BrainStation Capstone Project

Is the song "hot enough" for Billboard Hot100?

Predicting the probability of a song reaching the Billboard Hot100 chart.

GLOSSARY

Table schema for file "capstone_dataset_clean.csv"

Tables in secondary files like "capstone_dataset_artists.csv" or "capstone_dataset_genre.csv" are derived from this table below and use the same columns.

Column Name	Measurement	Description
billboard	$0\ or\ 1$	**This is our target variable.** Contains a binary output for songs that were on the Billboard Hot100 (1) and songs that were not (0)
billboard_year	integer	Year that song was on Billboard Hot100 in the format YYYY (if n/a, value is zero). This project covers the Billboard years of 2008-2021, as well as the first 4 months of 2022.
billboard_month	integer	Month that song was on Billboard Hot100 in the format MM (if n/a, value is zero)
peak_position	$1 \leq\ x\ \leq 100$	Highest position that song was ranked; 1 = highest and 100 = lowest (if n/a, value is zero)
weeks_on_chart	integer	Number of weeks that song was on Billboard Hot100 (if n/a, value is zero)
artist	string	Main artist of song
feat_artist	string	Featured artist of song
artist_uri	string	Unique artist ID in Spotify (if n/a, value is '-')
artist_popularity	$0 \leq\ x\ \leq 100$	Artist popularity score in Spotify
artist_followers	Integer	Number of followers that artist has in Spotify
song	string	Title of song
track_uri	string	Unique track ID in Spotify (if n/a, value is '-')
track_popularity	$0 \leq\ x\ \leq 100$	Track popularity score in Spotify

BrainStation Capstone Project

Is the song "hot enough" for Billboard Hot100?

Predicting the probability of a song reaching the Billboard Hot100 chart.

album_name	string	Name of album
release_year	integer	Year that song was released
release_month	integer	Month that song was released
duration_ms	integer	Duration of song in milliseconds (if n/a, value is zero)
danceability	$0 \leq x \leq 1$	How suitable a track is for dancing based on a combination of tempo, rhythm, stability, beat strength, and overall regularity
energy	$0 \leq x \leq 1$	Measurement of intensity and activity. Typically tracks that feel and sound fast, loud, and noisy.
key	-1 *to* 12	The key the track is in. If no key is detected, value is -1. Key starts from $0 = C$. Every $+1$ is a half-step up (e.g. $0 = C \rightarrow 1 = C\#$)
loudness	$-60 \leq x \leq 0$	Overall loudness of a track in decibels (dB)
mode	0 *or* 1	Modality of track or which scale the music is in (1 = major, 0 = minor). Typically, major scales evoke positive emotions, while minor scales sound more mellow/ominous.
speechiness	$0 \leq x \leq 1$	Presence of words spoken in track. $x > 0.66$ usually indicates a podcast, audio book, or similar. $x < 0.33$ usually indicates music or non-speech tracks. And anything in between contains both types in either sections or layers, including cases like rap music.
acousticness	$0 \leq x \leq 1$	A confidence measure of whether the track is acoustic
instrumentalness	$0 \leq x \leq 1$	Whether a track contains vocals. "Ooh" and "ahh" (vocalise) are treated as such in this case.
liveness	$0 \leq x \leq 1$	Detects presence of an audience in the recording
valence	$0 \leq x \leq 1$	Music positiveness where 1 is happy, cheerful, and euphoric and 0 is sad, depressed, and angry.
tempo	integer	Tempo of a track in beats per minute (BPM)
time_signature	$3 \leq x \leq 7$	How many beats are in each measure. In music composition, a song is divided into measures and each measure has a number of beats. Most songs have 4 beats per

		measure. 3 beats per measure is written as 3/4 (Think: during a waltz, dancers often count “1-2-3, 1-2-3”.)
major_label	0 *or* 1	New feature derived from release year. Artist gained recognition or released first song before 2016 (1). After that, artist is considered "new" or "emerging" (0)
genre_pop	0 *or* 1	Clustered Spotify's genres by keywords: pop, alt-z, idol, band, boy group, girl group, etc.
genre_screen	0 *or* 1	Clustered Spotify's genres by keywords: movie, hollywood, disney, broadway, soundtrack, Glee Club, etc.
genre_hiphop	0 *or* 1	Clustered Spotify's genres by keywords: hip hop, rap, grime, crunk, swag, etc.
genre_punk	0 *or* 1	Clustered Spotify's genres by keywords: punk, garage, permanent wave, industrial, new wave, etc.
genre_blues	0 *or* 1	Clustered Spotify's genres by keywords: blues, funk, reggae, swing, groove, bluegrass, etc.
genre_country	0 *or* 1	Clustered Spotify's genres by keywords: country, folk, cowboy, oldtime, etc.
genre_rnb	0 *or* 1	Clustered Spotify's genres by keywords: r&b, soul, afro-, motown, boogaloo
genre_rock	0 *or* 1	Clustered Spotify's genres by keywords: rock, metal, deathcore, hard, scream, grunge, progressive, etc.
genre_latin	0 *or* 1	Clustered Spotify's genres by keywords: latin, reggaeton, mexi-, espanol, dominicana, brasil, etc.
genre_dance	0 *or* 1	Clustered Spotify's genres by keywords: dance, club, disco, party, tropical, etc.
genre_indie	0 *or* 1	Clustered Spotify's genres by keywords: indie, alternative, singersongwriter, acoustic, solo wave, etc.
genre_elect	0 *or* 1	Clustered Spotify's genres by keywords: elect-, house, edm, tech-, -tronica, etc.
genre_gospel	0 *or* 1	Clustered Spotify's genres by keywords: gospel, worship, praise, ccm, christian
genre_chill	0 *or* 1	Clustered Spotify's genres by keywords: jazz, lounge, chill-, lofi, easy listening, middle earth, etc.