Medium Path Pipeline

Medium Path: Content-Based Filtering with Advanced Audio Features and Clustering

1. Data Input & Preprocessing:
- Load track metadata including full audio features (Mean, Variance, Skewness, Kurtosis, RMS Energy, Zero Crossing Rate, Tempo, Loudness) separately for vocal and non-vocal tracks.
- Normalize all numerical features using StandardScaler or RobustScaler.
- Encode categorical features (Genre, Artist) using embeddings or one-hot encoding.
- Apply PCA to reduce dimensionality of high-dimensional audio features.

2. Feature Engineering:
- Combine vocal and non-vocal track features into a joint feature vector.
- Cluster tracks and users using K-Means or Hierarchical Clustering on reduced features.
- Represent each user by cluster memberships or aggregated cluster centroids.

3. Model Selection & Training:
- Use Stratified K-Fold Cross-Validation for train-test splits.
- Train Random Forest classifiers or Gradient Boosting Machines on clustered feature sets.
- Integrate user preferences and cluster information for personalized recommendations.

4. Model Evaluation:
- Evaluate using Recall@K, F1-score, and cluster validity metrics.
- Conduct ablation studies comparing vocal-only, non-vocal-only, and combined feature sets.
- Analyze recommendation diversity and cluster cohesion.

Handling Vocal/Audio Tracks:
- Extract and preprocess vocal and audio features separately.
- Use separate pipelines for vocal and non-vocal features before merging.
- Experiment with feature importance to determine contributions from each track type.

ML Techniques Used:
- Random Forest
- Gradient Boosting Machines (GBM)
- K-Means Clustering
- Hierarchical Clustering
- PCA (Principal Component Analysis)
- Stratified K-Fold Cross-Validation
- Recall, F1-score
- One-hot Encoding / Embeddings
- Feature Normalization (StandardScaler, RobustScaler)