

# **Capstone Project Netflix**

**By** Sheetal Kharat

# **Index**

- 1. Executive Summary**
- 2. Problem Statement**
- 3. Dataset Information**
- 4. Objectives**
- 5. Methodology**
  - 5.1 Data Collection and Preparation
  - 5.2 Similarity Calculation
  - 5.3 Recommendation Algorithm
  - 5.4 Testing and Validation
- 6. Implications**
- 7. Conclusion**

# **1. Executive Summary**

This report outlines the development of a movie recommendation engine aimed at predicting user preferences and enhancing user engagement on streaming platforms. By analyzing user activity and ratings, the recommendation engine will provide personalized movie suggestions, thereby improving user satisfaction and retention.

## **2. Problem Statement**

Understanding customer behavior is vital for businesses in the digital age. Predictive analytics plays a significant role across various domains, including e-commerce, entertainment, and social media. Recommendation engines have emerged as critical tools that not only provide personalized content but also strategically enhance user interaction with platforms.

With the rise of Over-The-Top (OTT) services like Netflix and Amazon Prime, there is an increasing demand for efficient recommendation systems that cater to user preferences. These platforms analyze user behavior patterns and suggest relevant content, which can significantly influence user satisfaction and platform loyalty.

This project aims to create a recommendation engine from the ground up that will analyze user ratings and preferences to recommend movies tailored to individual users.

### 3. Dataset Information

The dataset utilized for this project comprises the following key components:

1. **ID**: Unique identifiers for customers and movies.
2. **Rating**: User ratings for all movies, reflecting their preferences.
3. **Genre**: Categories to which each movie belongs, facilitating genre-based recommendations.
4. **Movie Name**: Titles corresponding to the movie IDs, providing context for the recommendations.

## 4. Objectives

The primary objectives of this project are as follows:

1. **Identify Popular Genres:** Analyze the dataset to determine the most liked genres among users based on ratings.
2. **Develop a Recommendation Model:** Create a model that accurately identifies the best-suited movies for individual users in various genres.
3. **Analyze Rating Trends:** Investigate which genres receive the highest and lowest ratings, providing insights into user preferences and satisfaction.

## **5. Methodology**

### **5.1 Data Collection and Preparation**

- Gather data from the movie dataset, ensuring it is clean and structured for analysis.
- Create a pivot table that represents user ratings across different movie genres.

### **5.2 Similarity Calculation**

- Employ cosine similarity to quantify how similar each movie is to one another based on user ratings.
- Store the similarity scores in a DataFrame for easy access during the recommendation process.

### **5.3 Recommendation Algorithm**

- Develop a function that takes a movie ID as input and returns a list of recommended movies based on their similarity scores.
- Ensure the recommendations exclude the input movie to provide a diverse list of suggestions.

### **5.4 Testing and Validation**

- Test the recommendation engine using a subset of movie IDs to evaluate the accuracy and relevance of the recommendations.
- Gather user feedback to refine the model further.

## 6. Implications

The successful implementation of this recommendation engine has several potential implications:

1. **Enhanced User Experience:** By providing personalized movie suggestions, users are more likely to discover content that resonates with their interests, improving overall satisfaction with the platform.
2. **Increased Engagement:** Personalized recommendations can lead to higher user engagement, as users spend more time exploring suggested content.
3. **Competitive Advantage:** Organizations leveraging effective recommendation engines can gain a competitive edge in the crowded OTT market, attracting and retaining more users.
4. **Data-Driven Insights:** The analysis of user ratings and genre preferences can inform content acquisition strategies, helping platforms curate a library that meets user demand.



## **7. Conclusion**

The development of a movie recommendation engine is a significant step towards understanding and predicting user behavior in the entertainment sector. By leveraging data analytics and machine learning techniques, this project aims to create a valuable tool that enhances user engagement and satisfaction on streaming platforms.