collaborating Filtering recommendation systems(Final)

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

from mpl\_toolkits.mplot3d import Axes3D

from sklearn.feature\_extraction.text import CountVectorizer

from sklearn.metrics.pairwise import cosine\_similarity

# Loading the Cleaned E-Commerce Data of 23k Customers

data = pd.read\_csv('cleaned\_data.csv')

data.head()

# Calling the Variables from E-Commerce Data of 23k Customers

data = data[['Clothing ID', 'Age', 'Recommended IND', 'Review Text', 'Rating']]

data.dropna(subset=['Review Text'], inplace=True)

data.reset\_index(drop=True, inplace=True)

data['Review Text'] = data['Review Text'].str.lower()

# Calling the Clothing Items by using 'Clothing ID'here

Clothing\_id = data['Clothing ID'].unique()

# Designing the Customer and Clothing Matrix by focusing on preferences of 'Clothing ID', 'Age' and 'Rating' through pivot table

Clothing\_Customer\_matrix = data.pivot\_table(index='Clothing ID', columns='Age', values='Rating', fill\_value=0)

# Shifting the data from pivot table to Numpy array to move towards cosine similarlity

Clothing\_Customer\_matrix = Clothing\_Customer\_matrix.to\_numpy()

# Cosine Similarlity for Collborative filtering by taking Customer and Clothing matrix

Clothing\_similarity = cosine\_similarity(Clothing\_Customer\_matrix)

# Designing DataFrame by using Similiarity matrix data here

Clothing\_similarity\_df = pd.DataFrame(

Clothing\_similarity,

index=Clothing\_id,

columns=Clothing\_id

)

# Display of the Recommended Clothing ID of Collborative Filtering

print(Clothing\_similarity\_df)

# Locating the recommendation by using the 'Review Text'prefernces from data

vectorizer = CountVectorizer(stop\_words='english')

review\_matrix = vectorizer.fit\_transform(data['Review Text'])

review\_similarity = cosine\_similarity(review\_matrix)

review\_similarity\_df = pd.DataFrame(review\_similarity, index=data.index, columns=data.index)

# Calculating the Combined Similarity

alpha = 0.7

combined\_similarity = alpha \* Clothing\_similarity\_df + (1 - alpha) \* review\_similarity\_df

# Using 'recommendations' DataFrame and duplicating Clothing IDs

recommendations = recommendations.drop\_duplicates(subset='Clothing ID', keep='first')

# Displaying recommendations for specific 'Clothing ID'

clothing\_id = 420

print("Top Recommendations for Clothing ID", clothing\_id)

print(recommendations[['Clothing ID', 'Age', 'Rating']])

# Grouping of the DataFrame by preference of 'Clothing ID' and exploring the average rating for each selected Clothing\_id

average\_ratings = recommendations.groupby('Clothing ID')['Rating'].mean().reset\_index()

# Designing Bar Plot to visualise the recommended Clothing ID

plt.figure(figsize=(9, 6))

plt.bar(average\_ratings['Clothing ID'], average\_ratings['Rating'], color='red')

plt.xlabel('Clothing ID')

plt.ylabel('Average Rating')

plt.title('Average Ratings for Clothing IDs')

plt.xticks(rotation=90)

plt.grid(axis='y')

plt.show()