My Final Project
Classifying popular songs from Spotify using parametric and non-parametric models

By Clarissa Ache-Cabello

PROJECT MOTIVATION & OBJECTIVES

A music genre is a collection of patterns. Songs that share patterns can be grouped in a genre. It is likely that Spotify uses measures of this patterns along with listening patterns to assign songs to genres.

In theory, one can determine the genre of a song only by identifying these patterns, and that is what I am going to try to do in this project.

Project Objectives:

1. Find out what makes a song "belong" to a music genre.
2. Model the characteristics of popular songs to classify them into the genre Spotify has selected for them.

DATA

- Extracted from the Spotify Web API via the R package Spotifyr.
- My table for analysis has a sample of ~33K songs from 6 broad playlist genres:

Pop	**Rap**	**Latin**	**R&B**	**Rock**	**EDM**

- The data had the following 13 song measures:

track_popularity	danceability	energy	key	loudness	speechiness
acousticness	instrumentalness	liveness	valence	tempo	
mode	duration_ms				

EDA

Looked at the distributions of the main effects of predictors for the songs in each genre.

Findings:

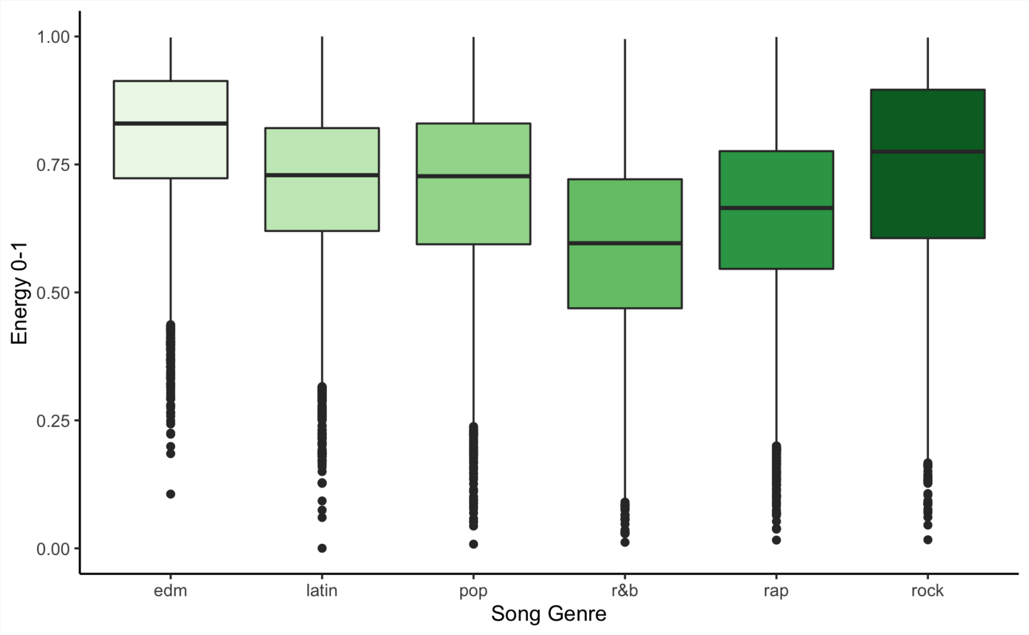
✓ Some predictors seem more useful at classifying than others:

Seemed very useful	Not very good at distinguishing songs
• **Energy** – perceptual measure of of intensity and activity • **Speechiness** – The presence of spoken words in a track. • **Acousticness** – A confidence measure from 0.0 to 1.0 of whether the track is acoustic (no electric instruments) • **Valence -** A measure from 0.0 to 1.0 describing the musical positiveness (happy, cheerful, euphoric) conveyed by a track.	• **Loudness** – The overall loudness of a track in decibels (dB) • **Instrumentalness** – Predicts whether a track contains no vocals. • **Liveness** – Detects the presence of an audience in the recording • **Duration** – Track duration in milliseconds • **Tempo** – Beats per minute (Bpm) • **Popularity** – Spotify's secret popularity score from 0 to 100

