

# Movie Recommendation System - Project Report

## Introduction

In this project, a movie recommendation system was developed to suggest movies to users based on their viewing preferences and historical data. The system utilizes data from movie and credit datasets and employs both collaborative filtering and content-based filtering methods to enhance recommendation accuracy. The project was designed using Python and key data analysis libraries.

## Objective

The primary objective of this project was to:

- Analyze large datasets to extract meaningful insights for generating movie recommendations.
- Improve the accuracy of recommendations by 15% over a baseline model.
- Discover and utilize key factors that impact user preferences.

## Datasets Used

- **Movie Dataset:** Includes details about movies such as title, genre, release year, and plot summary.
- **Credit Dataset:** Contains information about the cast, crew, and other production-related details for each movie.
- Both datasets were cleaned and processed before being used for analysis and modeling.

The project involved the following steps:

## Data Preprocessing

- **Data Cleaning:** Handled missing values, duplicates, and incorrect entries in both datasets.
- **Data Transformation:** Applied one-hot encoding for categorical variables such as genres, actors, and directors.
- **Feature Engineering:** Extracted important features like genre popularity and actor/director influence.
- **Collaborative Filtering:** Used user-item interaction data to recommend movies based on similar users' preferences.
- **Content-Based Filtering:** Recommended movies based on attributes of the movie itself (e.g., genre, cast, director).
- Combined both techniques to enhance overall recommendation accuracy.

During the analysis, several important insights were discovered:

- **Genre Popularity:** Genres like action and drama were more likely to be favored in recommendations due to their higher frequency in user interactions.
- **Actor/Director Influence:** Movies featuring popular actors and directors had a higher probability of being recommended.
- **User Behavior Patterns:** Identified patterns in user preferences based on genre, viewing habits, and movie ratings.

## Conclusion

This movie recommendation system successfully improved recommendation accuracy by 15% over the baseline model, utilizing both collaborative and content-based filtering techniques. Key insights were derived from user preferences and viewing patterns, enhancing the system's ability to provide relevant suggestions. The project demonstrates the potential of using data analysis techniques in building intelligent recommendation systems.