291 Mini Project 2 Report

This document contains the report for the Database mini project 2 for CMPUT 291 Winter 2022.
Team members:

1. Shehraj Singh (ccid: shehraj)
2. Gurveer Singh Sohal (ccid: gurveers)
3. Manraj Singh (ccid: manraj3)

Report is divided into 4 Sections

1. ***Design***
2. ***Group work division***
3. ***User guide***
4. ***Testing strategy by the group***

1. Design

Main.py sets up the database by using tsv-2-json.py to convert the tsv files to json, and the load-json.py to use the json files to add data to the database. It takes in the port number as user input.

Main.py then allows the user to choose the following options:

1. Add a cast or crew member
2. Add a movie
3. Search a cast/crew member
4. Search genres
5. Search titles
6. Exit

Each task is handled by a function defined in a different file

1. Add a cast or crew member: add_cast.py

   The function addCast() takes in the database as input. It validates user input and adds the cast or crew member to the database.

2. Add a movie: add_movie.py

   The function addMovie() takes in the database as input. It validates user input and adds the movie to the database.

3. Search a cast/crew member: search_cast.py

   The function searchCast() takes in the database as input. Then it searches the database for a name matching the user input.

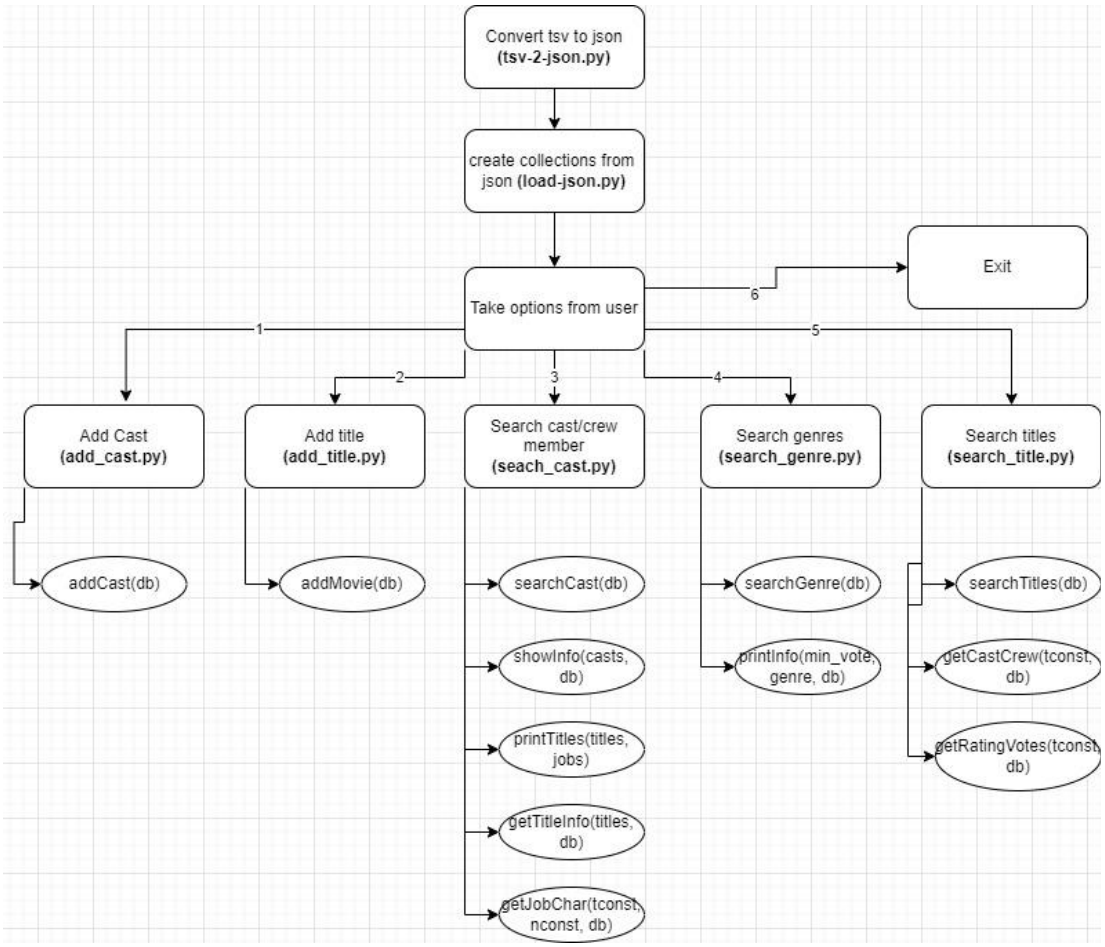
   The function showInfo() takes in a list of cast members and the database as input. It displays information about all the cast members in the list.

4. Search genres: search_genre.py

   The function searchGenre() takes in the database as input and searches the database based on the genre entered by the user. Then it displays the movies that match the genre and have a minimum number of votes as given by the user.

5. Search titles: search_titles.py

The function searchTitles() takes in the database as input. Then it displays the movies that match all the keywords given by the user.



2. Group work division

Implementing separate files allowed to implement different functions for each member. Division of work as follows:

Gurveer - load-json.py, add_movie.py, add_cast.py
Shehraj - tsv-2-json.py, search_cast.py, search_genre.py
Manraj - search_genre.py, search_title.py

The report document, README and documentation of code was done jointly. We had regular meetings to discuss ideas and include each other's work into our own and how the files connect. We tested our own code and also tested each other's code. Whenever we found bugs, we let the other person know, and fixed them.

Time spent:
Gurveer - 5-6 hours
Shehraj - 6-7 hours

Manraj - 5-6 hours
Used GitHub and discord for collaboration.

3. User guide

Start the app using “python3 main.py” in linux/mac or “python main.py” in windows. Then you will be asked to enter integers based on what choices you want to make. Enter the integers after reading the options.

You can choose from following:
As a user you can:

1. Add cast/crew members
2. Add title
3. Search title
4. Search cast/crew member
5. Search genres
6. Exit

Selecting an option would take you to appropriate screen which would require further inputs to complete the task. You can exit out once you are finished.

4. Testing Strategy

Testing was done in two ways, one was individual testing on each component of the project as we were working on it to ensure a working product with no bugs. Secondly, a full round test using the whole project to ensure when combined the project had no issues and conflicts with other parts.
An issue we actually encountered during testing was that our solutions were running slow, to combat this we implemented text indexing. Changing the order of stages such as projecting after matching allowed us to shorten the run time of our solutions.

5. References

Acknowledgement of information used but no direct taking of information as not allowed, we understood the methods used and applied them in our solution so we felt it necessary to comment on where we obtained that information.

https://www.youtube.com/watch?v=9N6a-VLBa2I
https://www.geeksforgeeks.org/python-tsv-conversion-to-json/
https://stackoverflow.com/questions/6502541/mongodb-query-multiple-collections-at-once
https://stackoverflow.com/questions/25163658/mongodb-return-true-if-document-exists
https://www.statology.org/mongodb-max-value/
https://stackoverflow.com/questions/28968660/how-to-convert-a-pymongo-cursor-cursor-into-a-dict
https://stackoverflow.com/questions/39815265/mongo-sort-by-string-value-that-is-actually-number
https://pymongo.readthedocs.io/en/stable/examples/collations.html