Predicting Hip-Hop Popularity

Nicholas Indorf

Client & Issue

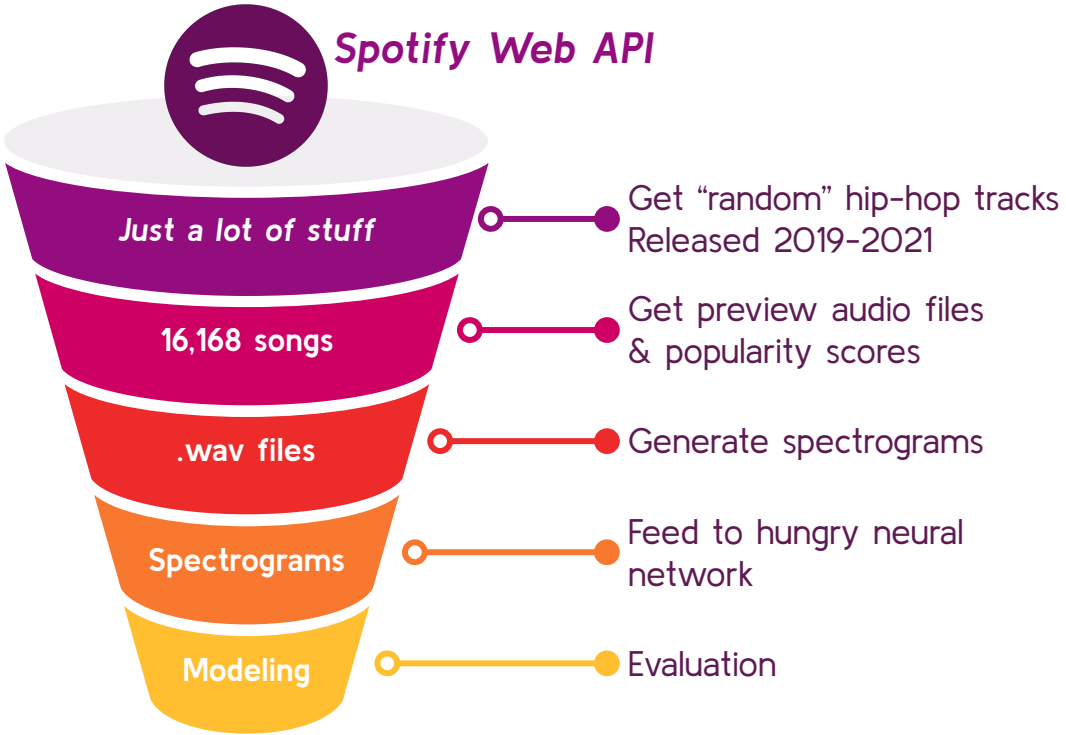
KC Makes Music

- 2018-present, 30 songs
- ~24.5k monthly listeners
- ~10 songs finished, unreleased

