

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter Movie-recommender-system Last Checkpoint: 08/27/2023 (autosaved)



Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted

Python 3 (ipykernel)

Code

```
In [3]: import pandas as pd
import numpy as np
```

```
In [4]: movies = pd.read_csv("tmdb_5000_movies.csv")
credits = pd.read_csv("tmdb_5000_credits.csv")
```

```
In [5]: movies.head()
```

```
Out[5]:
```

	budget	genres	homepage	id	keywords	original_language	original_title	overview	popularity	production_compar
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}, {"id": 14, "name": "Fantasy"}]	http://www.avatarmovie.com/	19995	[{"id": 1463, "name": "culture clash"}, {"id": 1464, "name": "culture clash"}]	en	Avatar	In the 22nd century, a paraplegic Marine is di...	150.437577	[{"name": "Ingeni Film Partners", "id": 28}
1	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "name": "Fantasy"}]	http://disney.go.com/disneypictures/pirates/	285	[{"id": 270, "name": "ocean"}, {"id": 726, "name": "na..."}]	en	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	139.082615	[{"name": "Walt Dis Pictures", "id": 2}, {"id": 28, "name": "Action"}]
2	245000000	[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}, {"id": 14, "name": "Fantasy"}]	http://www.sonypictures.com/movies/spectre/	206647	[{"id": 470, "name": "spy"}, {"id": 1464, "name": "culture clash"}]	en	Spectre	A cryptic message from Bond's past	107.376788	[{"name": "Colum Pictures", "id": 28}

Desktop/Movie Recommender system

Movie-recommender-system

localhost:8888/notebooks/Desktop/Movie%20Recommender%20system/Movie-recommender-system.ipynb

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter

Movie-recommender-system

Last Checkpoint: 08/27/2023 (autosaved)

Python 3 (ipykernel)

Logout

FileEditViewInsertCellKernelWidgetsHelp

ca...

In [6]: credits.head()

Out[6]:

	movie_id	title	cast	crew
0	19995	Avatar	[{"cast_id": 242, "character": "Jake Sully", "...	[{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	[{"cast_id": 4, "character": "Captain Jack Spa...	[{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	[{"cast_id": 1, "character": "James Bond", "cr...	[{"credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises	[{"cast_id": 2, "character": "Bruce Wayne / Ba...	[{"credit_id": "52fe4781c3a36847f81398c3", "de...
4	49529	John Carter	[{"cast_id": 5, "character": "John Carter", "c...	[{"credit_id": "52fe479ac3a36847f813eaa3", "de...

In [7]: movies=movies.merge(credits,on='title')

In [8]: movies=movies[['movie_id','title','overview','genres','keywords','cast','crew']]

In [9]: movies.head()

Out[9]:

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	[{"id": 1463, "name": "culture clash"}, {"id": ...	[{"cast_id": 242, "character": "Jake Sully", "...	[{"credit_id": "52fe48009251416c750aca23", "de...

Type here to search

01:13 PM

04-09-2023

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

```
In [10]: movies.isnull().sum()

Out[10]: movie_id    0
         title      0
         overview    3
         genres      0
         keywords    0
         cast        0
         crew        0
         dtype: int64

In [11]: movies.dropna(inplace=True)

In [12]: movies.duplicated().sum()

Out[12]: 0

In [13]: movies.iloc[0].genres

Out[13]: '[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}, {"id": 14, "name": "Fantasy"}, {"id": 878, "name": "Science Fiction"}]'

In [14]: import ast
         def convert(obj):
             lst = []
             for i in ast.literal_eval(obj):
                 lst.append(i['name'])
```

UPDATE
 Read [the migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
 Don't show anymore

```
In [14]: import ast
def convert(obj):
    lst = []
    for i in ast.literal_eval(obj):
        lst.append(i['name'])
    return lst
```

```
In [15]: movies['genres'] = movies['genres'].apply(convert)
```

```
In [16]: movies['keywords']=movies['keywords'].apply(convert)
```

```
In [17]: movies.head()
```

Out[17]:	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[{"cast_id": 242, "character": "Jake Sully", "...	[{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[{"cast_id": 4, "character": "Captain Jack Spa...	[{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[{"cast_id": 1, "character": "James Bond", "cr...	[{"credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	[{"cast_id": 2, "character": "Bruce Wayne / Ba...	[{"credit_id": "52fe4781c3a36847f81398c3", "de...

```
In [18]: import ast
def convert3(obj):
    lst = []
    c=0
    for i in ast.literal_eval(obj):
        if c !=3:
            lst.append(i['name'])
            c+=1
        else:
            break
    return lst
```

In [19]: `movies['cast']=movies['cast'].apply(convert3)`

In [20]: `movies.head()`

Out[20]:

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[Sam Worthington, Zoe Saldana, Sigourney Weave...	[{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[Johnny Depp, Orlando Bloom, Keira Knightley, ...	[{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[Daniel Craig, Christoph Waltz, Léa Seydoux, R...	[{"credit_id": "54805967c3a36829b5002c41", "de...

