James Smith

N01400606

COP4710 - Data Modeling

Final Project Report

# **Movie Recommendation Engine**

# **Objectives:**

The goal of this project was to create a movie recommendation engine based on a user's input. The program using pre-parsed CSV files from a lighter version of the IMBD database. This project aims to return intelligent and useful recommendations to the user using a similarity matrix to calculate similarity scores.

#### How to run:

To run the program, please ensure that all the CSV files, and the JAR file are included in the same directory. Do not change any names or remove any files, they are all necessary for the program to run. The program was built using Java 8, so that version of the JDK is required.

For compilation, please the contents of the zip into your favorite IDE (or navigate to the path via the command line), compile, and run. Give the program roughly 5 minutes to perform data insertions into MYSQL.

### 1/0:

No extra files are needed to run the program, and the program will also not produce any output files. All input is given via the keyboard, and the recommendations are sent directly to the console.

### Design:

First, a connection is made to MYSQL, and a database and multiple tables that correlate to the CSV files in the zip are created. Then, the data is taken from the CSV files that were parsed in part 1 of the project and inserted into tables, then using MYSQL queries, the data is then stored in a class called Movie, which holds all the relevant information needed for proper recommendations to be evaluated.

The program, after a loading period, shows a small menu which tells the user to enter a movie title, or the number 0 to exit the program. The program will either return a list of recommendations upon a valid input and after a brief evaluation, show an error message on an invalid input, or exit upon receiving a 0. If the program receives a 0, the database is dropped, and the program then exits.

### **Sample Run and Closing Thoughts:**

This is a small sample run of the program from beginning of execution, to exit it includes multiple valid inputs, and invalid input, and the exit command. I believe the engine works best when given movies that have no sequels, due to the fact that the similarity Matrix may overwhelm the results with sequels, as you can see in the results of a search of "Star Wars". So I would recommend users choose movies that may not have multiple sequels in order to get better recommendations from the engine.

```
Loading movies...
Welcome to my movie recommendation engine!
Recommendations for Tombstone
1 - The Legend of Hell's Gate: An American Conspiracy
3 - Young Guns
4 - 英雄
5 - Heaven's Gate
Recommendations for Avatar
2 - Aliens
3 - The Fifth Element
1 - The Empire Strikes Back
3 - Star Wars: Episode III - Revenge of the Sith
4 - Star Wars: Episode II - Attack of the Clones
5 - Star Wars: Episode I - The Phantom Menace
```

```
Enter a movie title to receive a list of recommendations, or 0 to exit:
1 - Training Day
2 - American Gangster
3 - GoodFellas
5 - Colombiana
Enter a movie title to receive a list of recommendations, or 0 to exit:
Recommendations for The Social Network
2 - The Town
3 - Blackhat
4 - Gone Girl
5 - The Dark Knight Rises
Enter a movie title to receive a list of recommendations, or 0 to exit:
Recommendations for Looper
1 - Minority Report
2 - Southland Tales
3 - Sin City: A Dame to Kill For
4 - Terminator Genisys
5 - I Robot
Recommendations for Inception
1 - The Dark Knight Rises
4 - Southland Tales
5 - Mad Max: Fury Road
Goodbye!
```