** Project Day - 19 **

[Movies Recommendation System]

Importing the dependencies

In [183…

```
import numpy as np
import pandas as pd
import difflib
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn. metrics.pairwise import cosine_similarity
```

Data Collection and Pre-Processing

In [186…

```
# Loading the dataset from the csv file to a pandas DataFrame
data=pd.read_csv('Movies.csv')
```

In [188…

```
# ptinting the first 5 rows of the dataframe
data.head()
```

Out[188…

	index	budget	genres	homepage	id	keywords	original_language	original_title	overview	popularity	...	runtime	spoken_languages	status	tagline	title	vote_average	v
0	0	237000000	Action Adventure Fantasy Science Fiction	http://www.avatarmovie.com/	19995	culture clash future space war space colony so...	en	Avatar	In the 22nd century, a paraplegic Marine is di...	150.437577	...	162.0	[{"iso_639_1": "en", "name": "English"}, {"iso...	Released	Enter the World of Pandora.	Avatar	7.2	
1	1	300000000	Adventure Fantasy Action	http://disney.go.com/disneypictures/pirates/	285	ocean drug abuse exotic island east india trad...	en	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	139.082615	...	169.0	[{"iso_639_1": "en", "name": "English"}]	Released	At the end of the world, the adventure begins.	Pirates of the Caribbean: At World's End	6.9	
2	2	245000000	Action Adventure Crime	http://www.sonypictures.com/movies/spectre/	206647	spy based on novel secret agent sequel mi6	en	Spectre	A cryptic message from Bond's past sends him o...	107.376788	...	148.0	[{"iso_639_1": "fr", "name": "Fran\u00e7ais"},...	Released	A Plan No One Escapes	Spectre	6.3	
3	3	250000000	Action Crime Drama Thriller	http://www.thedarkknightrises.com/	49026	dc comics crime fighter terrorist secret ident...	en	The Dark Knight Rises	Following the death of District Attorney Harve...	112.312950	...	165.0	[{"iso_639_1": "en", "name": "English"}]	Released	The Legend Ends	The Dark Knight Rises	7.6	
4	4	260000000	Action Adventure Science Fiction	http://movies.disney.com/john-carter	49529	based on novel mars medallion space travel pri...	en	John Carter	John Carter is a war-weary, former military ca...	43.926995	...	132.0	[{"iso_639_1": "en", "name": "English"}]	Released	Lost in our world, found in another.	John Carter	6.1	

5 rows × 24 columns

In [190…

```
# number of rows and columns in the data frame
data.shape
```

Out[190…

```
(4803, 24)
```

In [192…

```
selected_features=['genres','keywords','tagline','cast','director']
selected_features
```

Out[192…

```
['genres', 'keywords', 'tagline', 'cast', 'director']
```

In [194…

```
# replacing the null values with null string
for feature in selected_features:
    data[feature]=data[feature].fillna("")
```

In [196…

```
# combining all the 5 selected features
combined_features=data['genres']+' '+data['keywords']+' '+data['tagline']+' '+data['cast']+' '+data['director']
```

In [198…

```
combined_features
```

Out[198…

```
0       Action Adventure Fantasy Science Fiction cultu...
1       Adventure Fantasy Action ocean drug abuse exot...
2       Action Adventure Crime spy based on novel secr...
3       Action Crime Drama Thriller dc comics crime fi...
4       Action Adventure Science Fiction based on nove...
                              ...
4798    Action Crime Thriller united states\u2013mexic...
4799    Comedy Romance  A newlywed couple's honeymoon ...
4800    Comedy Drama Romance TV Movie date love at fir...
4801      A New Yorker in Shanghai Daniel Henney Eliza...
4802    Documentary obsession camcorder crush dream gi...
Length: 4803, dtype: object
```

In [200…

```
# converting the text data to feature vectors
vectorizer=TfidfVectorizer()
```

In [202…

```
feature_vectors= vectorizer.fit_transform(combined_features)
```

In [204…

```
print(feature_vectors)
```

```
  (0, 201)	0.07860022416510505
  (0, 274)	0.09021200873707368
  (0, 5274)	0.11108562744414445
  (0, 13599)	0.1036413987316636
  (0, 5437)	0.1036413987316636
  (0, 3678)	0.21392179219912877
  (0, 3065)	0.22208377802661425
  (0, 5836)	0.1646750903586285
  (0, 14378)	0.33962752210959823
  (0, 16587)	0.12549432354918996
  (0, 3225)	0.24960162956997736
  (0, 14271)	0.21392179219912877
  (0, 4945)	0.24025852494110758
  (0, 15261)	0.07095833561276566
  (0, 16998)	0.1282126322850579
  (0, 11192)	0.09049319826481456
  (0, 11503)	0.27211310056983656
  (0, 13349)	0.15021264094167086
  (0, 17007)	0.23643326319898797
  (0, 17290)	0.20197912553916567
  (0, 13319)	0.2177470539412484
  (0, 14064)	0.20596090415084142
  (0, 16668)	0.19843263965100372
  (0, 14608)	0.15150672398763912
  (0, 8756)	0.22709015857011816
  :	:
  (4801, 403)	0.17727585190343229
  (4801, 4835)	0.24713765026964
  (4801, 17266)	0.28860981849329476
  (4801, 13835)	0.27870029291200094
  (4801, 13175)	0.28860981849329476
  (4801, 17150)	0.3025765103586468
  (4801, 3511)	0.3025765103586468
  (4801, 13948)	0.3025765103586468
  (4801, 7269)	0.3025765103586468
  (4802, 11161)	0.17867407682173203
  (4802, 4518)	0.16784466610624255
  (4802, 2129)	0.3099656128577656
  (4802, 4980)	0.16078053641367315
  (4802, 6155)	0.18056463596934083
  (4802, 3436)	0.21753405888348784
  (4802, 4528)	0.19504460807622875
  (4802, 1316)	0.1960747079005741
  (4802, 12989)	0.1696476532191718
  (4802, 4371)	0.1538239182675544
  (4802, 6417)	0.21753405888348784
  (4802, 4608)	0.24002350969074696
  (4802, 2425)	0.24002350969074696
  (4802, 3654)	0.262512960498006
  (4802, 5367)	0.22969114490410403
  (4802, 6996)	0.5700048226105303
```

