**Student Name:** ROSHNI. B

**Register Number:**510623104087

**Institution:** C. ABDUL HAKEEM COLLEGE OF

ENGINEERING AND TECHNOLOGY

**Department:** COMPUTER SCIENCE ENGINEERING

**Date of Submission:**08-05-2025

**Github Repository Link:** <https://github.com/Rosh-0906/Movie-recommendations.git>

# 1. Problem Statement

The challenge lies in improving AI-driven personalization for movie recommendations. By analysing behavioural patterns, ratings, and sentiment analysis, the system can make refined and engaging matches, enhancing user experience.

# 2. Project Objectives

* Develop a hybrid recommendation model combining collaborative filtering and content-based filtering.
* Increase recommendation accuracy using user sentiment analysis from reviews. - Ensure adaptability with self-learning AI models that refine suggestions dynamically.
* Compare various machine learning approaches to optimize performance

# 3. Flowchart of the Project Workflow

**1.Start** – Initiate the process.

1. **Data Collection** – Gather movie details and user preferences.
2. **Data Processing** – Clean and encode relevant data features.

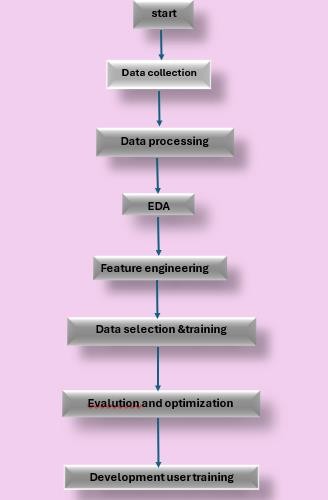
**4**.**EDA & Insights** – Analyse user-movie interactions to find patterns.

**5**.**Feature Engineering** – Enhance input data for improved recommendations.

**6**. **Model Building** – Implement and evaluate recommendation algorithms.

**7**.**Deployment & Visualization** – Present AI-driven movie suggestions.

**8**.**User Feedback Loop** – Refine recommendations based on user engagement.



# 4. Data Description

The dataset includes:

1.Movie metadata (genres, ratings, cast, director, release year).

2.User preferences (watch history, likes, dislikes).

3.Review sentiments (text analysis for emotional tone).

4.Streaming platform trends (popularity rankings).

# 5. Data Preprocessing

* Handling missing values (imputation).

* Encoding categorical features (genres, cast).

* Processing text for sentiment analysis (tokenization, stop-word removal).

* Feature scaling & normalization for enhanced model performance

# 6. Exploratory Data Analysis (EDA)

-Genre-watch frequency trends to recognize popular categories.

-Correlation mapping of ratings and user engagement.

-User behavior segmentation through clustering.

-Sentiment-based polarity scoring for enhanced personalization.

**7. Feature Engineering:**

-User-movie interaction matrices for collaborative filtering.

-TF-IDF feature extraction for analyzing movie descriptions. -Embedding-based similarity measures\*\* for deep personalization.

- Contextual tagging\*\* to ensure dynamic AI adaptability.

**8. Model Building :**

Comparison between approaches:

Matrix Factorization (SVD) for user-movie matchmaking.

Neural Collaborative Filtering for deep-learning personalization.

Hybrid Approach integrating multiple filtering techniques.

Graph-based modelsanalyzing structured connectivity insights.

# 9. Visualization of Results & Model Insights

Heatmaps for similarity tracking.

* Feature importance rankings showing influential parameters.

-Precision-recall curves assessing recommendation accuracy.

* Confusion matrices for classification validation.

# 10. Tools and Technologies Used

* **Sufiya Firdouse**: Python – The primary language for data handling and model building.
* **Roshni**: Pandas & NumPy – Efficient preprocessing and feature engineering.
* **Pooja Sri**: Matplotlib & Seaborn – Data visualization and exploratory data analysis (EDA).
* **Sandhiya**: Surprise & LightFM – Collaborative and hybrid recommendation model development.
* **Mudabbira Fathima**: TensorFlow / PyTorch – Deep learning models for advanced recommendations.
* **Monisha**: Flask / FastAPI – Backend development and API deployment for the recommendation system.

# 11. Team Members and Contributions

Here’s how our team:

**Team Members & Responsibilities:**

* **Sufiya Firdouse** – **Feature Engineering**: Extracting and transforming relevant attributes for model improvement.
* **Roshni** – **EDA (Exploratory Data Analysis)**: Identifying patterns, trends, and insights from the dataset.
* **Pooja Sri** – **Data Cleaning**: Handling missing values, inconsistencies, and preparing structured data for modeling.
* **Sandhiya** – **Model Development**: Implementing collaborative filtering and hybrid recommendation models.
* **Mudabbira Fathima** – **Advanced Model Optimization**: Exploring deep learning and fine-tuning recommendation algorithms.
* **Monisha** – **Documentation & Reporting**: Structuring workflows.