

Spotify User Behavior Dataset

Priscilla Wong

Problem Area

Research Question: Which user behaviors most strongly predict willingness to subscribe to Spotify Premium?

I chose this topic because I use Spotify on a daily basis, and it would be interesting to do a study on the correlation of user behavior and Spotify premium.

Subscription-based platforms like Spotify rely heavily on converting free users into paid subscribers. In general, this can be used to inform targeted marketing strategies and improve features on the app.

The Data

This is a free dataset that came from Kaggle.

It is a survey-based Spotify user behavior dataset. It is a collection of anonymized data that offers valuable insights into the behavior patterns and preferences of Spotify users.

Preliminary Assessment:

- 520 rows of data
- Ensured that there are both free and premium subscribers so that the data is not too skewed

Important Inputs:

- music_recc_rating: User satisfaction with recommendations (1–5 scale)
- spotify_usage_period: Length of platform loyalty (Years)
- music_lis_frequency: Primary engagement context (Time of day)
- Age: Age group of respondent

Important Output:

- Premium_sub_willingness: Yes/No to get or continue subscription



Analysis Plan

Exploratory Data Analysis

- Central tendency and spread
- Histograms and box plots to visualize distributions

Hypothesis Testing

- H0: There is no difference in the mean rating between users who are willing to subscribe and those who are not.
- H1: Users who are willing to subscribe have a statistically significant higher mean rating than those who are not.

Predictive Analysis

- k-Nearest Neighbors to classify users
- Decision trees to identify correlations

Pattern Discovery

- K-Means clustering to group users based on their numerical ratings and usage frequency

Successful Outcome

- Clear ranking of feature importance between age, rating, usage, and frequency



Thanks!

Any questions/comments/suggestions?