

# Stream-Way to Heaven



\$12.2 Billion US Music revenues in 2020



\$10 Billion in streaming revenue



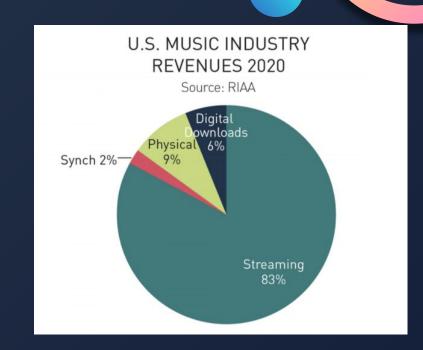
83 million US Spotify users



\$23 Million / Day for major US Record Labels



Top 100 songs worth \$450 million in revenue



# Goal: Create Song Ranking for Hit Identification

## **Process**







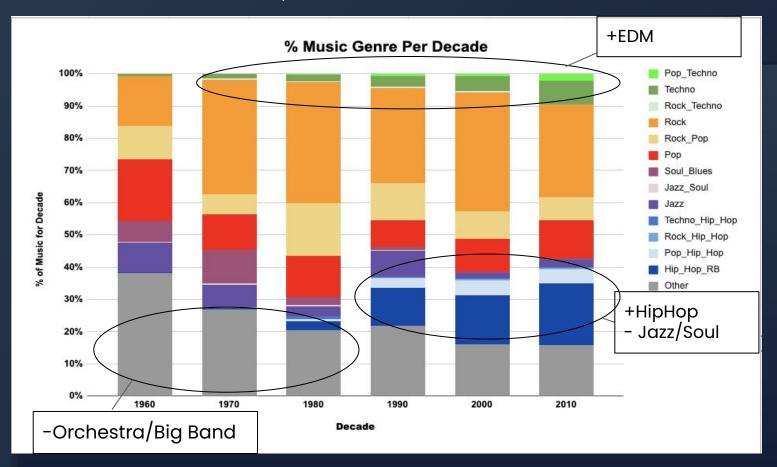




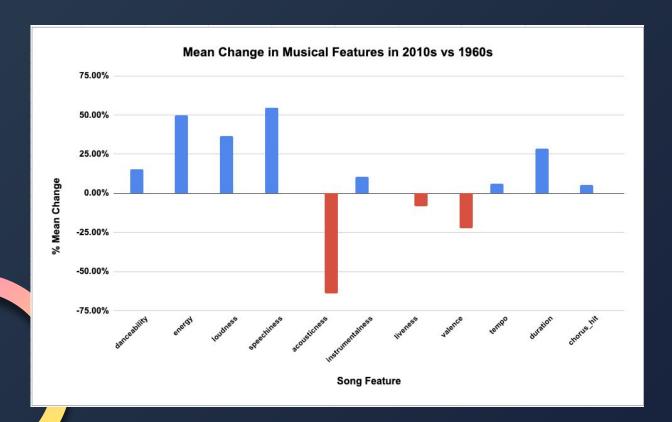
41,000 Spotify Songs with 15 music features (1960 to 2019) Define Songs "Flop" or "Hit" if in Billboard Top 100 Feature Engineer and find best algorithm to maximize AUC

Target: AUC ROC

## Out With The Old, In With the New



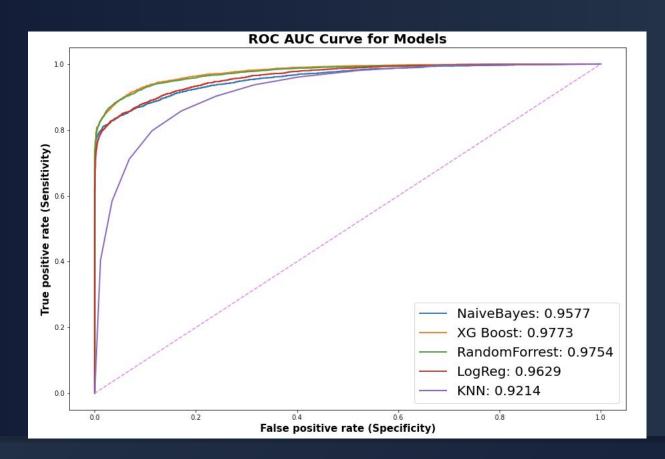
## Times They Are A Changin'



#### **Modern Music**

- +Verbal
- +Energetic
- +Louder
- +Longer
- +Danceable
- -Acoustic
- -Live
- -Happy

