Movie Recommendation System - Fetching Data via API

This documentation provides a detailed explanation of using Python to interact with the IMDb API to fetch movie data based on a cast name. The fetched data is converted into a DataFrame for further analysis or processing. The script demonstrates the use of the http.client library for making HTTPS requests and the pandas library for handling data.

Prerequisites

1. **IMDb API Access:** You need access to the IMDb API through RapidAPI.
2. **RapidAPI Key:** Ensure you have a valid RapidAPI key.
3. **Python Libraries:** Install the required libraries:
    - http.client: Standard library for HTTP requests.
    - ssl: For secure HTTPS connections.
    - json: To handle JSON responses.
    - pandas: For data manipulation (install using pip install pandas).

Script Description

Code

```
import http.client
import ssl
import json
import pandas as pd


# Create HTTPS connection
conn = http.client.HTTPSConnection(
    "imdb_api4.p.rapidapi.com",
    context=ssl._create_unverified_context()
)


# Set headers
headers = {
    'x-rapidapi-key': "bae03c19damshcdbccf3d5f16607p15b0f4jsndb1ebace834b",
    'x-rapidapi-host': "imdb_api4.p.rapidapi.com"
}


# Send request
conn.request("GET", "/get_movies_by_cast_name", headers=headers)
```

```
# Get response
res = conn.getresponse()
data = res.read()

# Decode response
json_data = json.loads(data.decode("utf-8"))

# Convert JSON to DataFrame
df = pd.DataFrame(json_data)

# Print DataFrame
print(df)
```

Steps

1. **Establish HTTPS Connection:**
    - An HTTPSConnection object is created using the IMDb API's host and a secure SSL context.
2. **Set API Headers:**
    - The x-rapidapi-key and x-rapidapi-host headers are used for authentication.
3. **Send GET Request:**
    - A GET request is sent to the endpoint /get_movies_by_cast_name to fetch movie data.
4. **Handle API Response:**
    - The response is read and decoded into JSON format using json.loads().
5. **Convert JSON to DataFrame:**
    - The JSON data is converted to a Pandas DataFrame for easier handling and analysis.
6. **Print DataFrame:**
    - The DataFrame is printed to display the fetched movie data.

Usage

1. **Replace API Key:**
    - Replace the placeholder RapidAPI key (bae03c19damshcdbccf3d5f16607p15b0f4jsndb1ebace834b) with your valid API key.
2. **Run the Script:**
    - Execute the script in your Python environment.
3. **Analyze Data:**
    - Use the DataFrame (df) to analyze or manipulate the movie data as needed.

Notes

- The ssl._create_unverified_context() is used here to bypass SSL verification. Use this cautiously in a production environment.
- Ensure your API key is kept secure and not hardcoded in scripts shared publicly.

Dependencies

- Python 3.x
- Pandas library (pip install pandas)

License

This script is for educational and personal use. The data fetched from the IMDb API is subject to IMDb's terms of use.

Acknowledgments

- IMDb API provided by RapidAPI.
- Python community for the libraries used in this script.