Movie Recommendation Report

Objective

This project implements a movie recommendation system inspired by the Pixie algorithm, which uses random walks on a user–movie graph to generate relevant suggestions. The system takes a movie title as input and recommends similar movies based on viewing patterns of other users.

Dataset Description

The system uses three datasets:

- `movies.csv`: Contains movie metadata including titles and IDs.

- `ratings.csv`: Contains user IDs, movie IDs, and their corresponding ratings.

- `users.csv`: Contains user information (used to construct user nodes in the graph).

These datasets were preprocessed to form a bipartite graph of users and movies.

Methodology

A graph was built where:

- Nodes = Users and Movies

- Edges = User-to-Movie ratings

A random walk starts from the selected movie node and alternates between users and other movies they have rated. This is repeated over multiple iterations and walk lengths. Movies frequently reached (excluding the start movie) are considered strong recommendations.

Key Results

Below are sample outputs from the system:

- Input: `Jurassic Park (1993)`

  Recommendations:

  - Mighty Aphrodite (1995)

  - Twelve Monkeys (1995)

  - Mission: Impossible (1996)

- Input: `Dead Man Walking (1995)`

  Recommendations:

  - Mighty Aphrodite (1995)

  - Star Trek V: The Final Frontier (1989)

  - Twelve Monkeys (1995)

These results show that the system consistently finds contextually similar or co-watched movies.

Insights

- Movies like "Mighty Aphrodite (1995)" appear frequently, suggesting it's commonly co-rated across diverse user paths.

- Random walks enable diverse but relevant recommendations even for obscure or less-rated movies.