Spotify Genre Analysis

By: Mitchell Ren

Overview

Business Context

Data Exploration

Modeling

Results Evaluation

Key Takeaways & Next Steps

Business Context

What separates Spotify from its competitors is the platform's **data-driven consumer experience**.

Opportunity

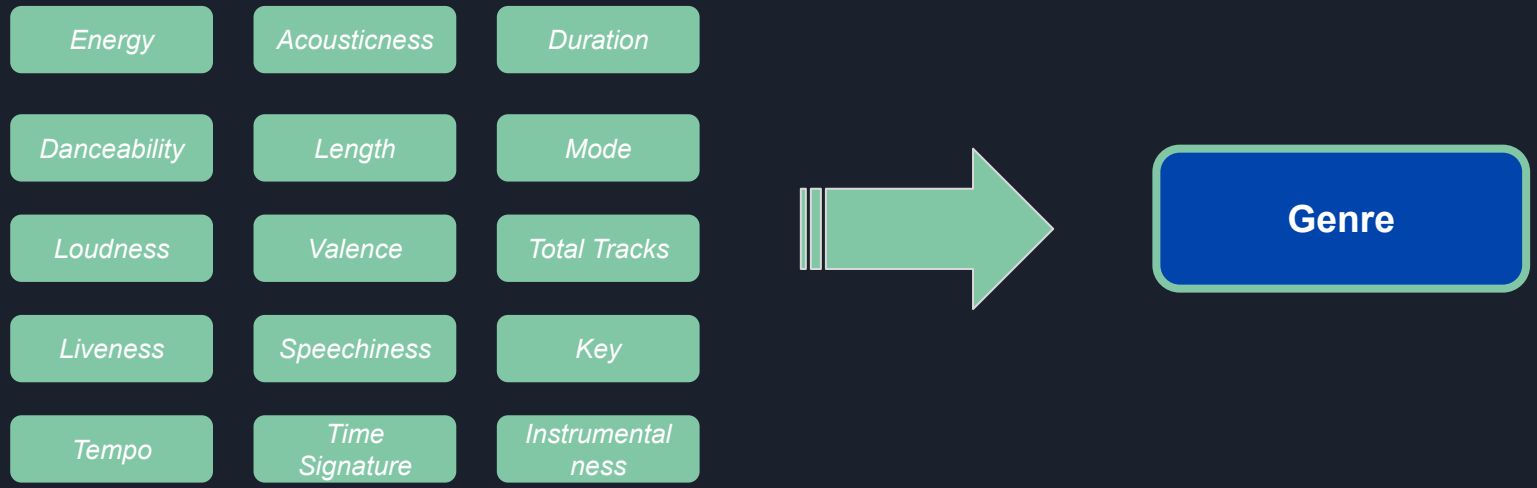
Spotify can further enhance its customer experience by improving upon its **song recommendations** for genre-based playlists and radio