Cosine Similarity

In [207…

```
# getting the similarity scores using cosine similarity
similarity=cosine_similarity(feature_vectors)
```

In [209…

```
print(similarity)
```

```
[[1.         0.07219487 0.037733   ... 0.         0.         0.        ]
 [0.07219487 1.         0.03281499 ... 0.03575545 0.         0.        ]
 [0.037733   0.03281499 1.         ... 0.         0.05389661 0.        ]
 ...
 [0.         0.03575545 0.         ... 1.         0.         0.02651502]
 [0.         0.         0.05389661 ... 0.         1.         0.        ]
 [0.         0.         0.         ... 0.02651502 0.         1.        ]]
```

In [211…

```
similarity.shape
```

Out[211…

```
(4803, 4803)
```

In [213…

```
# getting the movie name from the user
movie_name=input('Enter Your Favourite movie name : ')
```

In [215…

```
list_of_all_titles=data['title'].tolist()
```

In [217…

```
# print(list_of_all_titles)
```

In [219…

```
# finding the close match for the movie name given by the user
find_close_match= difflib.get_close_matches(movie_name,list_of_all_titles)
```

In [221…

```
print(find_close_match)
```

```
['Iron Man', 'Iron Man 3', 'Iron Man 2']
```

In [223…

```
close_match=find_close_match[0]
```

In [225…

```
print(close_match)
```

```
Iron Man
```

In [227…

```
index_of_the_movie= data[data.title==close_match]['index'].values[0]
print(index_of_the_movie)
```

```
68
```

In [229…

```
# getting a list of similar movies
similarity_score=list(enumerate(similarity[index_of_the_movie]))
```

In [231…

```
len(similarity_score)
```

Out[231…

```
4803
```

In [233…

```
# shorting the movies baesd on their similarity score
sorted_similar_movies = sorted(similarity_score,key =lambda x:x[1],reverse=True)
```

In [235…

```
# print the name of similar movies based on the index

print('Movies suggested for you : \n')

i=1

for movie in sorted_similar_movies:
    index=movie[0]
    title_from_index =data[data.index ==index]['title'].values[0]
    if i<30:
        print(i, ".",title_from_index)
        i+=1
```

```
Movies suggested for you :

1 . Iron Man
2 . Iron Man 2
3 . Iron Man 3
4 . Avengers: Age of Ultron
5 . The Avengers
6 . Captain America: Civil War
7 . Captain America: The Winter Soldier
8 . Ant-Man
9 . X-Men
10 . Made
11 . X-Men: Apocalypse
12 . X2
13 . The Incredible Hulk
14 . The Helix... Loaded
15 . X-Men: First Class
16 . X-Men: Days of Future Past
17 . Captain America: The First Avenger
18 . Kick-Ass 2
19 . Guardians of the Galaxy
20 . Deadpool
21 . Thor: The Dark World
22 . G-Force
23 . X-Men: The Last Stand
24 . Duets
25 . Mortdecai
26 . The Last Airbender
27 . Southland Tales
28 . Zathura: A Space Adventure
29 . Sky Captain and the World of Tomorrow
```

Movie Recommendation System

In [240…

```
movie_name= input('Enter your fevourite movie name: ')
list_of_all_titles=data['title'].tolist()
find_close_match=difflib.get_close_matches(movie_name, list_of_all_titles)
close_match =find_close_match[0]
index_of_the_movie=data[data.title ==close_match]['index'].values[0]

similarity_score=list(enumerate(similarity[index_of_the_movie]))

sorted_similar_movies =sorted(similarity_score,key=lambda x:x[1], reverse =True)

print('Movies suggested for you : \n')

i=1

for movie in sorted_similar_movies:
    index=movie[0]
    title_from_index =data[data.index ==index]['title'].values[0]
    if i<30:
        print(i, ".",title_from_index)
        i+=1
```

```
Movies suggested for you :

1 . Batman
2 . Batman Returns
3 . Batman & Robin
4 . The Dark Knight Rises
5 . Batman Begins
6 . The Dark Knight
7 . A History of Violence
8 . Superman
9 . Beetlejuice
10 . Bedazzled
11 . Mars Attacks!
12 . The Sentinel
13 . Planet of the Apes
14 . Man of Steel
15 . Suicide Squad
16 . The Mask
17 . Salton Sea
18 . Spider-Man 3
19 . The Postman Always Rings Twice
20 . Hang 'em High
21 . Spider-Man 2
22 . Dungeons & Dragons: Wrath of the Dragon God
23 . Superman Returns
24 . Jonah Hex
25 . Exorcist II: The Heretic
26 . Superman II
27 . Green Lantern
28 . Superman III
29 . Something's Gotta Give
```

END - PROJECT

In [ ]: