

'movie': ['inception', 'avatar', 'titanic', 'the prestige', 'interstellar'],

'description': [

'dreams mind-bending thriller',

'sci-fi alien planet',

'romantic ship sinking',

'magic rivalry thriller',

'space time travel'

]

}

df = pd.DataFrame(data)

# vectorize descriptions

tfidf = TfidfVectorizer()

tfidf\_matrix = tfidf.fit\_transform(df['description'])

# similarity calculation

similarity = cosine\_similarity(tfidf\_matrix)

# recommend movies similar to 'inception'

movie\_index = df[df.movie == 'inception'].index[0]

similar\_scores = list(enumerate(similarity[movie\_index]))

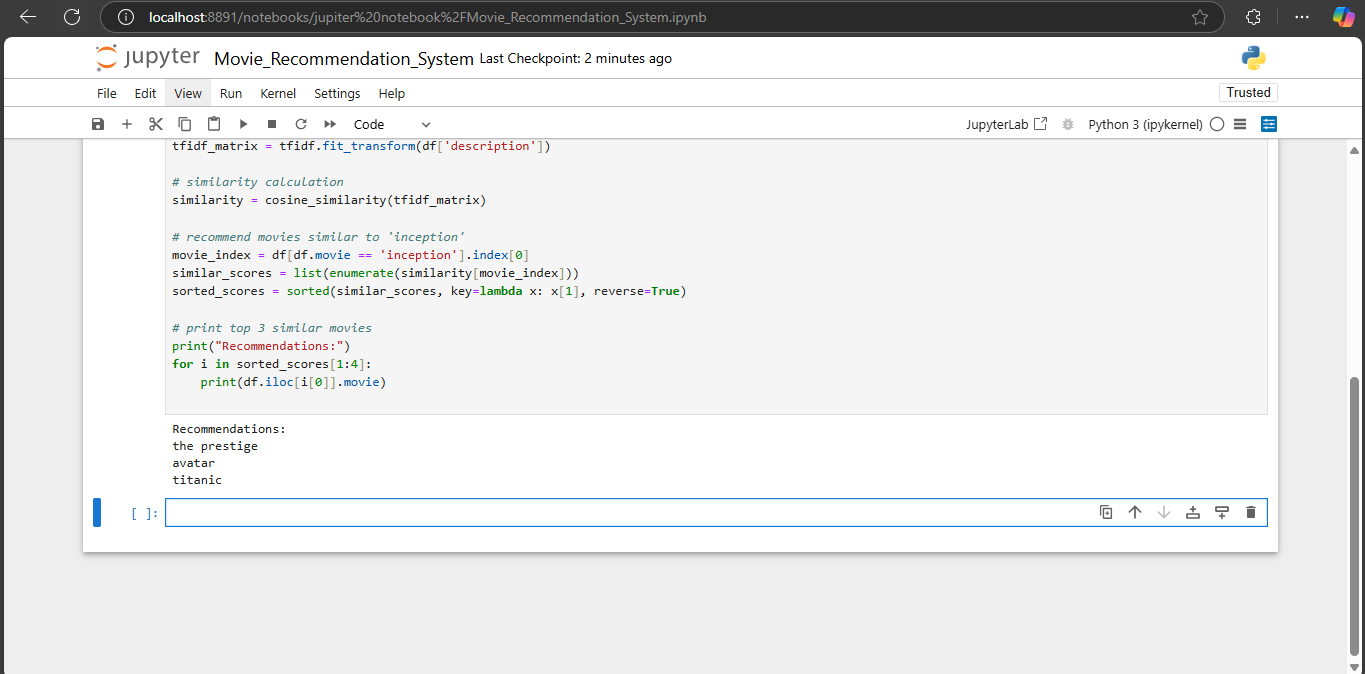
sorted\_scores = sorted(similar\_scores, key=lambda x: x[1], reverse=True)

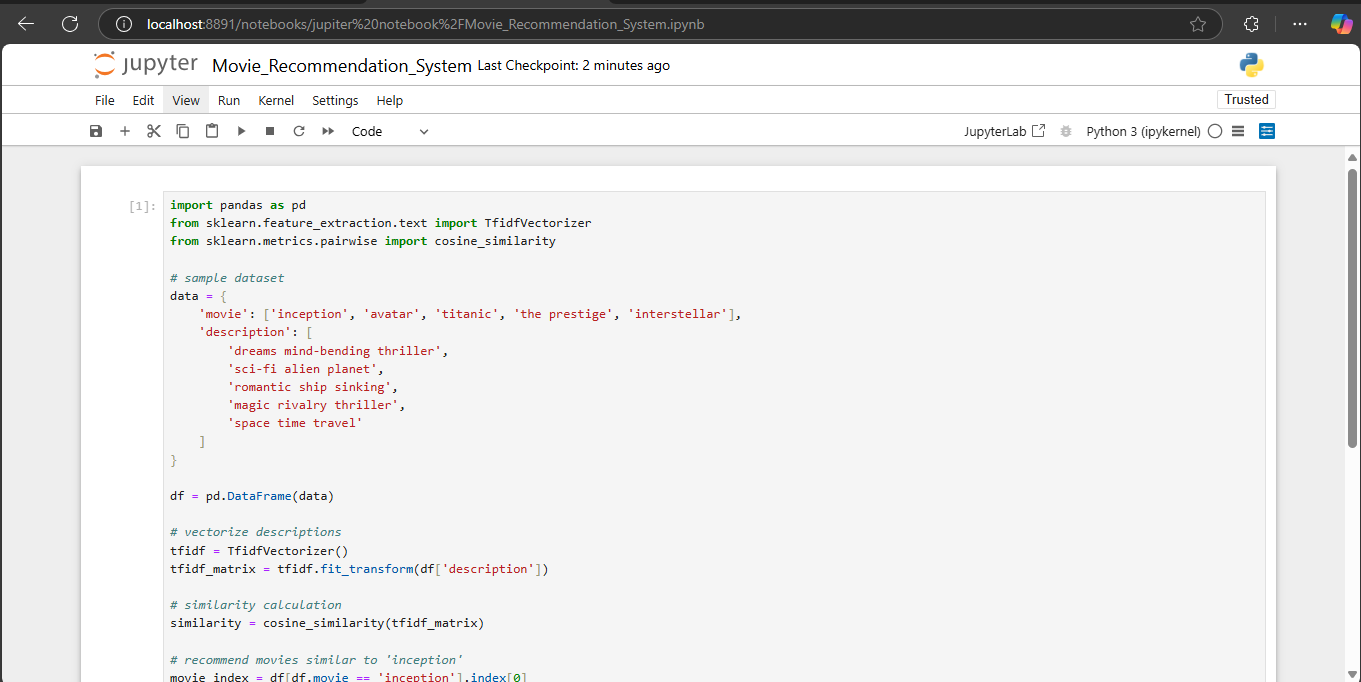
# print top 3 similar movies

print("Recommendations:")

for i in sorted\_scores[1:4]:

print(df.iloc[i[0]].movie)

**Output:**



**Conclusion:**  
hence, we successfully implemented a movie recommendation system using content-based filtering in python.