TASK NO: 4

TASK NAME: **Data Science Lifecycle Example**

**Entertainment: Audience Analysis & Content Recommendation**

**1. Problem Definition**

**Objective:** Develop a product that analyzes viewer preferences and streaming behavior to provide personalized content recommendations. This will enhance user engagement and satisfaction by tailoring content to individual interests.

**2. Data Collection**

Gather data from streaming platforms, social media interactions, user reviews, and viewing history. Collect both structured data and unstructured data. Ensure the data is comprehensive and representative of the user base to capture diverse preferences.

**3. Data Cleaning**

Remove duplicates, handle missing values, and standardize data formats to ensure consistency. Filter out irrelevant data points that do not contribute to understanding viewer preferences.

**4. Data Exploration**

Use visualizations to identify trends, patterns, and correlations in viewing behavior. Understand peak viewing times, popular genres, and user demographics to inform content recommendations.

**5. Data Preparation**

Create new features that capture user behavior, such as average viewing time or genre preferences. Scale and normalize data to ensure consistency and improve model performance. Divide data into training and testing sets to evaluate model performance accurately.

**6. Modeling**

Choose algorithms suitable for recommendation systems, such as collaborative filtering or content-based filtering. Train models on historical data to capture user preferences and viewing patterns. Fine-tune hyperparameters to improve model accuracy and performance.

**7. Evaluation**

Use evaluation metrics like precision, recall, and F1-score to assess model performance. Conduct A/B testing to compare different models and select the best-performing one.

**8. Deployment**

Integrate the model into the streaming platform’s recommendation engine. Implement real-time monitoring to track model performance and user feedback. Establish a pipeline for periodic model updates with new data to maintain accuracy.

**9. Monitoring and Maintenance**

Continuously collect user feedback to improve the recommendation system. Address data drift and model decay by regularly retraining the model with fresh data. Ensure user data is handled ethically, addressing privacy concerns and biases in data collection and analysis.