Popularity

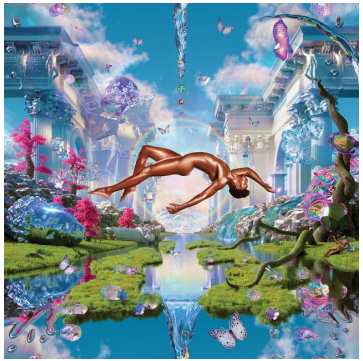
based solely on audio

Data Gathering & Preparation

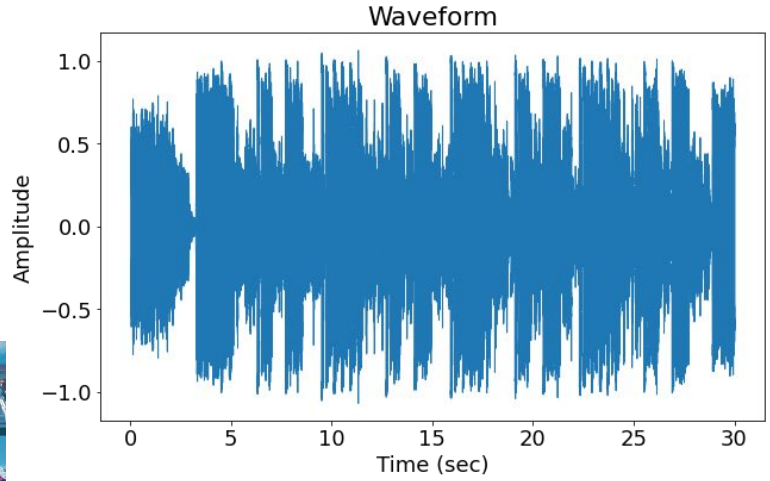


Key Data: Preview audio & Popularity scores

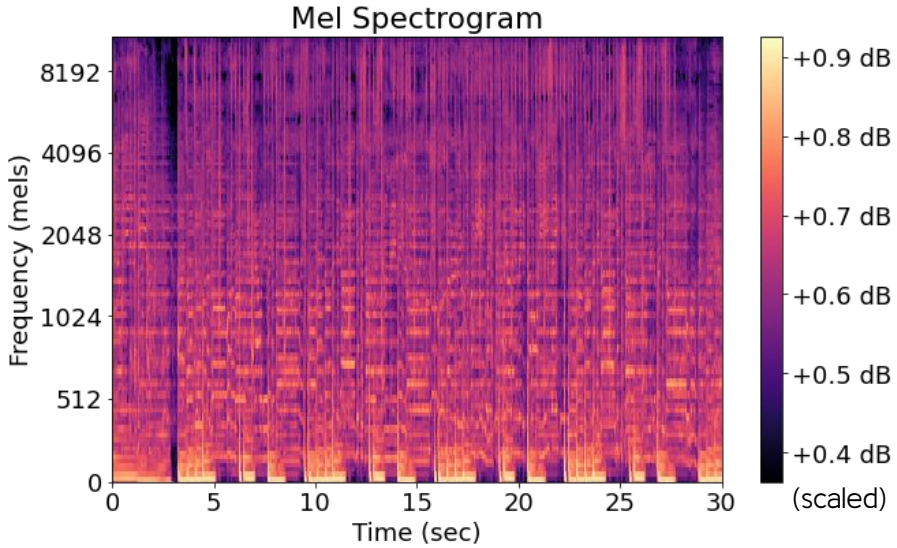
Preview Audio



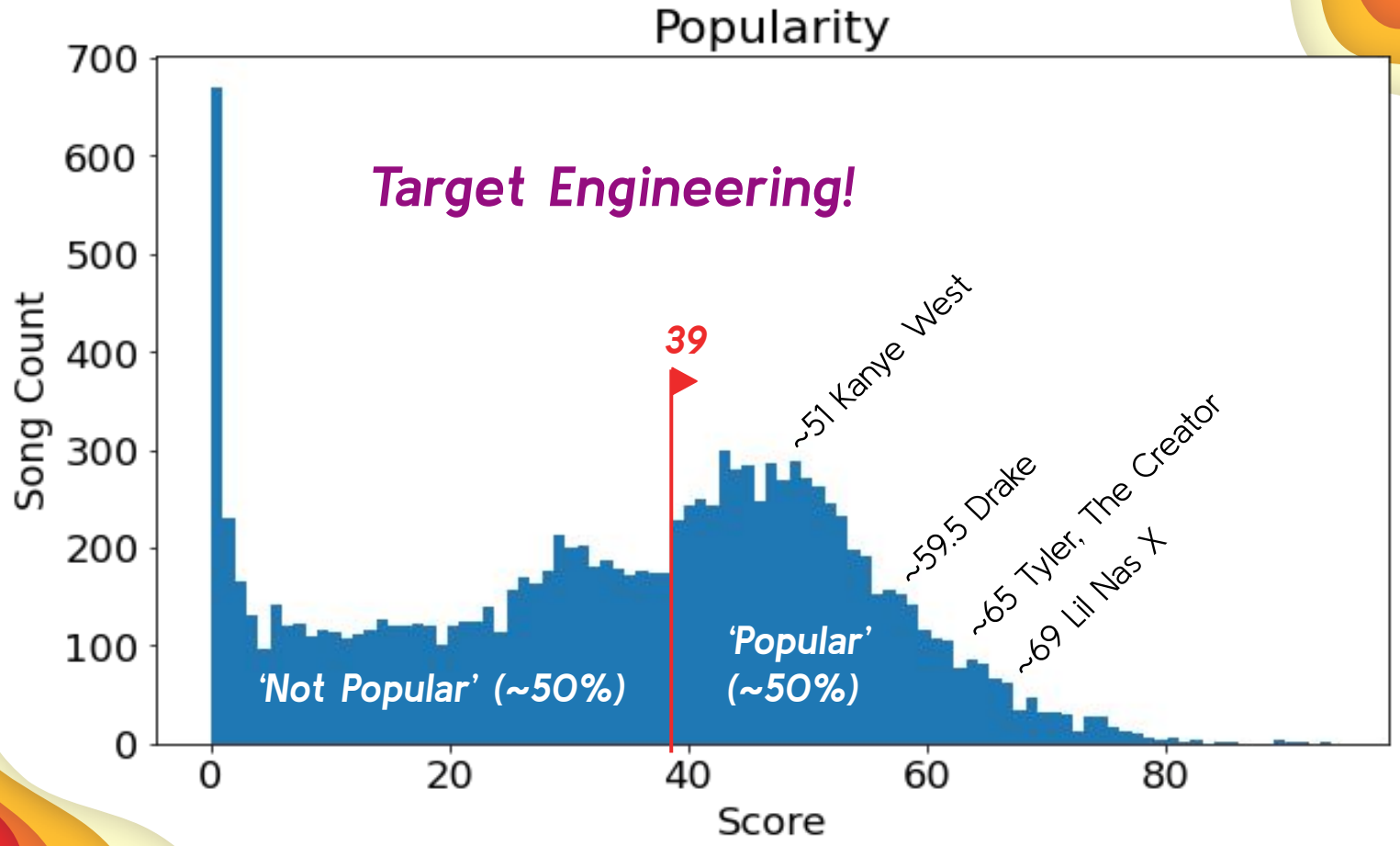
INDUSTRY BABY (feat. Jack Harlow) - Lil Nas X



