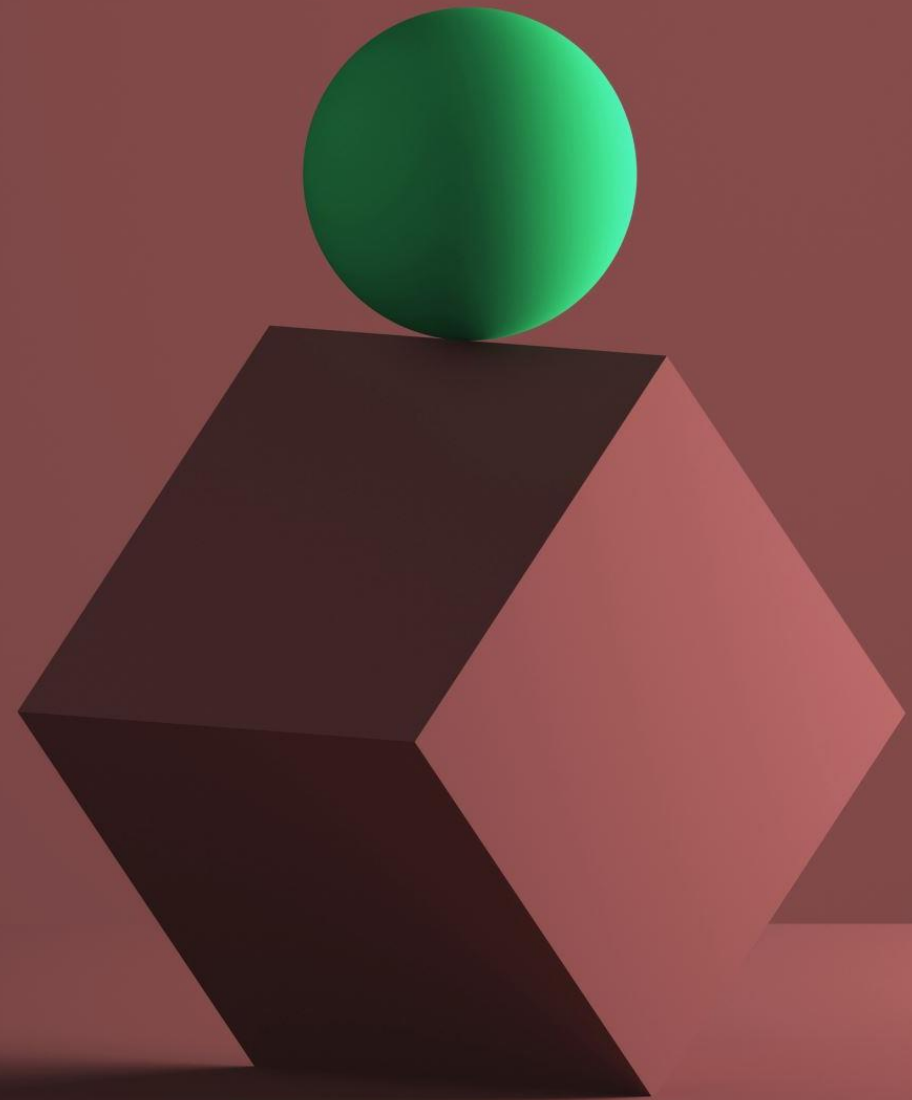


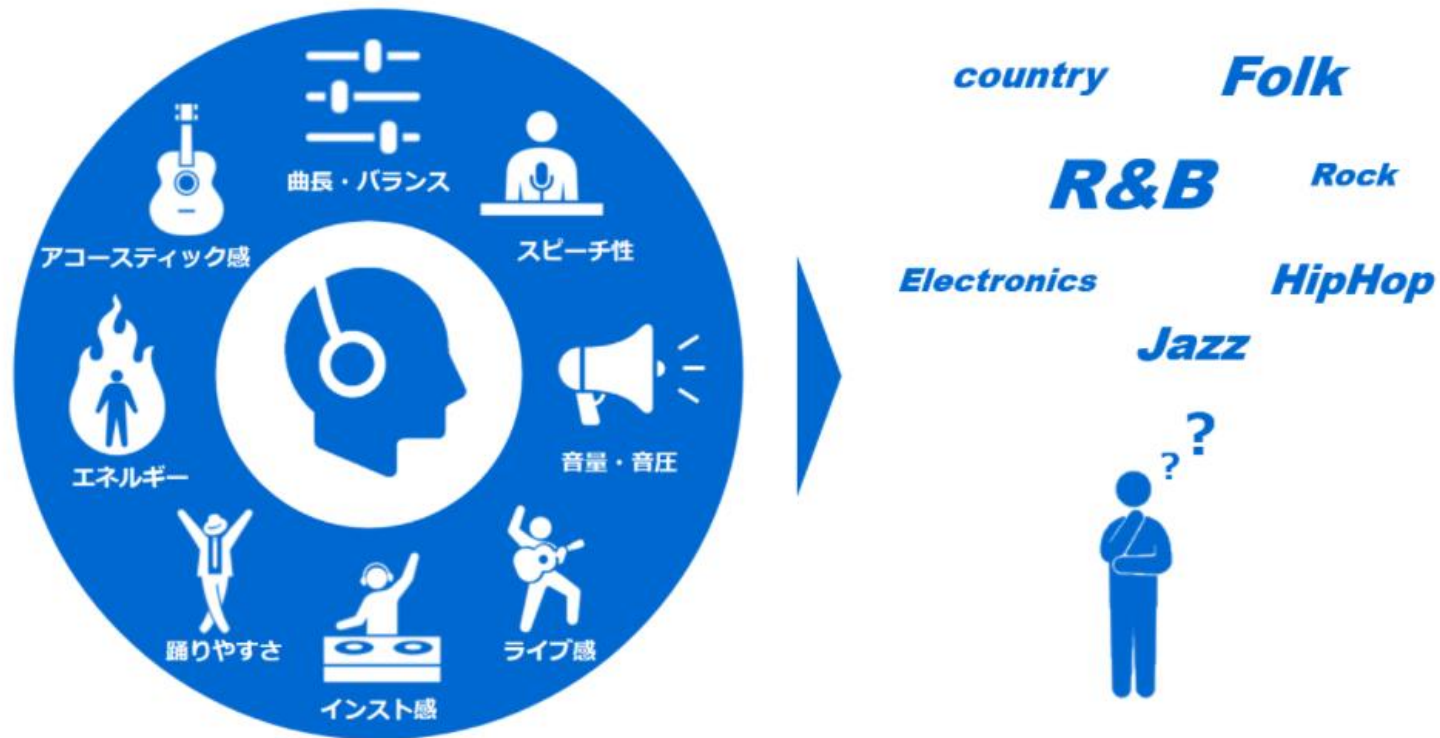
# Student Cup 2021

インサイト部門

楽曲の印象による分類法の検討

yama09





## 概要

楽曲の特徴からジャンルを分類したい。この時に何かいいアイデアはないだろうか。本発表では、楽曲の特徴から印象を割り出し特徴量に追加することで精度向上を図る。

そもそも...

音楽はどのような特徴で構成されているか？

調性  
(key)

テンポ  
(tempo)

音高  
(pitch)

リズム  
(rhythm)

和声  
(harmony)

旋律  
(melody)



構成要素  
が分かる

分類は  
簡単では？

ということで...

