Capstone 3 - Final Project Report

Solving the Music Discovery Gap: A Deep Latent-Space Recommendation Engine

Dylan Brown

1. Problem Statement Streaming platforms face a "Cold Start" and "Long Tail" problem where new or niche songs are rarely recommended because they lack user interaction data. This project aims to build a content-based recommendation engine using Deep Learning to analyze the "DNA" of a song (audio features) rather than relying on popularity or user history.

2. Approach I utilized the Spotify Tracks Dataset, focusing on 12 key audio features (e.g., danceability, energy, tempo). I implemented a **Deep Autoencoder** neural network to compress these high-dimensional features into a 6-dimensional "Latent Space." By calculating **Cosine Similarity** within this space, the model identifies musically similar tracks regardless of their play count.

3. Key Findings * The Deep Autoencoder significantly outperformed the PCA baseline, reducing reconstruction error (MSE) by approximately 40%.

- t-SNE visualization confirmed that the model naturally clustered genres like "Classical" and "Acoustic" together based on sonic properties rather than labels.
- The system successfully recommended "low-popularity" tracks that shared the same audio profile as mainstream hits.

4. Recommendations for Stakeholders * **Recommendation 1:** Integrate the Latent Space model into the "Discover Weekly" algorithm to boost the visibility of independent artists.

- **Recommendation 2:** Use the genre clustering data to automate the tagging of unclassified tracks uploaded by users.
- **Recommendation 3:** Develop a "Sonic Slider" feature for users to adjust their discovery settings based on specific audio features (e.g., increasing "Energy" for workout playlists).