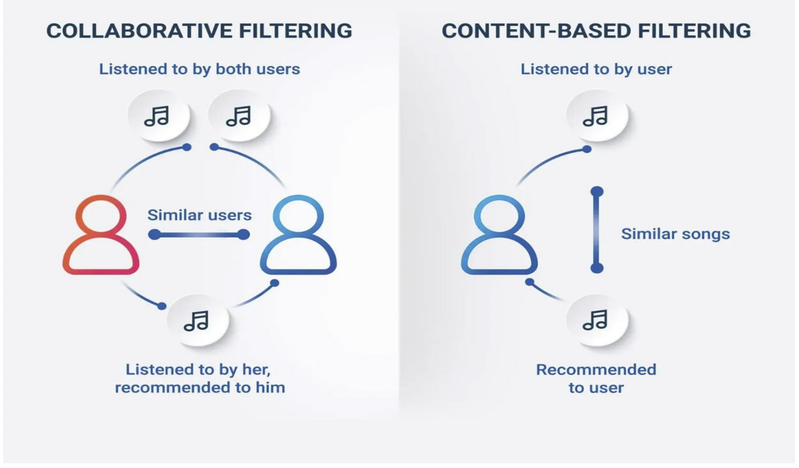
Machine Learning for Recommendation System

Susmita Poudel

Objective

The goal of this project is to develop a robust music recommender system leveraging two distinct machine learning approaches - content-based filtering and collaborative filtering. The system enhances the user experience by delivering personalized song recommendations by analyzing track metadata and user listening patterns.

Collaborative vs. Content



https://images.app.goo.gl/QKuxbESpWuJdimJ67

Dataset

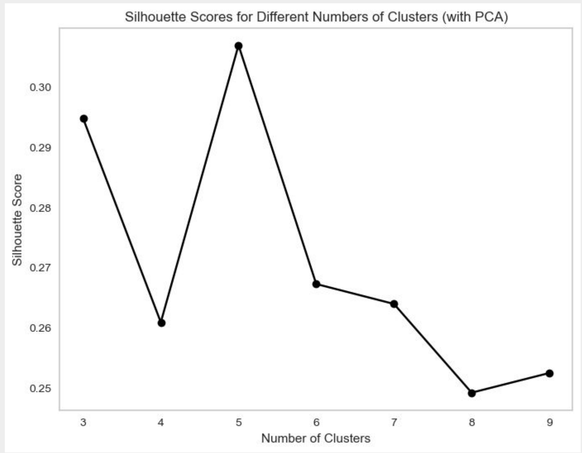
- Content and collaborative approaches were trained on separate datasets sourced from kaggle.
- The dataset for content-based filtering had metadata for tracks. **Used 9 features**
- The dataset for collaborative filtering had user and playlist. Interaction was added for user-song. **13147 users *11053 songs interactions**

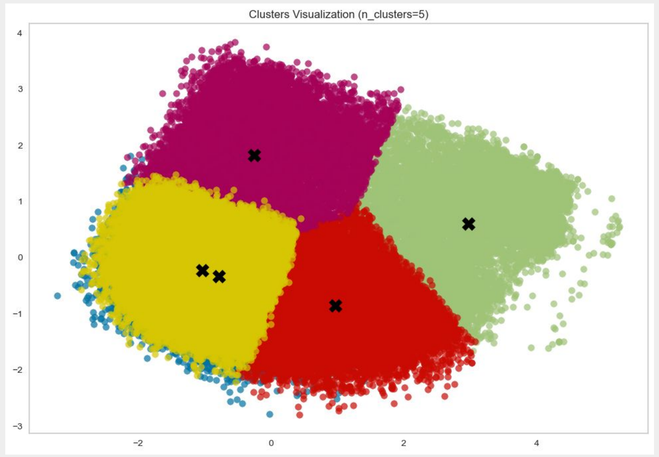
Acknowledgements

I would like to thank Dr.Heath for his guidance, and my classmates in DSC 500 for their feedback throughout the duration of the project.

Model & Results

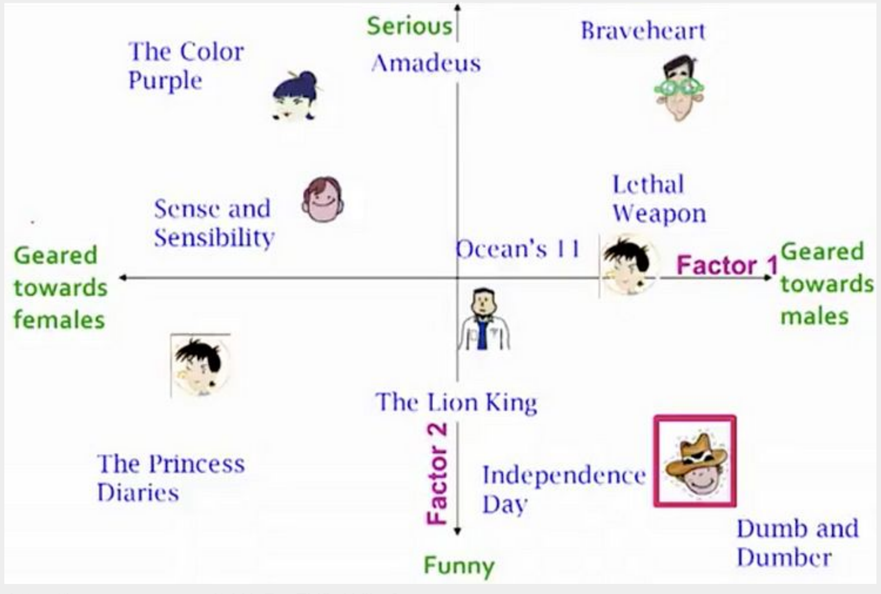
1) Content-based approach





- Silhouette scores were used to determine the optimal number of clusters for kMeans algorithm. The highest Silhouette score achieved was **0.307**, indicating moderate cluster separation.
- Content based filtering groups songs using **kMeans clustering** with **five distinct clusters** based on features such as popularity, genre, etc.
- This approach lacks an established metric for prediction accuracy.

2) Collaborative approach



https://images.app.goo.gl/yYJa7mtEVSqkVjLo6

- Utilizes Singular Value Decomposition **(SVD)** for a latent factor model.
- Optimization approach with SVD
- Captures user-item similarities for personalized recommendations.
- Evaluated using Root Mean Square Error: **RMSE = 0.05**