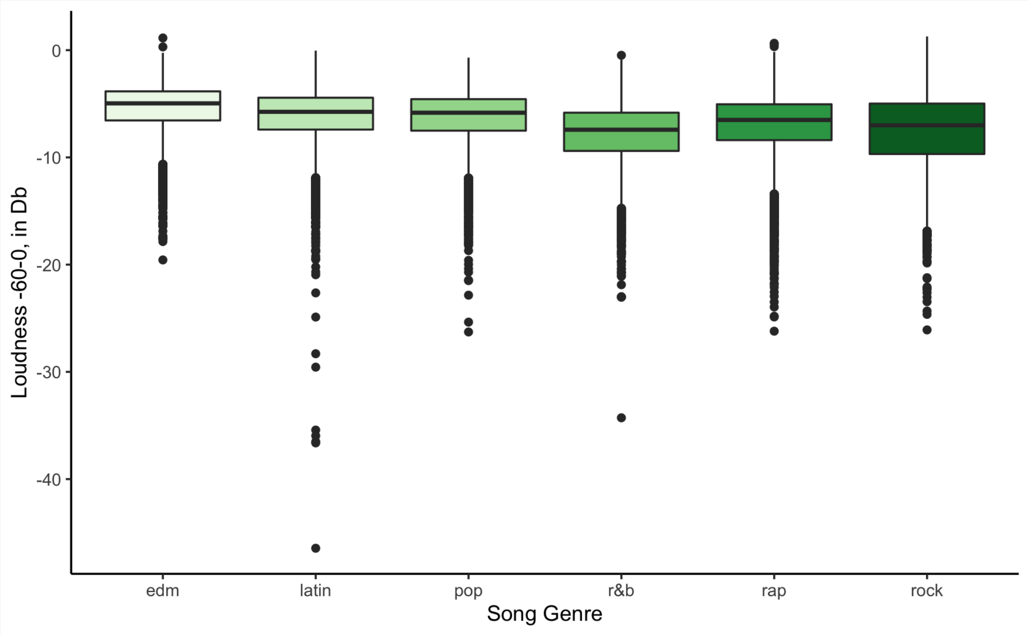
USEFUL VS. NOT SO USEFUL VARAIBLES

(ONLY LOOKING AT MAIN EFFECTS)

Probably a **useful** classifier:
Energy



Probably **not the best** classifier:
Loudness



APPROACH

- Used **2 parametric models** – Multinomial regressions
  - One with only main effects, one with interactions
- Used **3 non-parametric ML models** – Classification trees
  - CART, Bagging, Random Forest

Summary of models

	Parametric		**Non-Parametric**		
	Multinomial regression with main effects only	**Multinomial regression with interactions**	**CART** *(single tree)*	**Bagging**	**Random Forest**
Accuracy	47.0%	48.2%	36.3%	57.0%	56.5%

MULTINOMIAL REGRESSION MODEL (MAIN EFFECTS ONLY)

- A logistic regression for each level (genre) using EDM as the baseline.
- All p-values very significant
- Assumptions are plausible

```
genres1 <- multinom(playlist_genre ~ track_popularity + danceability + energy + key + loudness + speechiness +
acousticness + instrumentalness + liveness + valence + tempo + duration_ms + mode, data=songs)
```