Approach

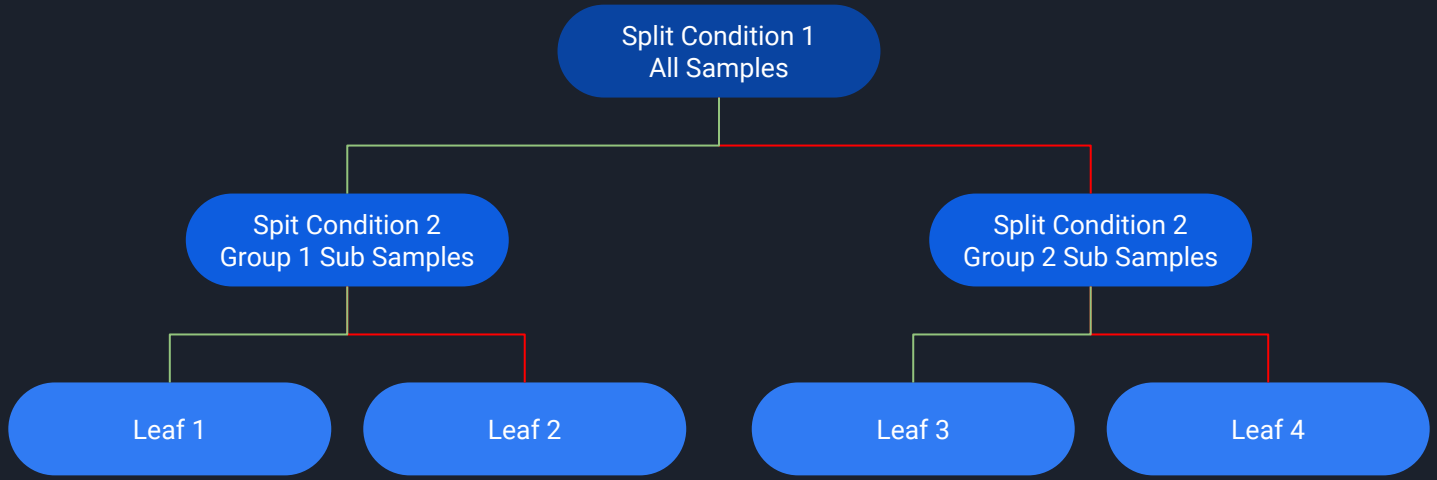
In order to capitalize on this opportunity I built classifier models to group Spotify songs into genres

Data Exploration

The dataset I used for this analysis comes directly from **Spotify's API** and includes arious categorical and numerical song features.

