Spotify Song Recommender

Dominykas Kepalas

Overview & How It Works

Project Snapshot

A CLI-based recommender that suggests similar Spotify tracks or lists the most popular songs in a chosen genre. The system operates on a cleaned Spotify dataset, scales and weights audio features, and applies a cosine k-nearest-neighbors model to identify similar tracks. Optional visual outputs include radar charts, feature histograms, a correlation heatmap, and 2D/3D UMAP projections.

Data Preparation

The raw Spotify dataset is preprocessed by removing unused metadata and retaining only relevant audio features. All numeric features are standardized prior to modeling to ensure fair similarity comparisons across dimensions.

Recommendation Logic

User input follows the format ’Artist - Title’. The system attempts an exact match and falls back to fuzzy string matching when necessary. Similarity is computed using cosine distance in a weighted feature space, with optional genre-based filtering and artist-specific recommendations.

Command-Line Interface

The application is launched via ‘python main.py‘. Users may request similarity-based recommendations or retrieve the most popular songs within a selected genre. Interactive prompts guide the workflow, and visual comparisons can be generated on demand.

Visual Analysis

Visualizations support exploratory analysis and interpretation. Histograms summarize feature distributions, the correlation heatmap reveals relationships between audio attributes, and UMAP projections illustrate the global structure of the song feature space.

Quick Start

Create a Python 3.10+ environment and install pandas, scikit-learn, rapidfuzz, matplotlib, and umap-learn. Ensure the cleaned dataset is available, then run ‘python main.py‘ to begin.