~30 Seconds



Popularity


Popularity
Target Engineering!
39
~51 Kanye West
~59.5 Drake
~65 Tyler, The Creator
~69 Lil Nas X
'Not Popular' (~50%)
'Popular' (~50%)
Song Count
Score
0
100
200
300
400
500
600
700
0
20
40
60
80

Not Popular

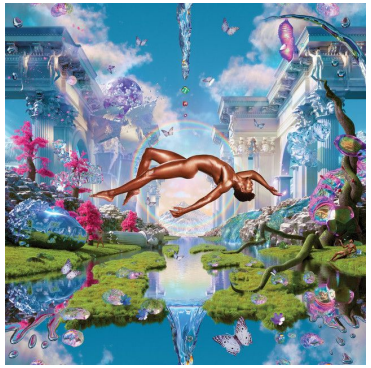
Uknowhatimsayin¿
– Danny Brown

Pop = 35

Look Over Your Shoulder (feat. Kendrick Lamar)
– Busta Rhymes

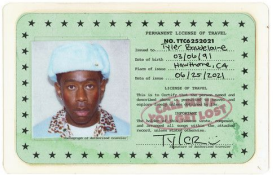
Pop = 33

Popular



INDUSTRY BABY (feat. Jack Harlow)
– *Lil Nas X*

Pop = 90

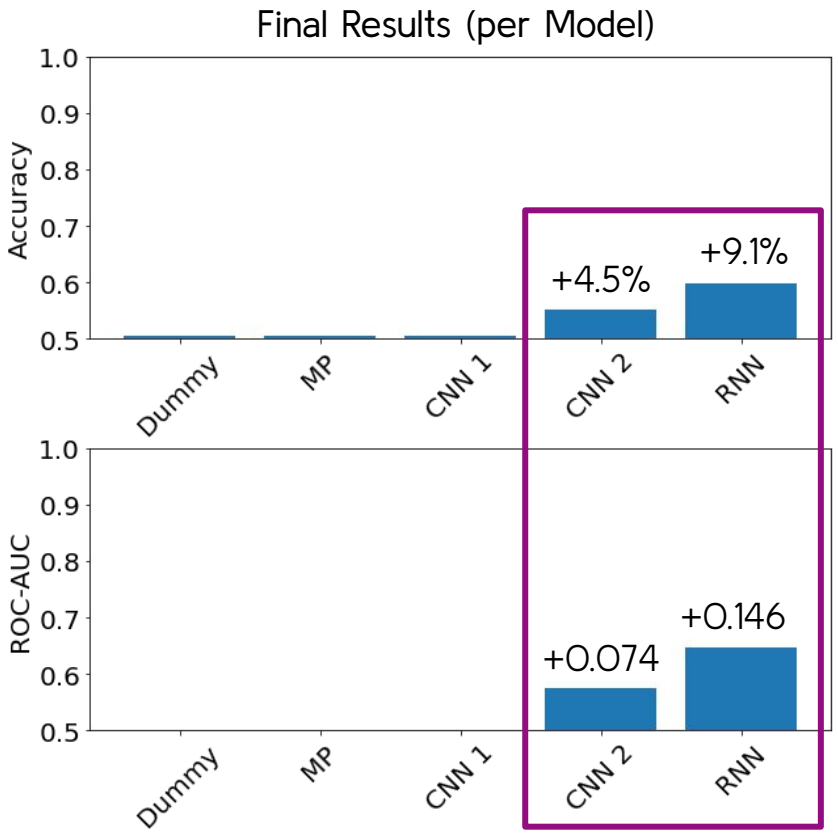


WUSYANAME (feat. Youngboy Never Broke Again & Ty Dolla $ign)
– Tyler, The Creator

Pop = 79

Modeling & Evaluation

- Model: Neural Network
- Metric:
  - Accuracy
  - ROC-AUC

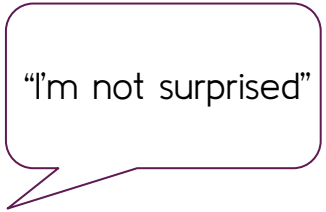


Promising!