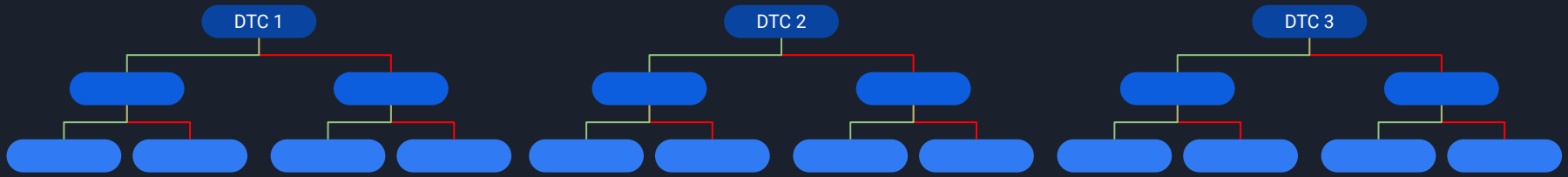


Modeling



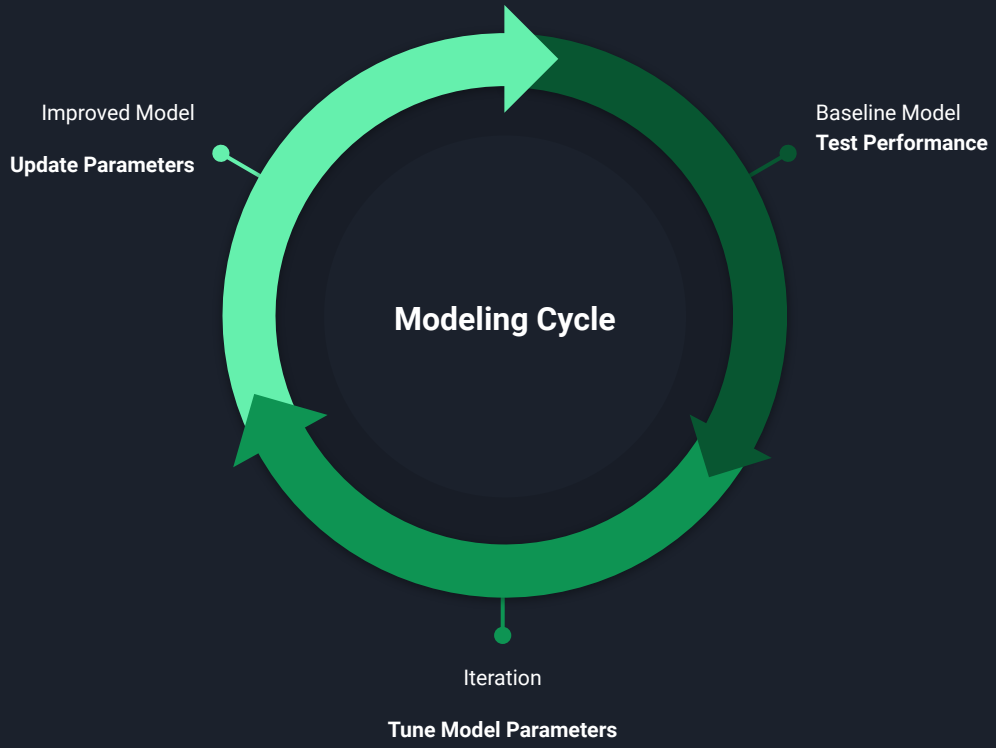
Decision Tree Classifier

Modeling (Cont.)

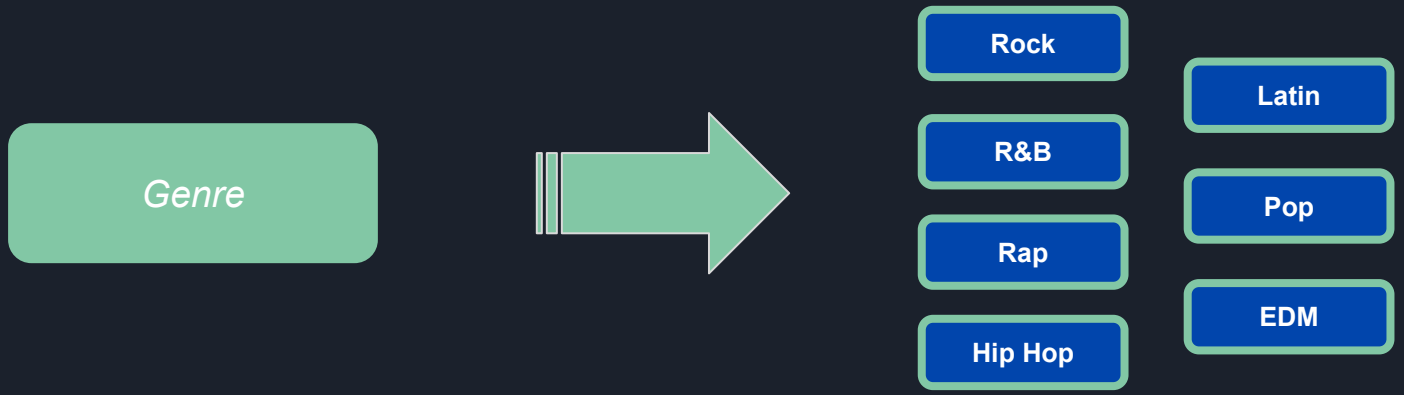


Random Forest Classifier

Modeling (Cont.)



Modeling (Cont.)



Modeling (Cont.)

Results & Evaluation (Multi-Class)

Model	Accuracy
Baseline DTC	Train: 0.90 | Test: 0.42
Tuned DTC	Train: 0.58 | Test: 0.49
Baseline RFC	Train: **0.91** | Test: **0.57**
Tuned RFC	Train: **0.74** | Test: **0.59**

Results & Evaluation (Binary)

Model	Accuracy
Rock vs. Not Rock - RFC	Train: 0.97 | Test: 0.91
EDM vs. Not EDM - RFC	Train: 0.97 | Test: 0.93
Pop vs. Not Pop - RFC	Train: 0.93 | Test: 0.82
Latin vs. Not Latin - RFC	Train: 0.94 | Test: 0.86
Hip Hop vs. Not Hip Hop - RFC	Train: 0.93 | Test: 0.85
Rap vs. Not Rap - RFC	Train: 0.93 | Test: 0.83
R&B vs. Not R&B - RFC	Train: 0.96 | Test: 0.89