予測部門のコンペで構成要素は活躍したのか確認してみる

## 予測部門コンペの特徴量重要度

- ・ 構成要素に近い特徴量が上位
- ・ 加工した特徴量 < 加工なしの特徴量



構成要素の特徴量が活躍！！



図1 コンペの特徴量データ

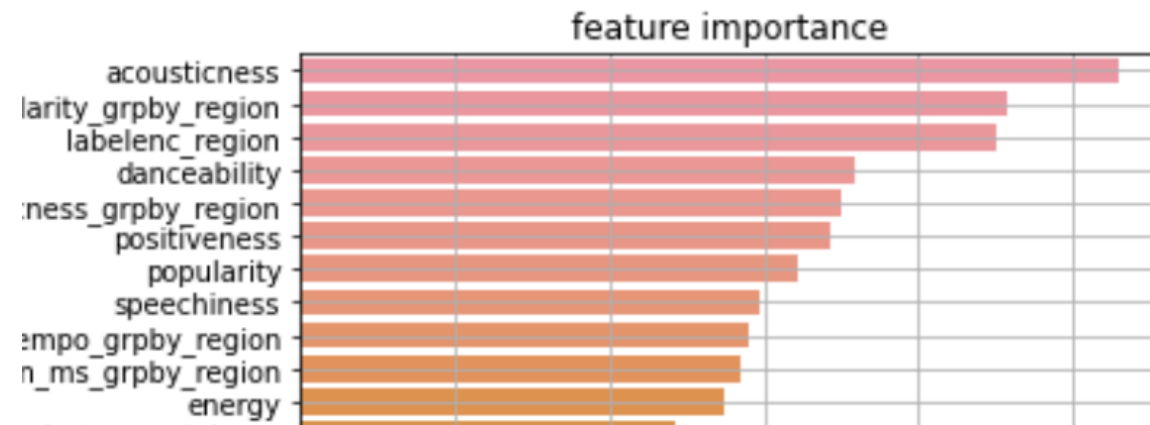


図2 最終 14 位の特徴量重要度

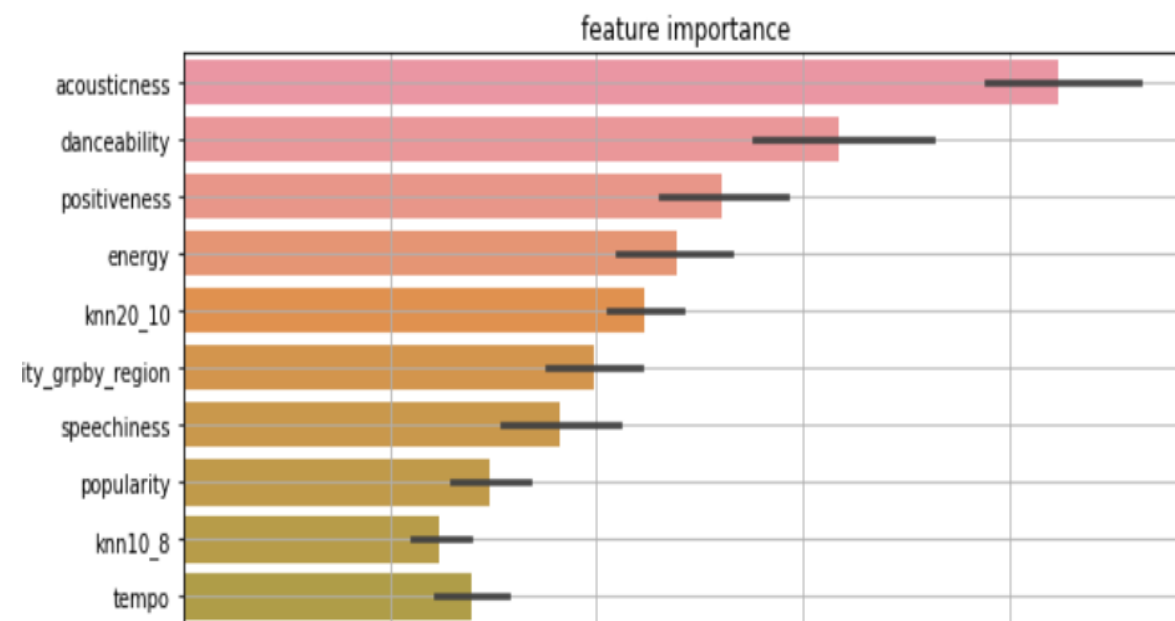


図3 最終 5 位の特徴量重要度

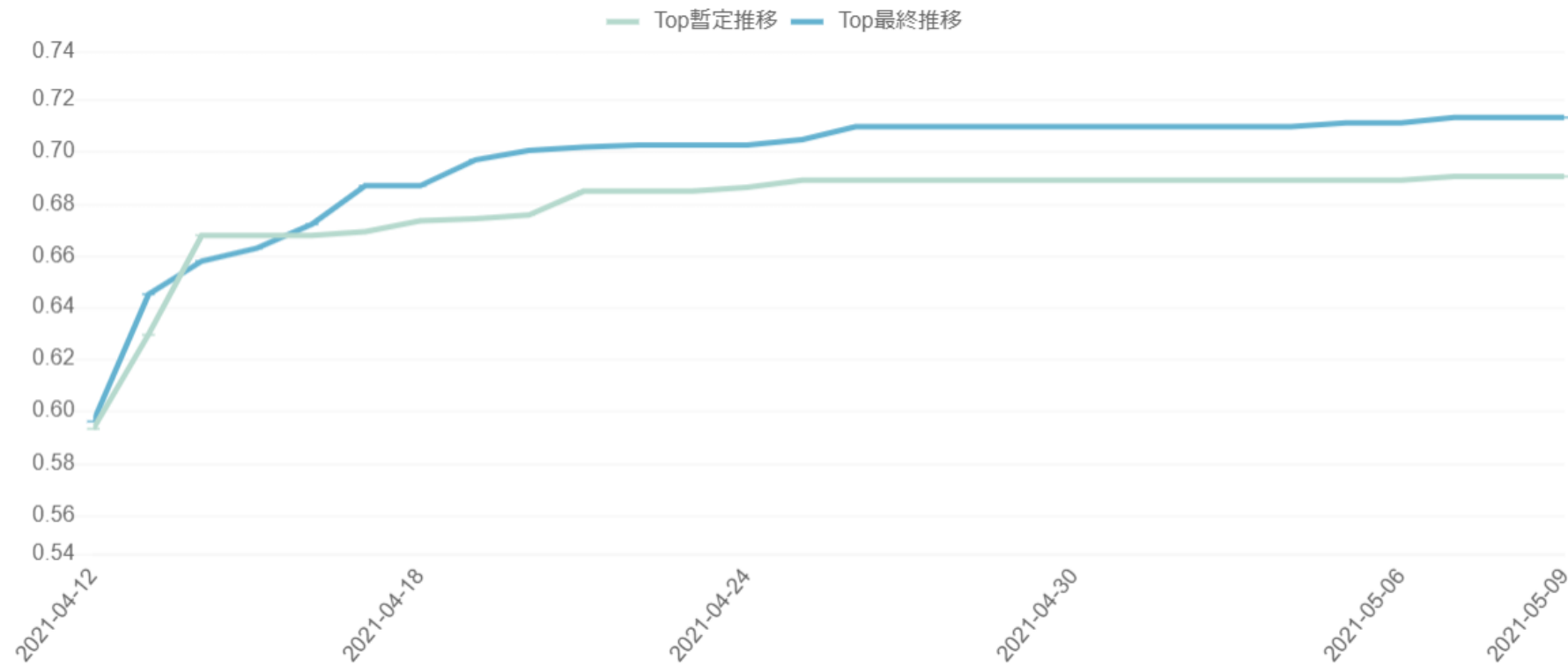


図4 予測部門でのTOP予測精度の推移

- ・ 構成要素の活躍はあったが70%の精度
- ・ 精度を上げるのが難しいのではないか？

構成要素を使用した予測精度には限界がある？

構成要素の影響ではなく、そもそも  
音楽ジャンル分類に限界があるという研究データがある

(音響ベースの音楽信号の分類 George Tzanetakis 訳：角尾衣未尾)



精度が低い原因は音楽のあいまい性

## あいまい性とは

---

### ①作曲家によるきめつけた分類

- ・正しい分類と間違った分類が混じる

---

### ②ジャンルの境界線が世代で変化

- ・時代により分類基準が変わる



# あいまい性を収束させるには？

- ・ 構成要素の他に重要な特徴量が必要  
(条件：誰もが同じ認識の特徴)



音楽の印象はみんな同じ認識である

音楽の構成要素と印象には**相関関係**がある (※1, ※2)  
よって、構成要素から**印象というメタデータ**を獲得できる

- ※1 “Experimental Studies of the Elements of Expression in Music” Kate Hevner
- ※2 “The affective value of pitch and tempo in music” Kate Hevner



## 分類する方法

---

### 3段階で分類を行う

---

**Step1** : 楽曲と印象の相関を調べる

---

**Step2** : 相関からメタデータを作成

---

