# **Project Report: Movie Recommendation System**

## **Project Title: Movie Recommendation System (Based on Genre and Ratings)**

## **Problem We Are Solving**

Choosing a movie to watch can be difficult because there are so many options. This project solves that problem by recommending movies based on the genres the user prefers and their average ratings. It does not use machine learning. Instead, it uses basic data filtering and sorting logic to suggest movies.

## **Tools and Libraries Used**

* **pandas**: For reading and working with CSV data
* **re (regular expressions)**: For extracting genre names from text
* **Google Colab**: An online platform to write and run Python code
* **files.upload()**: A Colab function to upload CSV files

## **Datasets Used**

Three small CSV files from the MovieLens Dataset were used:

1. **ratings\_small.csv** – Contains user ratings for movies
2. **movies\_metadata.csv** – Contains movie titles and genres
3. **links\_small.csv** – Connects different movie IDs from the datasets

Dataset Source:  
 Kaggle - The Movies Dataset  
<https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset>

## **What the Code Does (Solution Explanation)**

1. **Load the data**: The program reads the three CSV files using pandas.
2. **Clean the data**: It removes missing or incorrect values and converts columns to the correct data types.
3. **Merge the data**: It combines the ratings, metadata, and link files into one dataset using common keys.
4. **Extract genres**: The genres column contains JSON-like strings. The code uses regular expressions to pull out the actual genre names.
5. **User input**: The user is asked to enter one or more genres they like.
6. **Filter and sort**: The program filters the movies that match the selected genres and sorts them by rating in descending order.
7. **Display results**: It shows the top 5 highest-rated movies that match the user's preferences.

## **Final Result**

When the user selects their favorite genres, the system displays five top-rated movies that belong to those genres. This is a simple and useful tool to help people discover good movies without using complex algorithms or machine learning.