Results & Evaluation (Binary)

Model	F-1 Score
Rock vs. Not Rock - RFC	(True) Test: **0.74**
EDM vs. Not EDM - RFC	(True) Test: **0.79**
Pop vs. Not Pop - RFC	(True) Test: **0.46**
Latin vs. Not Latin - RFC	(True) Test: **0.58**
Hip Hop vs. Not Hip Hop - RFC	(True) Test: **0.48**
Rap vs. Not Rap - RFC	(True) Test: **0.49**
R&B vs. Not R&B - RFC	(True) Test: **0.33**

Key Takeaways

93% Accuracy

79% F-1 Score

EDM vs. Not EDM

Most important feature to distinguishing EDM from other genres was Total Tracks

91% Accuracy

74% F-1 Score

Rock vs. Not Rock

Most important feature to distinguishing Rock from other genres was Danceability

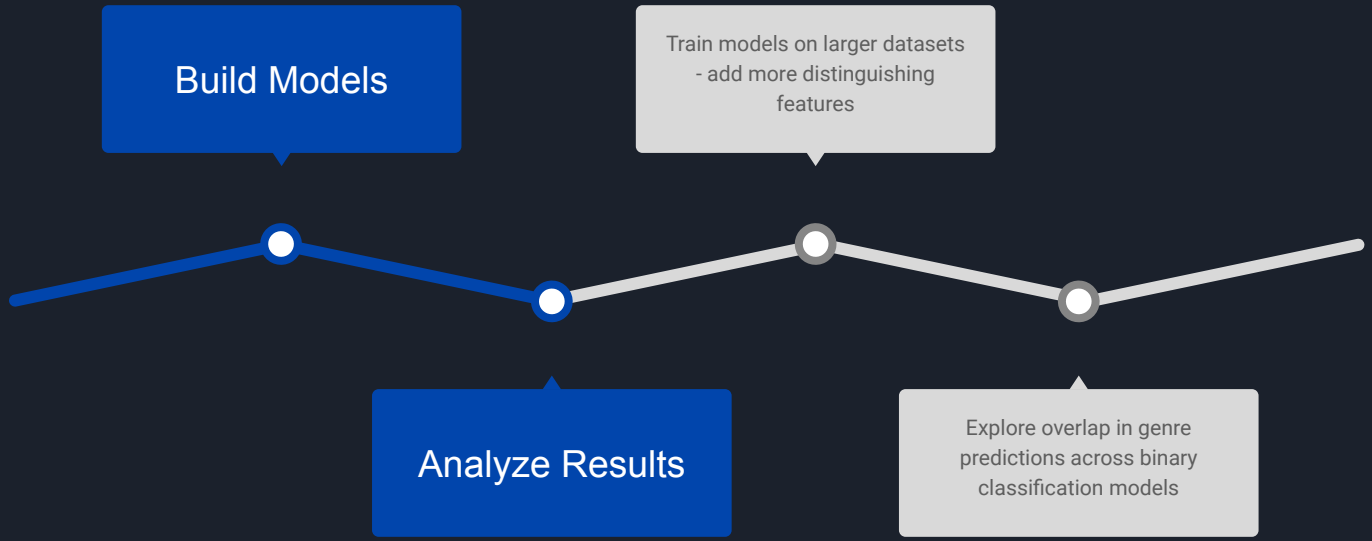
86% Accuracy

58% F-1 Score

Latin vs. Not Latin

Most important feature to distinguishing Rock from other genres was Acousticness

Next Steps


Build Models
Train models on larger datasets - add more distinguishing features
Analyze Results
Explore overlap in genre predictions across binary classification models

Thank You

Contact: mitchell.rensei@gmail.com