**Step3** : 検証の実施

# Step1: 楽曲と印象の相関

- ・ 印象を 8 つに分類
- ・ C1 ～ C8 と定義する

		c6 bright cheerful gay happy joyous merry	
	c7 agitated dramatic exciting exhilarated impetuous passionate restless sensational soaring triumphant	c5 delicate fanciful graceful humorous light playful quaint sprightly whimsical	c4 calm leisurely lyrical quiet satisfying serene soothing tranquil
c8 emphatic exalting majestic martial ponderous robust vigorous	c1 awe-inspiring dignified lofty sacred serious sober solemn spiritual	c2 dark depressing doleful frustrated gloomy heavy melancholy mournful pathetic sad tragic	c3 dreamy longing plaintive sentimental tender yearning yielding

図 5 Henver の 8 印象語群

# Step1: 楽曲と印象の相関

- ・ 印象と楽曲を対応
- ・ Henverの研究から  
変換行列を作成

楽曲構成 要素名	印象語群名							
	C1	C2	C3	C4	C5	C6	C7	C8
key	長調	短調	短調	長調	長調	長調	—	—
tempo	遅い	遅い	遅い	遅い	速い	速い	速い	速い
pitch	低い	低い	高い	高い	高い	高い	低い	低い
rhythm	固定	固定	流動	流動	固定	流動	固定	固定
harmony	単純	複雑	単純	単純	単純	単純	複雑	複雑
melody	上昇	—	—	上昇	下降	—	下降	下降