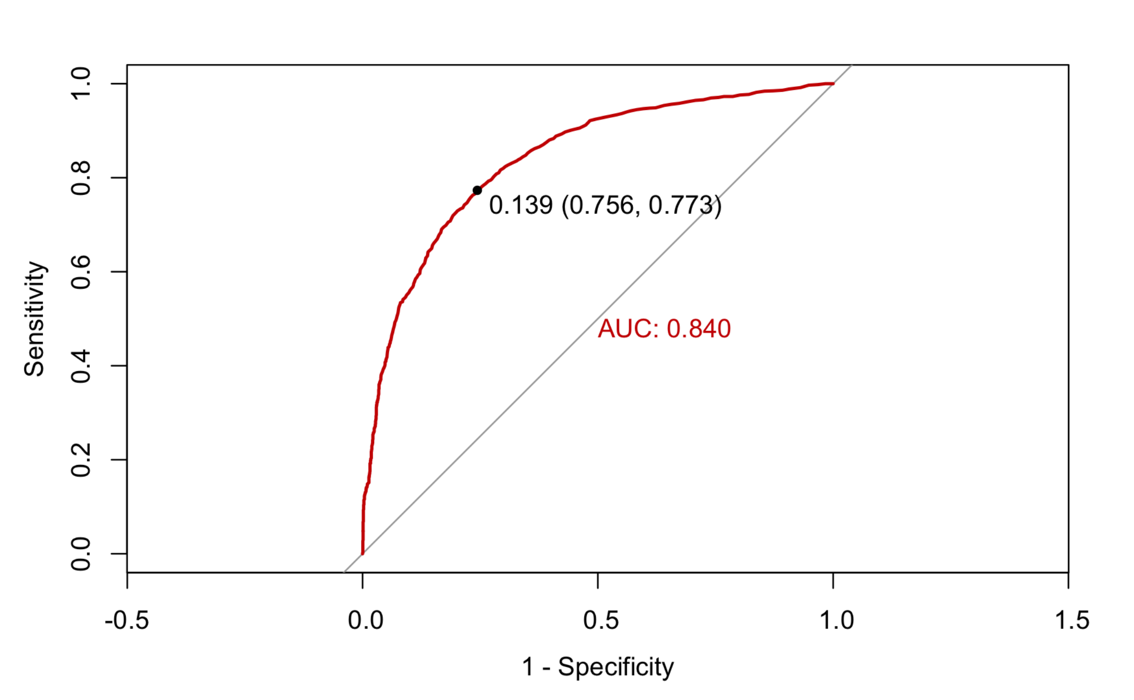
Multinomial Regression Model 1 coefficients (exponentiated)

Genre	(Intercept)	track_ popularity	danceability	energy	key	loudness	speechiness	acousticness	instrumentalness	liveness	valence	tempo	duration_ ms	mode1
latin	1.12	1.01	**1.32**	0.09	1.01	0.92	4.86	10.86	0.09	0.58	42.81	0.99	1.00	1.23
pop	1.24	1.02	0.02	0.04	1.00	0.93	0.06	2.76	0.13	0.36	15.81	0.99	1.00	1.28
r&b	1.15	1.00	0.06	0.00	1.00	0.89	61.11	4.60	0.02	0.53	40.44	0.99	1.00	1.01
rap	1.11	1.01	**2.94**	0.01	1.00	0.92	11060.57	2.88	0.25	0.89	5.48	0.99	1.00	1.07
rock	**0.30**	1.01	0.00	2.26	1.00	0.69	0.00	2.15	0.04	0.43	161.53	0.99	1.00	1.89

THE BAGGING MODEL

- Is a tree-based method that instead of doing one tree classification does a number of them using bootstrapped datasets and averages them out to increase accuracy and reduce variance.

```
genres_bagg <- randomForest(as.factor(playlist_genre) ~ track_popularity + danceability + energy + key + loudness
+ speechiness + acousticness + instrumentalness + liveness + valence + tempo + duration_ms + mode, data=train1,
mtry=4)
```



Genre	Sensitivity (Recall)	Specificity
edm	71%	92%
latin	44%	92%
pop	35%	90%
r&b	48%	91%
rap	68%	90%
rock	77%	93%

Overall Accuracy
57%

CONCLUSIONS (SO FAR), LIMITATIONS, & NEXT STEPS

- I wanted to try the Boosting model but I got a warning saying this:

```
Warning: Setting `distribution = "multinomial"` is ill-advised as it is currently broken.
It exists only for backwards compatibility. Use at your own risk.
```

- The most important takeaways are that:
  - Rap is speechy
  - Latin is most acoustic
  - R&B is always in between
  - EDM is loud
  - Rock is happy and easier to identify
  - What is Pop?