📑 Literature Review – The Million Playlist Dataset Challenge

1. What is the Dataset?
    - Released by Spotify Research (2018).
    - Contains 1 million user-generated playlists (24 million tracks).
    - Playlist metadata includes:
    - Playlist titles
    - Track IDs
    - Artist information
    - Album information
    - Track positions
    - The dataset is in JSON format (easy to parse with Python).

2. What was the Challenge?
    - Task: Automatic Playlist Continuation (APC)
    - Given a partial playlist, recommend additional tracks.
    - Hosted on the ACM RecSys Challenge 2018.
    - Evaluated recommendation quality using metrics such as:
    - R-precision
    - Normalized Discounted Cumulative Gain (NDCG)
    - Recommended Songs Clicks (RSC)

3. Key Methods Explored
    - Collaborative Filtering: Focuses on user-item interactions and matrix factorization.
    - Content-Based Filtering: Utilizes track audio features, metadata, and lyrics.
    - Hybrid Models: Combines collaborative and content-based approaches.
    - Sequence Models: Implements RNNs and session-based recommendations.
    - Use of deep learning embeddings for tracks.

4. Challenges & Difficulties
    - Cold Start Problem: New songs/playlists with little history.
    - Scalability: The dataset is extensive (1 million playlists).
    - Diversity vs. Accuracy: Balancing recommendations between popular hits and personalized tracks.
    - Evaluation: Different metrics favor different strategies, complicating assessment.

5. Relevance to My Project
    - My project also uses the Million Playlist Dataset (MPD) as a foundation.
    - I will extend this by:
        - ❖ Adding mood/context-awareness through lyrics and sentiment analysis.
        - ❖ Going beyond playlist continuation to develop mood-based playlist generation.
        - ❖ The challenge paper provides baseline methods (collaborative and content-based), which I aim to improve with NLP mood detection and hybrid models.

6. Key Takeaway
    - The MPD challenge established a benchmark for playlist recommendation research.
    - My project will build on this by making recommendations based not only on playlist history but also on user emotions (mood detection).