Spotify Skip Predictions

Data Analysis and Skip Predictor Model

Keep users on the app by limiting skips

About Me

Rachael McCue

Data Scientist

B.S. International Business
Customer Support
Non-Profits
F&B

Genres of choice:

Techno, D&B, Punk,
Psych/Indie Rock, Rap

Agenda

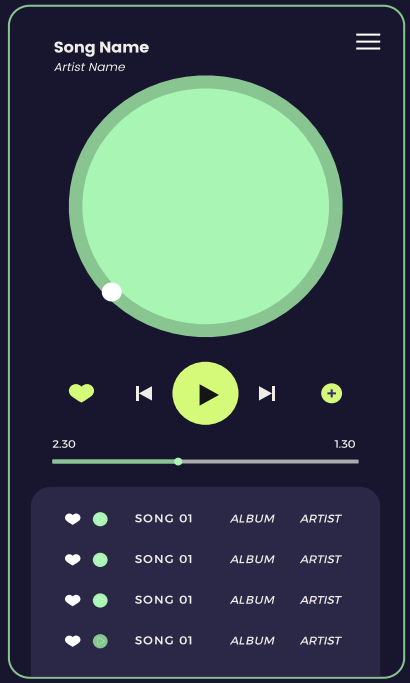
01 Business Understanding

02 Data Overview & Analysis

03 Introduce models

04 Recommendations and future insights

More than half of songs in a listening session are skipped in the first 30 seconds.



