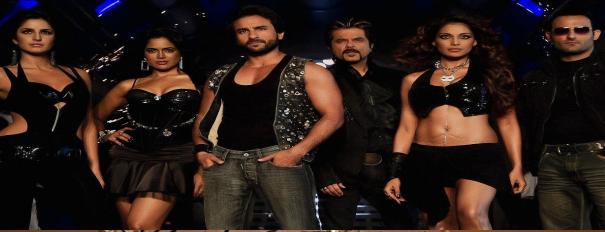


International Pop Genre Classifier

BY ANGELA CHEN, SAM
KARIMBIL, PRAGATI
TOPPO
CSC 466



INTRODUCTION

Why do we care?



Cultural Pop Becoming More Mainstream

International pop is becoming more and more mainstream – as it becomes easier to share across countries and cultures exposure to other types of music is more common.



Barrier to Entry - Language + ‘Cold Start’

However, it might be hard for some people to start listening to international pop – whether it's because of the language barrier, or even just because they don't know where to start.



Music is Universal!

But, music is universal! People can enjoy the rhythm or melody just as much even if it's in another language! We can start introducing new songs based on existing preferences.

OVERVIEW

Music Genres + Models

- ❖ **K-pop**
- ❖ **J-pop**
- ❖ **C-pop**
- ❖ **I-pop**

- ❖ **Dummy**
- ❖ **RandomForest**
- ❖ **SVM**
- ❖ **XGBoost**

We chose these genres because we are the most familiar with them. These classifiers are all commonly used models, with the 'dummy' as our baseline (always recommending Indian) to compare between models.

Features we used

Features that are audio related, eg: tempo, acousticness, danceability, energy

Sentiment from lyrics

Features we did not use

Features that reveal too much: eg: artist, title, language

Translated lyrics into english to avoid language biases

FINDINGS

Features

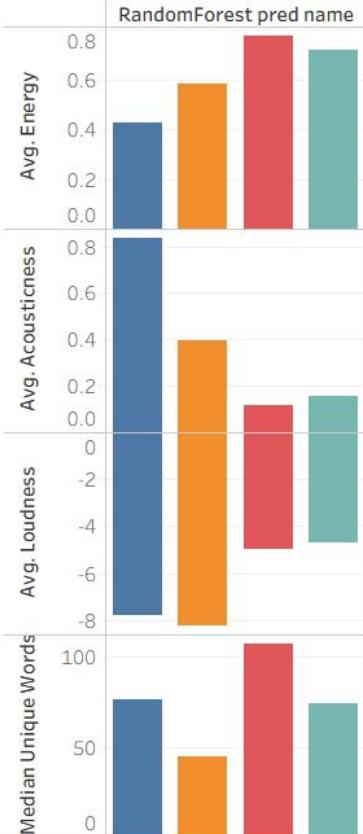
C-pop
lowest energy,
highest acousticness

I-pop
high loudness,
least unique words

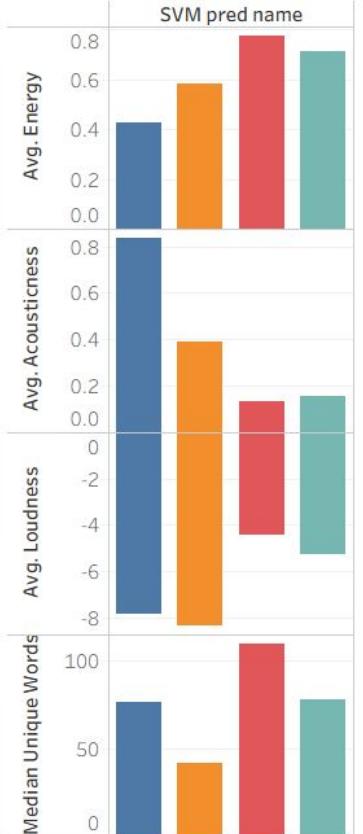
J-pop
highest energy,
lowest acousticness,
highest unique words

