import numpy as np

import pandas as pd

from sklearn.neighbors import NearestNeighbors

import matplotlib.pyplot as plt

import seaborn as sns

# Disable FutureWarnings

import warnings

warnings.filterwarnings('ignore', category=FutureWarning)

# Load the movie ratings data and the movie data from CSV files

ratings = pd.read\_csv('ratings.csv')

movies = pd.read\_csv('movies.csv')

# Create a mapping from movie IDs to strings

movie\_mapper = {}

movie\_inv\_mapper = {}

for i in range(len(movies)):

movie\_mapper[movies.loc[i, 'title']] = movies.loc[i, 'movieId']

movie\_inv\_mapper[movies.loc[i, 'movieId']] = movies.loc[i, 'title']

# Create a matrix of movie vectors

X = np.zeros((len(movies), 100))

for i in range(len(movies)):

movie\_id = movies.loc[i, 'movieId']

movie\_ratings = ratings[ratings['movieId'] == movie\_id]['rating']

X[i] = movie\_ratings.mean()

# Define a function to convert a movie ID to a string

def movie\_id\_to\_string(movie\_id):

return movies.loc[movies['movieId'] == movie\_id, 'title'].to\_string()

# Define a function to find similar movies using KNN

def find\_similar\_movies(movie\_id, X, k, metric='cosine', show\_distance=False):

"""Finds the k most similar movies to the given movie using KNN.

Args:

movie\_id: The ID of the movie to find similar movies for.

X: The matrix of movie vectors.

k: The number of similar movies to find.

metric: The metric to use to measure similarity.

show\_distance: Whether to show the distance between each movie and the given

movie.

Returns:

A list of the k most similar movie IDs.

"""

neighbour\_ids = []

movie\_ind = movie\_mapper[movie\_id]

movie\_vec = X[movie\_ind]

KNN = NearestNeighbors(n\_neighbors=k, algorithm="brute", metric=metric)

KNN.fit(X)

neighbour = KNN.kneighbors(movie\_vec, return\_distance=show\_distance)

for i in range(0, k):

n = neighbour.item(i)

neighbour\_ids.append(movie\_inv\_mapper[n])

return neighbour\_ids

# Example usage:

# Get the movie ID for "The Shawshank Redemption"

movie\_id = movie\_mapper["The Shawshank Redemption"]

# Find the 5 most similar movies to "The Shawshank Redemption"

similar\_ids = find\_similar\_movies(movie\_id, X, k=5)

# Print the titles of the similar movies

for i in similar\_ids:

print(movie\_titles[i])