PLAYLIST

01 *Business Understanding*

02 ***Data Analysis***

03 *Model*

04 *Looking Ahead*

Data Overview

PLAYLIST

01 *Business Understanding*

02 ***Data Analysis***

03 *Model*

04 *Looking Ahead*

Data Overview

WSDM, Spotify, AiCrowd - 2018

10k listening sessions, 10-20 songs each

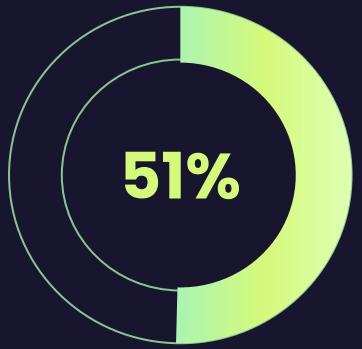
User behavior and environment

Track information

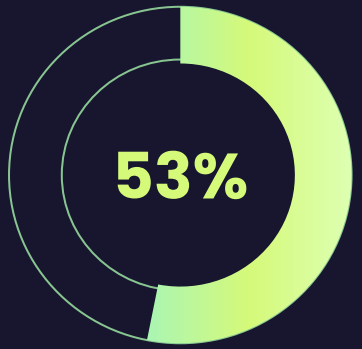
Limitations

Premium and Free users skip close to the same amount

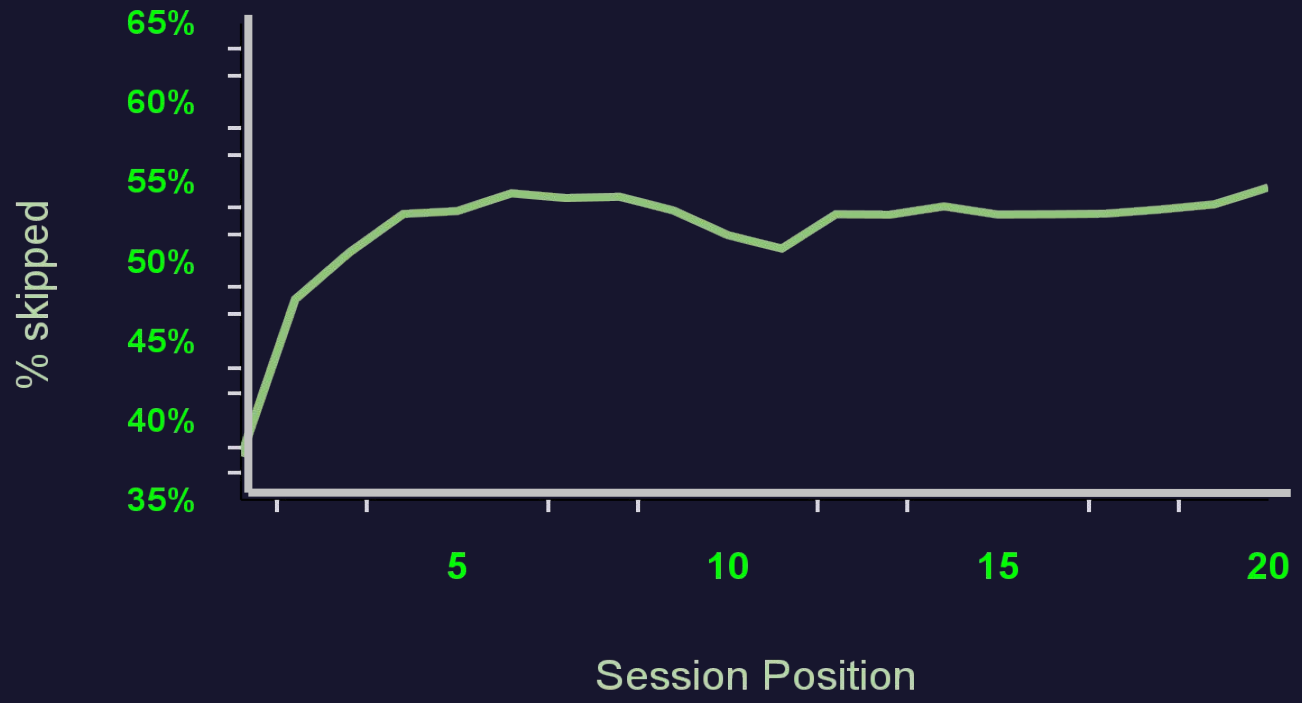
Premium



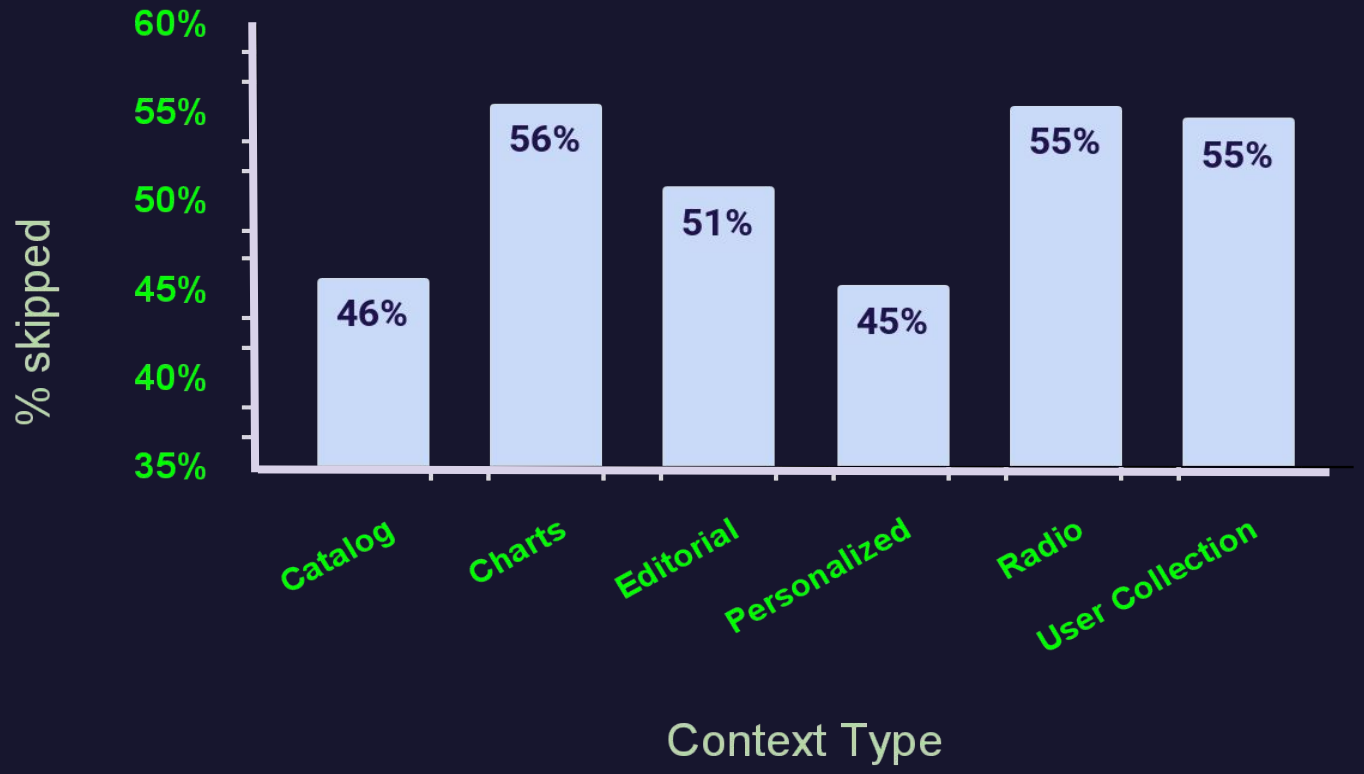
Free



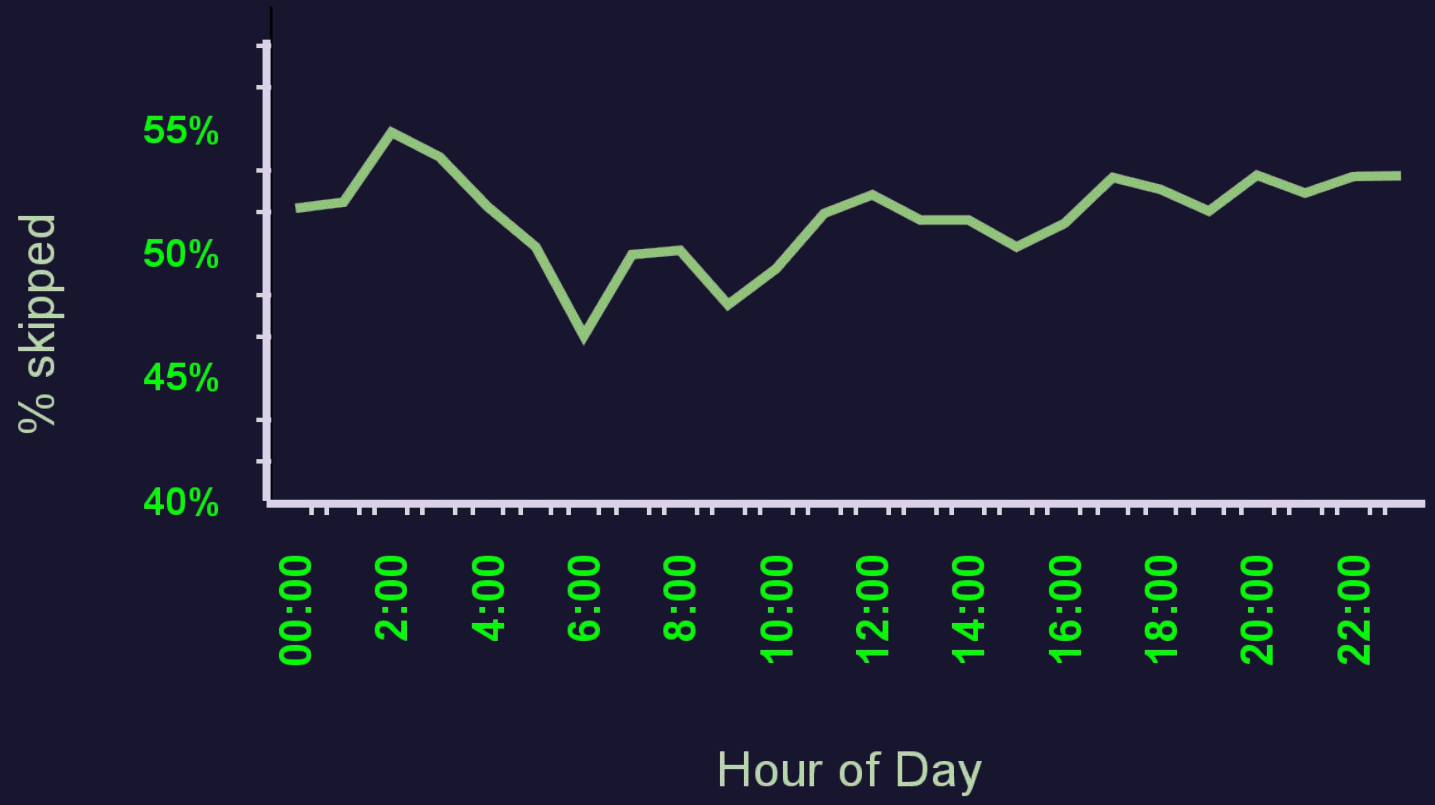
Skips increase as session lengthens



Radio and Chart playlists are skipped the most



Fewer skips in the morning



PLAYLIST

01 *Business Understanding*

02 *Data Analysis*

03 ***Model***

04 *Looking Ahead*

Skip Predictor Model

PLAYLIST

01 Business Understanding

02 Data Analysis

03 ***Model***

04 Looking Ahead

01

Binary Classifier

PLAYLIST

01 *Business Understanding*

02 *Data Analysis*

03 ***Model***

04 *Looking Ahead*

Binary Classifier

Accuracy Score - **56%**
F1 Score - **61%**

Did well on unseen data

General skipping behavior, no musical features, ignores sequencing

PLAYLIST

01 *Business Understanding*

02 *Data Analysis*

03 ***Model***

04 *Looking Ahead*

02

Neural Net

Neural net explained

PLAYLIST

01 *Business Understanding*

02 *Data Analysis*

03 ***Model***

04 *Looking Ahead*

Neural Net

Accuracy Score - **70%**
F1 Score - **61%**

Less robust than Binary Classifier on unseen data

Uses mostly all features

Model is designed for sequential data

PLAYLIST

01 *Business Understanding*

02 *Data Analysis*

03 *Model*

04 ***Looking Ahead***

Breakdown

AM sessions and Morning hours have higher skips

Skip behavior is robust to type of account

Deliver more appropriate content to the user based on recent activity and their mood

PLAYLIST

01 *Business Understanding*

02 *Data Analysis*

03 *Model*

04 ***Looking Ahead***

Next Steps...

Multiclass Classifier

Genre & Podcasts

Thank you

Any questions?

https://github.com/rachaelmmccue/Spotify-Sequential-Skip-Prediction