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter Movie-recommender-system Last Checkpoint: 08/27/2023 (autosaved)



Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted

Python 3 (ipykernel)



```
In [21]: import ast
def director(obj):
    lst = []
    for i in ast.literal_eval(obj):
        if i['job']=='Director':
            lst.append(i['name'])
            break
    return lst
```

```
In [22]: movies['crew']=movies['crew'].apply(director)
```

```
In [23]: movies.head()
```

Out[23]:	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[Sam Worthington, Zoe Saldana, Sigourney Weave...	[James Cameron]
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[Johnny Depp, Orlando Bloom, Keira Knightley, ...	[Gore Verbinski]
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[Daniel Craig, Christoph Waltz, Léa Seydoux, R...	[Sam Mendes]
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	[Christian Bale, Michael Caine, Gary Oldman, A...	[Christopher Nolan]
4	49529	John Carter	John Carter is a war-weary, former military ca...	[Action, Adventure, Science Fiction]	[based on novel, mars, medallion, space travel...	[Taylor Kitsch, Lynn Collins, Samantha Morton, ...	[Andrew Stanton]



Desktop/Movie Recommender s

Movie-recommender-system -


+

localhost:8888/notebooks/Desktop/Movie%20Recommender%20system/Movie-recommender-system.ipynb

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore











jupyter

Movie-recommender-system Last Checkpoint: 08/27/2023 (autosaved)

 Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)



Code

In [24]:

movies['overview']=movies['overview'].apply(lambda x:x.split())

In [25]:

movies.head()

Out[25]:

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[Sam Worthington, Zoe Saldana, Sigourney Weave...	[James Cameron]
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[Johnny Depp, Orlando Bloom, Keira Knightley, ...	[Gore Verbinski]
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[Daniel Craig, Christoph Waltz, Léa Seydoux, R...	[Sam Mendes]
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	[Christian Bale, Michael Caine, Gary Oldman, A...	[Christopher Nolan]
4	49529	John Carter	[John, Carter, is, a, war-weary,, former, mili...	[Action, Adventure, Science Fiction]	[based on novel, mars, medallion, space travel...	[Taylor Kitsch, Lynn Collins, Samantha Morton,...	[Andrew Stanton]

In [26]:

movies['genres']=movies['genres'].apply(lambda x:[i.replace(" ", "")for i in x])
movies['keywords']=movies['keywords'].apply(lambda x:[i.replace(" ", "")for i in x])
movies['cast']=movies['cast'].apply(lambda x:[i.replace(" ", "")for i in x])
movies['crew']=movies['crew'].apply(lambda x:[i.replace(" ", "")for i in x])


In [27]:





movies.head()

Out[27]:

	movie_id	title	overview	genres	keywords	cast	crew
--	----------	-------	----------	--------	----------	------	------

Type here to search



 ENG 01:14 PM 04-09-2023

movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...	[Action, Adventure, Fantasy, ScienceFiction]	[cultureclash, future, spacewar, spacecolony, ...]	[SamWorthington, ZoeSaldana, SigourneyWeaver, ...]
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...	[Adventure, Fantasy, Action]	[ocean, drugabuse, exoticisland, eastindiadrad...	[JohnnyDepp, OrlandoBloom, KeiraKnightley, Ste...
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...	[Action, Adventure, Crime]	[spy, basedonnovel, secretagent, sequel, mi6, ...]	[DanielCraig, ChristophWaltz, LéaSeydoux, Ralp...
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...	[Action, Crime, Drama, Thriller]	[dccomics, crimefighter, terrorist, secretiden...	[ChristianBale, MichaelCaine, GaryOldman, Anne...
4	49529	John Carter	[John, Carter, is, a, war-weary,, former, mili...	[Action, Adventure, ScienceFiction]	[basedonnovel, mars, medallion, spacetravel, p...	[TaylorKitsch, LynnCollins, SamanthaMorton, Wi...

```
In [29]: movies.head()
```


UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter Movie-recommender-system Last Checkpoint: 08/27/2023 (autosaved)



Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted

Python 3 (ipykernel)

Code

In [29]: `movies.head()`

	movie_id	title	overview	genres	keywords	cast	crew	tags
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...	[Action, Adventure, Fantasy, ScienceFiction]	[cultureclash, future, spacewar, spacecolony, ...	[SamWorthington, ZoeSaldana, SigourneyWeaver, ...	[JamesCameron]	[In, the, 22nd, century,, a, paraplegic, Marin...
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...	[Adventure, Fantasy, Action]	[ocean, drugabuse, exoticisland, eastindiatrad...	[JohnnyDepp, OrlandoBloom, KeiraKnightley, Ste...	[GoreVerbinski]	[Captain, Barbossa,, long, believed, to, be, d...
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...	[Action, Adventure, Crime]	[spy, basedonnovel, secretagent, sequel, mi6, ...	[DanielCraig, ChristophWaltz, LéaSeydoux, Ralp...	[SamMendes]	[A, cryptic, message, from, Bond's, past, send...
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...	[Action, Crime, Drama, Thriller]	[dccomics, crimefighter, terrorist, secretiden...	[ChristianBale, MichaelCaine, GaryOldman, Anne...	[ChristopherNolan]	[Following, the, death, of, District, Attorney...
4	49529	John Carter	[John, Carter, is, a, war-weary,, former, mili...	[Action, Adventure, ScienceFiction]	[basedonnovel, mars, medallion, spacetravel, p...	[TaylorKitsch, LynnCollins, SamanthaMorton, Wi...	[AndrewStanton]	[John, Carter, is, a, war-weary,, former, mili...

