



# Predicting popularity and preference of Mandarin music with Machine Learning

指導教授：朱威達

學生：鄭至善

## • Introduction

Exploring the potential success of a given song based on its acoustic characteristics is both interesting from a scientific point of view and something that could be beneficial to the music industry. Therefore, this work primarily evaluates the Mandarin music which are published on Youtube platform. With the approach of exploiting low- and high- level audio features, this project models the prediction through classification and regression methods.

## • Dataset

The experiments make use of 1923 Mandarin songs released in second half of 2019. To extract the low- and high- level audio features, it relies on representative 30 seconds samples for each song and extract features by Essentia framework. Furthermore, we take the subscriber count of given song's uploader into account in order to improve prediction results.

Category	Features
Basic	Replay Gain, BPM, Dissonance, Pitch Saliency
Mood	Acoustic, Aggressive, Electronic, Happy, Party, Relaxed, Sad, Hu and Downie's 5 clusters of mood
Genre	Alternative, Blues, Electronic, FolkCountry, FunkSoulRNB, Jazz, Pop, RapHippop, Rock
Rhythm	Danceability
Voice	Gender, Instrumental, Atonal, Timbre

Fig 1. Audio features computed by Essentia.

\*The first category contains solely low-level features, whereas the other categories contain high-level features.

## • Experiments

Approach	RMSE	MAE	R <sup>2</sup>
Linear regression	4573398.56	1414113.07	-0.15
Support Vector regression	4652718.97	1009487.51	-0.05

Fig 2. Result of experiment 1

Approach	Accuracy	Precision	Recall	F1-score
Logistic regression	0.537129	0.557757	0.342543	0.417422
K nearest neighbor	0.679527	0.733602	0.564820	0.632266
Support vector machine	0.701435	0.697224	0.711664	0.702229

Fig 3. Result of experiment 2

Approach	Accuracy	Precision	Recall	F1-score
Logistic regression	0.588113	0.568564	0.768051	0.651757
K nearest neighbor	0.659712	0.669574	0.638698	0.651487
Support vector machine	0.690777	0.692992	0.696477	0.692789

Fig 4. Result of experiment 3

## • Conclusion

Three experiments both show that audio features have minimal influence on popularity and preference. Perhaps considering new features such as lyrics, more singer related information may lead to a better outcome.

\*Features with strikethrough format are deprecated.  
Adopting those features is helpless to improve accuracy.

### Experiment 1 - Predicting view count

Feature selection: Subscribe\_cnt, Rhythm, ~~Genre~~, ~~Mood~~, ~~Voice~~, ~~Basic~~

### Experiment 2 - Classification by view count(160000)

Feature selection: Subscribe\_cnt, ~~Genre~~, ~~Mood~~, ~~Rhythm~~, ~~Voice~~, ~~Basic~~

### Experiment 3 - Classification by degree of preference(97.85%)

Feature selection: Subscribe\_cnt, Rhythm, ~~Genre~~, ~~Mood~~, ~~Voice~~, ~~Basic~~

$$\begin{aligned} & \text{*Degree of preference} \\ &= \frac{\text{like count}}{\text{like count} + \text{dislike count}} * 100\% \end{aligned}$$