```
import json #import necessary packages
import numpy as np
from sklearn.metrics.pairwise import cosine_distances
from sklearn.metrics.pairwise import euclidean_distances
import pandas as pd

# Initialize empty lists to store JSON objects
movieData = []

# Load JSON data from the file line by line
with open('imdb_movies_2000to2022.prolific.json', 'r') as json_file:
    for line in json_file:
        movie = json.loads(line)  # Parse each JSON object individually
        movieData.append(movie)

genre = set() #sets to store actor id and genre
actorId = set()
for movie in movieData:  #extract ids from data
    for id, _ in movie["actors"]:
        actorId.add(id)

    for j in movie["genres"]: #extracts genres from data
        genre.add(j)

# Convert sets to lists
actorId = list(actorId)
genre = list(genre)

df = pd.DataFrame(0, columns=genre, index=actorId) #create data frame

for movie in movieData:
    for id, _ in movie["actors"]: #adds genres to data frame
        for j in movie["genres"]:
            df.at[id, j] += 1

queryId = "nm1165110" #query actor id(chris hemsworth)
qActor = df.loc[queryId] #retrieve genre count for query id

cosMatrix = cosine_distances([qActor], df) #calculate cos distance

simIndices = np.argsort(cosMatrix.flatten())[1:11] # 10 most similar based on cos
distance
simActors = df.iloc[simIndices]
print("10 most similar actors based on cos distance")
print(simActors)
```