In [30]: `new_df = movies[['movie_id', 'title', 'tags']]`

In [31]: `new_df.head()`

	movie_id	title	tags
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...



Desktop/Movie Recommender system - x

Movie-recommender-system - x


+

localhost:8888/notebooks/Desktop/Movie%20Recommender%20system/Movie-recommender-system.ipynb

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter

Movie-recommender-system Last Checkpoint: 08/27/2023 (autosaved)

 Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)

4

49529

John Carter [John, Carter, is, a, war-weary,, former, mili...

In [32]:

```
new_df['tags'] = new_df['tags'].apply(lambda x: " ".join(x))
```

C:\Users\Abhishek\AppData\Local\Temp\ipykernel_24524\3089450492.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
new_df['tags'] = new_df['tags'].apply(lambda x: " ".join(x))
```

In [33]:

```
new_df.head()
```

Out[33]:

	movie_id	title	tags
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...
2	206647	Spectre	A cryptic message from Bond's past sends him o...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...
4	49529	John Carter	John Carter is a war-weary, former military ca...

In [34]:

```
new_df['tags'][0]
```

Out[34]:

'In the 22nd century, a paraplegic Marine is dispatched to the moon Pandora on a unique mission, but becomes torn between following orders and protecting an alien civilization. Action Adventure Fantasy ScienceFiction cultureclash future spacewar spacecolony society spac

Windows

Type here to search



 ENG

01:14 PM
04-09-2023


Desktop/Movie Recommender s...Movie-recommender-system - x

localhost:8888/notebooks/Desktop/Movie%20Recommender%20system/Movie-recommender-system.ipynb












UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter

Movie-recommender-system Last Checkpoint: 08/27/2023 (autosaved)

 Logout

FileEditViewInsertCellKernelWidgetsHelp



Code

449529John CarterJohn Carter is a war-weary, former military ca...

In [34]:

new_df['tags'][0]

Out[34]:

'In the 22nd century, a paraplegic Marine is dispatched to the moon Pandora on a unique mission, but becomes torn between following ord
ers and protecting an alien civilization. Action Adventure Fantasy ScienceFiction cultureclash future spacewar spacecolony society spac
etravel futuristic romance space alien tribe alienplanet cgi marine soldier battle loveaffair antiwar powerrelations mindandsoul 3d Sam
Worthington ZoeSaldana SigourneyWeaver StephenLang MichelleRodriguez GiovanniRibisi JoelDavidMoore CCHPounder WesStudi LazAlonso Dileep
Rao MattGerald SeanAnthonyMoran JasonWhyte ScottLawrence KellyKilgour JamesPatrickPitt SeanPatrickMurphy PeterDillon KevinDorman Kelso
Henderson DavidVanHorn JacobTomuri MichaelBlain-Rozgay JonCurry LukeHawker WoodySchultz PeterMensah SoniaYee JahnelCurfman IlramChoi Ky
laWarren LisaRoumain DebraWilson ChrisMala TaylorKibby JodieLandau JulietLamm CullenB.Madden JosephBradyMadden FrankieTorres AustinWilso
n SaraWilson TamicaWashington-Miller LucyBriant NathanMeister GerryBlair MatthewChamberlain PaulYates WrayWilson JamesGaylyn MelvinLeno
ClarkIII CarvonFutrell BrandonJelkes MicahMoch HanniyahMuhammad ChristopherNolen ChristaOliver AprilMarieThomas BravitaA.Threatt ColinB
leasdale MikeBodnar MattClayton NicoleDionne JamieHarrison AllanHenry AnthonyIngruber AshleyJeffery DeanKnowsley JosephMika-Hunt TerryW
otary KaiPantano LoganPithyou StuartPollock Raja GarethRuck RhianSheehan T.J.Storm JodieTaylor AliciaVela-Bailey RichardWhiteside Nikie
Zambo JuleneRenee JamesCameron'

In [35]:

new_df['tags'] = new_df['tags'].apply(lambda x:x.lower())


C:\Users\Abhishek\AppData\Local\Temp\ipykernel_24524\3214958533.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
new_df['tags'] = new_df['tags'].apply(lambda x:x.lower())

In [36]:

import nltk

Type here to search



01:14 PM

04-09-2023