## XG Boost: Best of the Rest

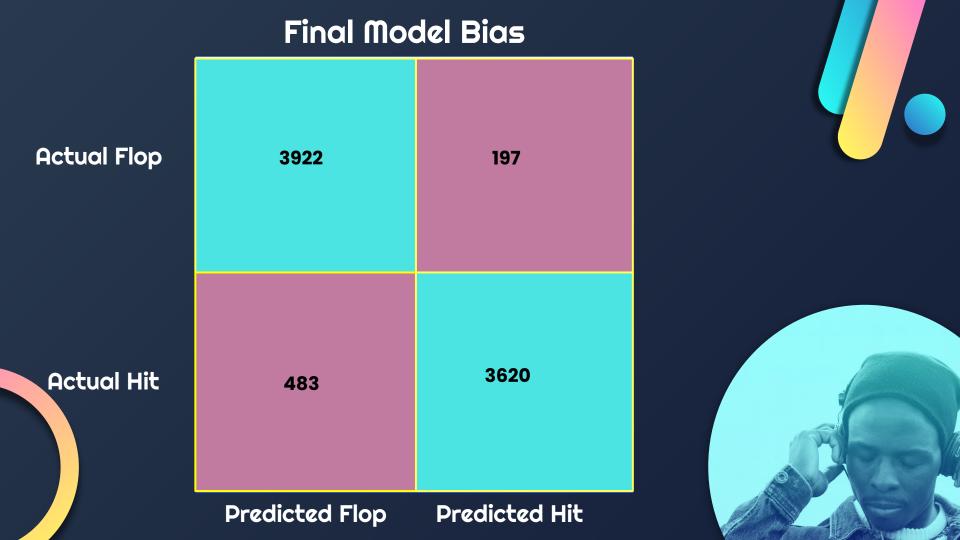


Final AUC Score: .9765

**Hit Accuracy: .91** 

#### Feature Importance

- Previous Hits
- Genre
- Instrumentalness



"But how will this model translate outside of this dataset?"

-Someone Famous



# **Predicting the Billboard 100**

-.03

85

Correlation

Predicted Hits

# Predicting Success: To Pimp A Butterfly

.28

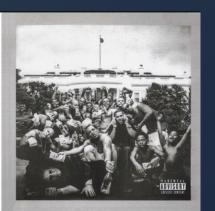
9

3

Correlation

Predicted Hits

**Actual Hits** 





Collect More songs

Artist with and without hits, include star flops

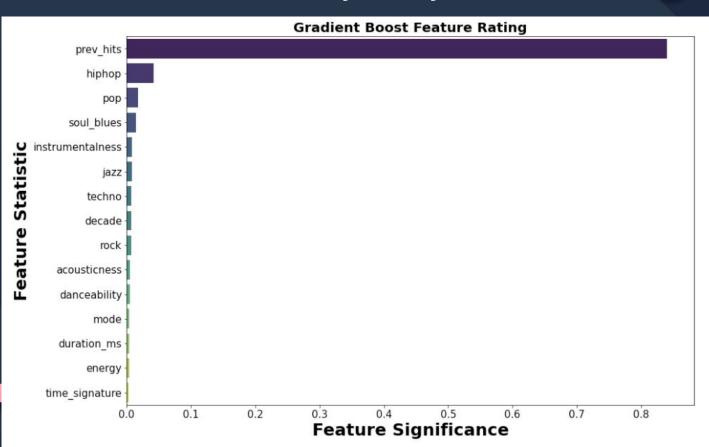
"Realistic" Target distribution

Focus in on recent decades for sales today

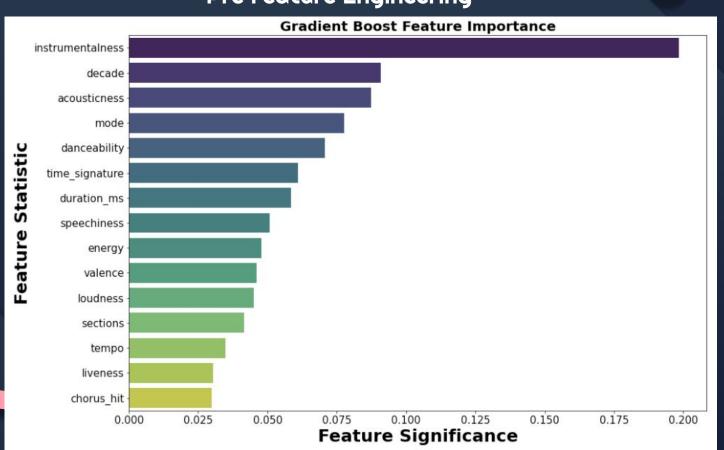
**Total Streams** 



### Post Feature Engineering



### Pre Feature Engineering



# Turn It Up

