Experiment No. 5

Aim: Develop Content based social media analytics model for business.

Theory:

This project is a Movie Recommender System which recommends movies to the user:- Based on

1.Popularity:- Correlation
2.Collaborative Filtering:- KNN
3.Content Based :- NLP and Cosine Similarity

Using Classification algorithm , NLP and Correlation.

Methods Used:

- Feature Engineering:-
  Feature engineering is the process of transforming raw data into features that are suitable for machine learning models. In other words, it is the process of selecting, extracting, and transforming the most relevant features from the available data to build more accurate and efficient machine learning models.

- Machine Learning:-
  Machine learning (ML) is a discipline of artificial intelligence (AI) that provides machines with the ability to automatically learn from data and past experiences while identifying patterns to make predictions with minimal human intervention.

- Data Visualization:-
  Data visualization is the representation of data through use of common graphics, such as charts, plots, infographics, and even animations. These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand.

- NLP:-
  Natural language processing (NLP) is a machine learning technology that gives computers the ability to interpret, manipulate, and comprehend human language. Organizations today have large volumes of voice and text data from various communication channels like emails, text messages, social media news feeds, video, audio, and more. They use NLP software to automatically process this data, analyze the intent or sentiment in the message, and respond in real time to human communication.

Result:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

```
[6] movie=pd.read_csv('movies.csv')
movie.head()
```

	movieId	title	genres
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	2	Jumanji (1995)	Adventure|Children|Fantasy
2	3	Grumpier Old Men (1995)	Comedy|Romance
3	4	Waiting to Exhale (1995)	Comedy|Drama|Romance
4	5	Father of the Bride Part II (1995)	Comedy

Next steps: View recommended plots

```
[7] df=pd.read_csv('ratings.csv')
df.head()
```

	userId	movieId	rating	timestamp
0	1	1	4.0	964982703
1	1	3	4.0	964981247
2	1	6	4.0	964982224
3	1	47	5.0	964983815
4	1	50	5.0	964982931

```
rate_count=df.groupby('title')['rating'].count().sort_values(ascending=False)
rate_count
```

```
title
Forrest Gump (1994)                          329
Shawshank Redemption, The (1994)             317
Pulp Fiction (1994)                          307
Silence of the Lambs, The (1991)             279
Matrix, The (1999)                           278
                                            ...
King Solomon's Mines (1950)                    1
King Solomon's Mines (1937)                    1
King Ralph (1991)                              1
King Kong Lives (1986)                         1
À nous la liberté (Freedom for Us) (1931)      1
Name: rating, Length: 9719, dtype: int64
```

Recommending user based on user ratings using correlation

```python
movie_mat=pd.pivot_table(df,values='rating',columns='title',index='userId')
movie_mat.head(10)
```

title userId	'71 (2014)	'Hellboy': The Seeds of Creation (2004)	'Round Midnight (1986)	'Salem's Lot (2004)	'Til There Was You (1997)	'Tis the Season for Love (2015)	'burbs, The (1989)	'night Mother (1986)	(500) Days of Summer (2009)	*batteries not included (1987)	...	Zulu (2013)	[REC] (2007)	[REC]^2 (2009)	[REC]^3 3 Génesis (2012)	anohana: The Flower We Saw That Day - The Movie (2013)	eXistenZ (1999)	xXx (2002)	xXx: State of the Union (2005)	¡Three Amigos! (1986)	À nous la liberté (Freedom for Us) (1931)
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	4.0	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	1.0	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Suggesting movie for Matrix, The (1999)

```python
x=corr['Matrix, The (1999)'].sort_values(ascending=False)
x
```

```
title
Haywire (2011)                                           1.0
Back to the Beach (1987)                                 1.0
Demolition (2016)                                        1.0
Gerry (2002)                                             1.0
Wolf Children (Okami kodomo no ame to yuki) (2012)       1.0
                                                        ...
Zoom (2015)                                              NaN
Zulu (2013)                                              NaN
[REC]³ 3 Génesis (2012)                                  NaN
anohana: The Flower We Saw That Day - The Movie (2013)   NaN
À nous la liberté (Freedom for Us) (1931)                NaN
Name: Matrix, The (1999), Length: 9719, dtype: float64
```

Recommending user using Collaborative Filtering

```python
df.head()
```

	userId	movieId	rating	title	genres
0	1	1	4.0	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	5	1	4.0	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
2	7	1	4.5	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
3	15	1	2.5	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
4	17	1	4.5	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy

Suggesting movie for Matrix, The (1999)

```
movie_mat_new[movie_mat_new['title']=='Matrix, The (1999)']
```

userId	title	1	2	3	4	5	6	7	8	9	...	601	602	603	604	605	606	607	608	609	610
5512	Matrix, The (1999)	5.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	...	5.0	0.0	5.0	0.0	0.0	5.0	5.0	5.0	0.0	5.0

1 rows × 611 columns

Recommending User based on content based filtering

```
[32] df=pd.read_csv('tmdb_movies.csv')
```

```
df.head()
```

	budget	genres	homepage	id	keywords	original_language	original_title	overview
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.avatarmovie.com/	19995	[{"id": 1463, "name": "culture clash"}, {"id":...	en	Avatar	In the 22nd century, a paraplegic Marine is di...
1	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "...	http://disney.go.com/disneypictures/pirates/	285	[{"id": 270, "name": "ocean"}, {"id": 726, "na...	en	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...
2	245000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.sonypictures.com/movies/spectre/	206647	[{"id": 470, "name": "spy"}, {"id": 818, "name...	en	Spectre	A cryptic message from Bond's past sends him o...
3	250000000	[{"id": 28, "name": "Action"}, {"id": 80, "nam...	http://www.thedarkknightrises.com/	49026	[{"id": 849, "name": "dc comics"}, {"id":	en	The Dark Knight Rises	Following the death of District Attorney Harve...

```
df1=df[['original_title','overview']]
df1.head()
```

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

```
[48] dic={'index':ind,'value':val}
     rec_df=pd.DataFrame(dic)
```

```
rec=rec_df.merge(df1,on='index')
rec.head()
```

	index	value	original_title	overview
0	0	0.024995	Avatar	In the 22nd century, a paraplegic Marine is di...
1	1	0.000000	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...
2	2	0.000000	Spectre	A cryptic message from Bond's past sends him o...
3	3	1.000000	The Dark Knight Rises	Following the death of District Attorney Harve...
4	4	0.010433	John Carter	John Carter is a war-weary, former military ca...

Next steps: View recommended plots

```
[50] a=rec.sort_values(by='value',ascending=False).head(7)
```

```
[52] print(a['original_title'])

3                       The Dark Knight Rises
65                           The Dark Knight
299                           Batman Forever
428                           Batman Returns
1359                                  Batman
3854    Batman: The Dark Knight Returns, Part 2
119                            Batman Begins
Name: original_title, dtype: object
```

Conclusion:

The Content-Based Social Media Analytics Model provides a powerful tool for businesses to analyze various content types like text, emoticons, images, audio, and video. Through advanced techniques such as NLP and machine learning, the model identifies topics, detects issues, tracks trends, and analyzes sentiment. It also delves into multimedia content, offering insights into consumer preferences and engagement patterns. Ultimately, this model enables businesses to make informed decisions, optimize marketing strategies, and enhance customer experiences in the ever-evolving digital world.