Dense Vectors, Dot Products, and One-Hot Encoding in the Spectral Neighbors Recommender

In the Spectral Neighbors Recommender, each song is represented as a dense vector that captures its underlying sonic identity. These vectors are built from numerical audio features such as MFCCs, spectral centroid, roll-off, flatness, zero-crossing rate, and tempo. Each value in the vector represents a measurable property of the track's sound, and together they form a compact mathematical representation of what the song "is." Dense vectors are powerful because they allow the system to express similarity between different pieces of audio not just by identity but by shared acoustic characteristics. Two songs that sound alike will have feature vectors that point in similar directions in the feature space.

The relationship between two dense vectors is measured using the dot product or more specifically the cosine similarity, which is a normalized form of the dot product. Cosine similarity computes the angle between two vectors and returns a value between −1 and 1. When the vectors point in the same direction, the value approaches 1, indicating a high degree of similarity. When they are orthogonal, the value is 0, showing no relation. In the context of this recommender, cosine similarity measures how closely two songs align in terms of their timbral and spectral profile rather than their absolute loudness or amplitude. Because all feature vectors are standardized, the cosine metric focuses purely on relative feature composition. This is what allows the system to identify songs that sound similar even when they differ slightly in volume, mixing, or dynamic range.

By contrast, one-hot encoding represents each song as a long vector filled with zeros except for a single one at the position corresponding to its unique ID. One-hot vectors are useful for identity but cannot express similarity because different songs always have orthogonal representations, meaning their dot product is zero. In a one-hot space, there is no way to tell that two songs might sound alike or belong to the same artist. Dense representations solve this limitation by encoding multiple acoustic attributes into continuous numerical dimensions. This enables meaningful comparison through cosine similarity, where proximity in vector space directly corresponds to similarity in sound.

In this prototype, the use of dense feature vectors and cosine similarity demonstrates the same conceptual foundation used in large-scale recommendation systems. Instead of representing songs as categorical IDs, they are expressed as continuous points in an audio feature space, and similarity is measured through vector alignment. This design allows the system to retrieve tracks that share timbral and rhythmic traits, producing coherent and musically relevant recommendations. It also lays the groundwork for future extensions where feature vectors could be learned embeddings rather than hand-crafted descriptors, creating a bridge between classical signal processing and modern machine learning approaches in music recommendation.