

# PROJECT REPORT

## Movie Recommendation System using Python (Content-Based Filtering)

### 1. Introduction

A recommendation system suggests relevant items to users. This project implements a Content-Based Movie Recommendation System that recommends movies similar to a given movie using metadata such as overview, genres, keywords, cast, and crew.

### 2. Objective

To build a movie recommendation engine using TF-IDF vectorization and cosine similarity.

### 3. Technologies Used

- Python
- Jupyter Notebook
- Pandas
- Scikit-Learn
- TMDB 5000 Movies & Credits Dataset

### 4. Methodology

#### Step 1: Import Libraries

Pandas for data handling, Scikit-Learn for TF-IDF vectorization and cosine similarity.

#### Step 2: Load Datasets

Loaded tmdb\_5000\_movies.csv and tmdb\_5000\_credits.csv.

#### Step 3: Merge Datasets

Merged on movie title.

#### Step 4: Data Preprocessing

Extracted director name and combined features including overview, genres, keywords, cast, and director.

## 5. Feature Extraction using TF-IDF

Converted combined text features into numerical vectors.

## 6. Similarity Calculation

Cosine similarity was used to compute similarity between movies.

## 7. Recommendation Function

A function was built to return top 10 similar movies based on cosine similarity.

## 8. Testing the System

Tested with:

- Avatar
- The Dark Knight

Both provided relevant top-10 movie recommendations.

## 9. Results

The system provides accurate movie recommendations based on textual content similarity.

## 10. Conclusion

The project successfully demonstrates a content-based recommendation approach using Python.

## 11. Future Enhancements

- Add collaborative filtering
- Deploy using Streamlit or Flask
- Integrate posters and ratings
- Use deep learning embeddings