K-pop
most 'average' –
high energy, middling
loudness/unique
words

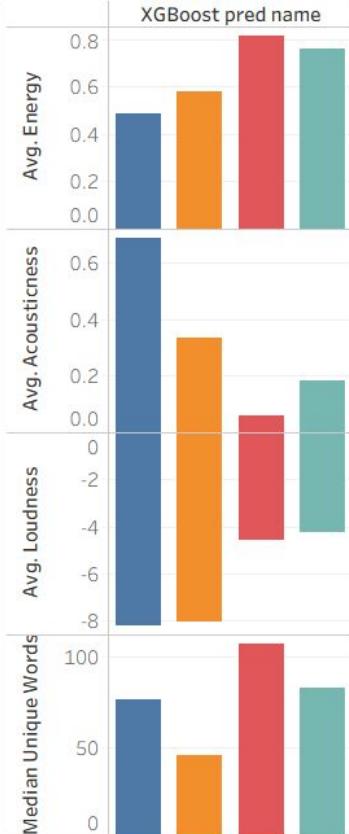
RandomForest



SVM



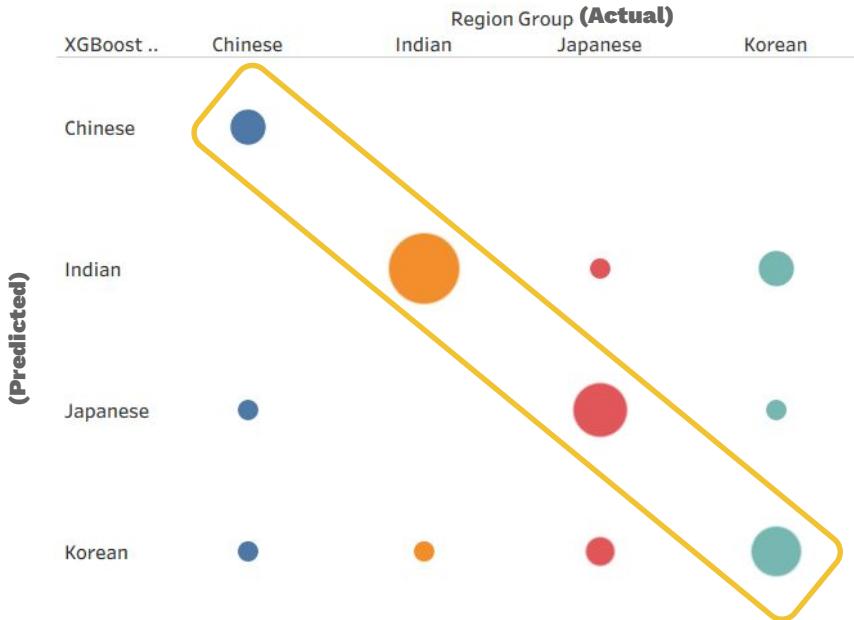
XGBoost



Model Selection

XGBoost

Actual vs Predicted XGBoost



```
XGBoost
Classes: ['Chinese', 'Indian', 'Japanese', 'Korean']
Confusion Matrix (rows=true, cols=pred):
[[ 3  0  1  1]
 [ 0 12  0  1]
 [ 0  1  7  2]
 [ 0  3  1  6]]
Accuracy: 0.7368
Macro F1: 0.7286
Weighted F1: 0.7336
Macro ROC AUC: 0.8934711122211123

Classification Report:
              precision    recall  f1-score   support
Chinese        1.0000   0.6000   0.7500      5
Indian         0.7500   0.9231   0.8276     13
Japanese       0.7778   0.7000   0.7368     10
Korean          0.6000   0.6000   0.6000     10

accuracy           0.7368      -      -
macro avg       0.7819   0.7058   0.7286     38
weighted avg    0.7507   0.7368   0.7336     38
```

APPLICATIONS

American Pop Songs and their Respective Classification

Here are a couple of examples of what we found when we attempted classifying American pop songs using our model

Believer - Imagine Dragons
Firework - Katy Perry
Dynamite - BTS

K-pop

Piano Man - Billy Joel
Put Your Head on My Shoulder - Paul Anka
La Vie En Rose - Daniela Andrade

C-pop

Smells Like Teen Spirit - Nirvana
Bring Me to Life - Evanescence
Welcome to the Black Parade - My Chemical Romance

J-pop

Sick Boy - The Chainsmokers
No Role Modelz - J. Cole
Sympathy for the Devil - The Rolling Stones

I-pop

Thank you

