Re-Pollinating: Making Room for Random

An Argument Against Best Fit in Music Recommendation

UNIVERSITY OF MICHIGAN

Elizabeth Bradford, **elizbr**
SI671 Data Mining

MOTIVATION

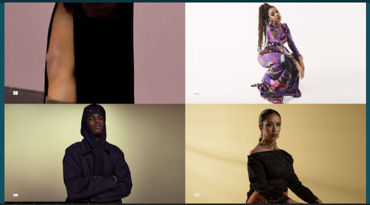
Music streaming services like Spotify overfit recommendations and do not allow for diversity in recommendation genre or content.

how much variance in a recommendation system is appropriate?

Accepting the Spotify Track Radio playlist as the baseline, overfit model, I created competing models to balance similarity encountered in a content-based recommendation system with an added, more disruptive element.

Variance and less predictive content is something that a counterculture of consumers seek out.

Best in Class Player: COLORS X STUDIOS
"All COLORS, no genres."



Pulling a subset of the Million Song Dataset's last.fm data, the following track info was utilized:

- Track title, id, artist and date published
- Track tags and confidence scores (0-100)
- Similar Tracks with confidence scores

METHODS

Based on a selected "seed track" the first 10 tracks returned by the each method is compared to the baseline (Spotify).

SEED TRACK:
"ROUND AND ROUND"

BASELINE

The first 10 tracks of "Track Radio" via Spotify.

TOPIC MODELING

10 sampled tracks for LSI similarity scores based on Track Tags using NLP.

CLUSTERING

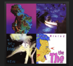
Creating a Network from a Track's listed similars and using network evaluations to create KNN clusters.

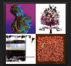
MANUAL

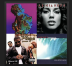
Using listed similar tracks and tags to scrape dataset for other tracks with similar notes via Re Library.

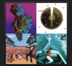
Exploration & Tuning	# of Topics Similarity Score Threshold, *e.g. > .95*	Network Metric Scores, *e.g. Centrality measures*	Similar Track & Tag confidence score thresholds

LISTEN FOR YOURSELF



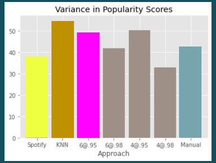


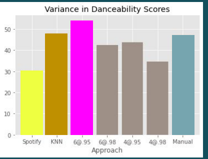


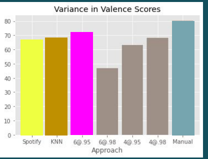


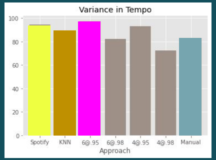
EVALUATION

Qualitatively, each method's tracklist is listened to and scored for genre variety and content variety. The metrics scored for variance are the following: popularity, danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, valence, liveness and tempo.

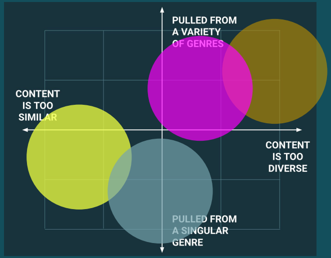








Qualitatively, each method's tracklist is listened to and scored for genre & content variety.



BASELINE
TOPIC MODEL
CLUSTERING
MANUAL

OUTCOMES

Most appropriate variety for ten track recommendations is the Topic Modeling using Latent Semantic Indexing. This method proved to balance track diversity in genre and content very well while the clustering model sometimes created too much diversity.

All in all, this is a question for every individual user. While some may be content with Spotify's recommender, others may be looking for more cross pollination and creativity in their lives. Give the sample track lists a shot by following the QR codes to the left. Which amount of variance is right for you?