Increasing model complexity →

Discussion & Recommendation

- Only audio → ***very limiting***
- Popularity → ***confusing***
- Exposure is huge → TikTok, etc.



Client

Sieve-type model & Continued development

Next Steps

- Incorporate non-audio features
  - E.g. exposure, release timing
- Expand dataset
- Feature importance analysis
  - What did it pick up on?

Thank You

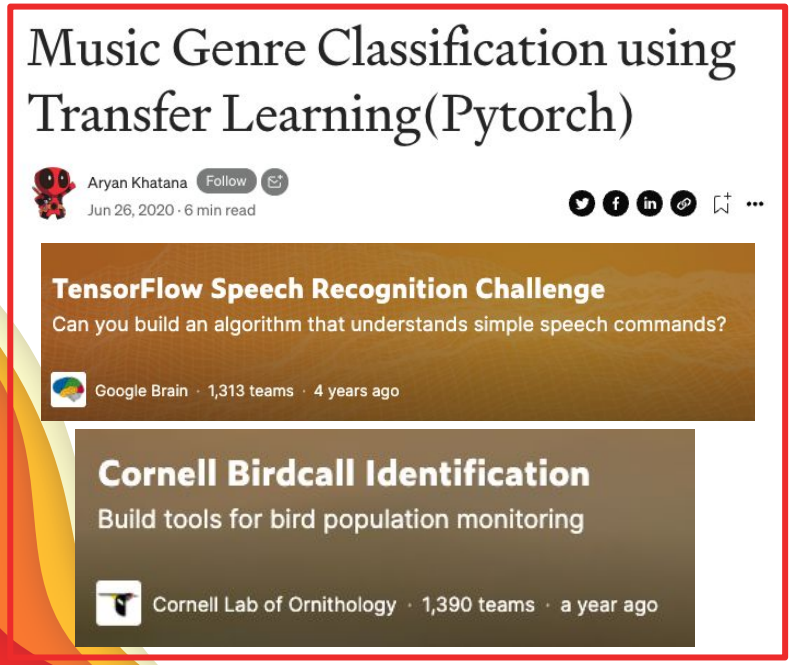

linkedin.com/nicholas-indorf-data-scientist


nicholasindorf@gmail.com

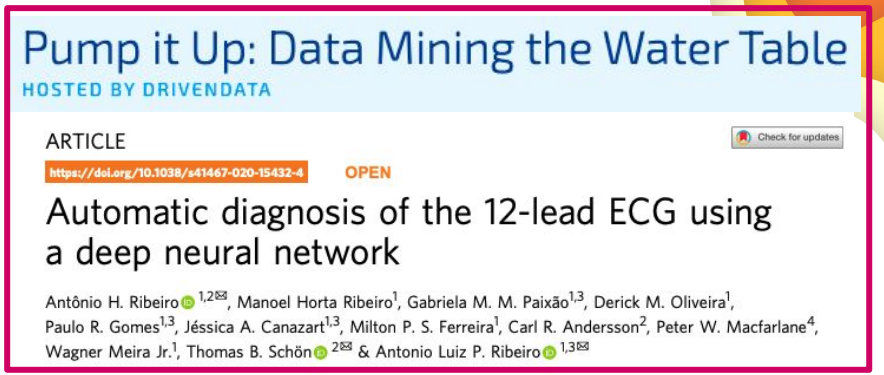

github.com/Nindorph

Apdx. - Context

Audio

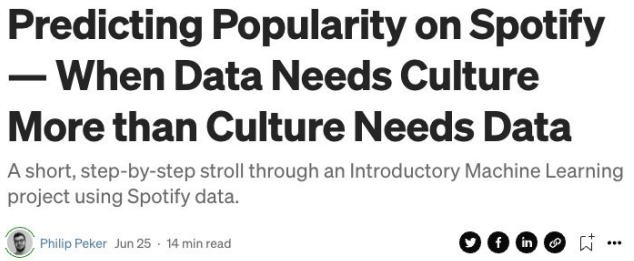


Fine Detail



Similar

→



- Audio + Finely detailed insight

Very Difficult!

(Links to these projects on my GitHub)