

Lab 1

Date 03-03-25
Page 1

LAB-1

Examples of any 5 datasets (explain) and write down the problem statement and its attributes

1. Student Performance and Behaviour dataset

- * The dataset involves key attributes necessary for exploring patterns, calculations and insights related to academic performance.

- * Attributes:

StudentID, First-name, Last-name, Email, Gender, Age, Department, Attendance, Midterm-score, Final-score, Quizzes-avg, total-score, grade, internet-at-home, parent-education-level, Family-income-level.

2. Netflix Interaction

- * Netflix recommendation is based on a combination of user behaviour analysis i.e. tracking watch history, browsing pattern, content based filtering. This tool allows users to group similar movies based on themes, ratings and user interactions.

- * Attributes:

N-id, Title, maingenres, subgenres, Release year, maturity rating, original audio, Recommendations

3. Top Spotify songs in 73 countries

- * The dataset presents the top songs currently trending for ~~time~~ over 73 countries. It identifies the most popular genres of music, analyse the change in rankings of the song over time, analyse the relationship between song features, and predict the future popularity of songs.

* Attributes:

Spotify-id, name, artists, daily-rank, daily-movement, weekly-movement, country, snapshot-date, popularity, is-explicit.

4. Daily food and Nutrition Dataset

- * This is designed for health-related machine learning models to analyse dietary patterns, nutritional intake, and their impact on health.

* Attributes:

date, user-id, food-item, category, calories, proteins, fat, carbohydrates, sugars, sodium, cholesterol, meal-type, water-intake.

5. Bollywood Movies

- * The default dataset contains details of 1698 hindi language movies released in India between in India between 2005 and 2017.

* Attributes:

genre, franchise status, debut contributors (actors, directors, music directors) and financial metrics (budget, revenue, no. of screens, etc).

- * Libraries used - pandas, sklearn