図6 構成要素と印象の相関表



	<i>key'</i>	<i>tempo'</i>	<i>pitch'</i>	<i>rhythm'</i>	<i>harmony'</i>	<i>melody'</i>
<i>C<sub>1</sub></i>	4	-14	-10	18	3	4
<i>C<sub>2</sub></i>	-12	-12	-19	3	-7	0
<i>C<sub>3</sub></i>	-20	-16	6	-9	4	0
<i>C<sub>4</sub></i>	3	-20	8	-2	10	3
<i>C<sub>5</sub></i>	21	6	16	8	12	-3
<i>C<sub>6</sub></i>	24	20	6	-10	16	0
<i>C<sub>7</sub></i>	0	6	-13	10	-8	-8

図7 相関表を基にした変換行列

## Step2: メタデータ作成

- ・ 相関値から印象のデータを作成
- ・ ラベル付けを行う

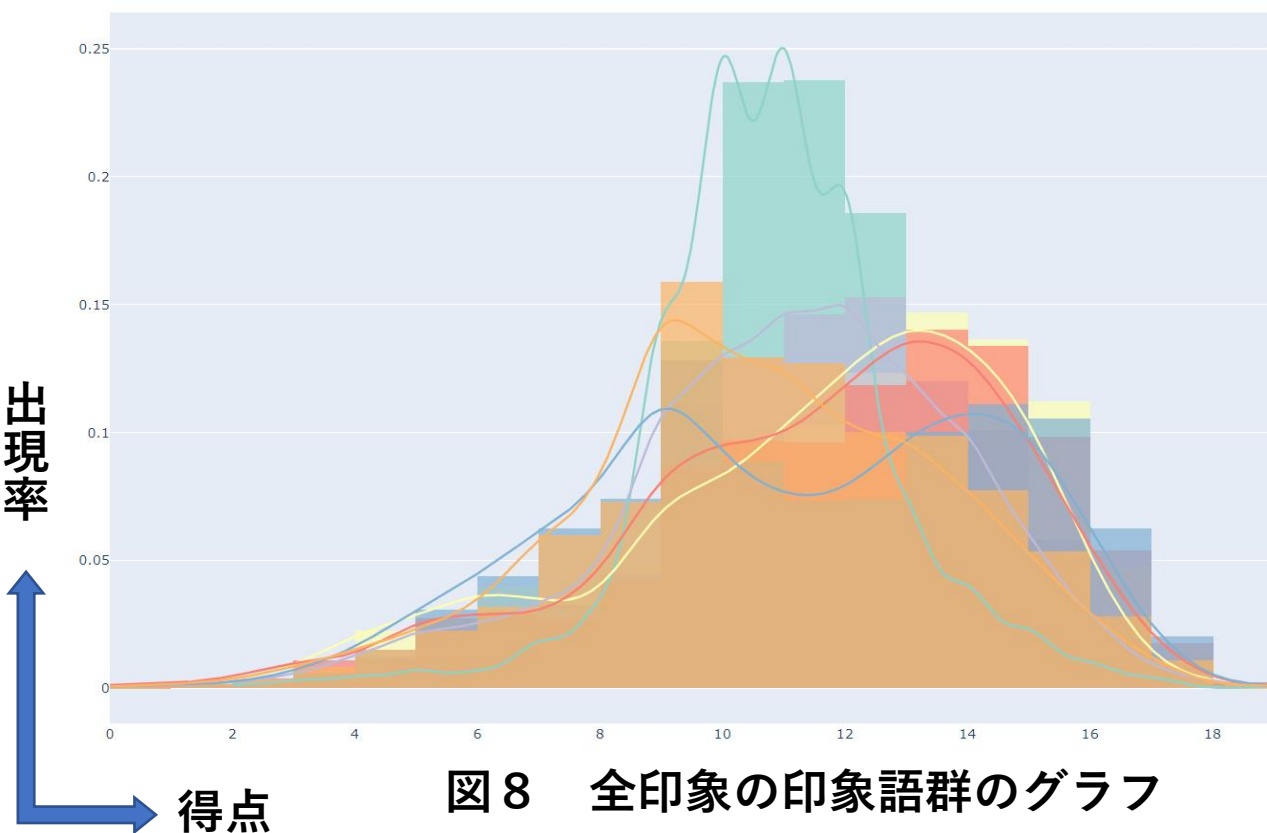
- ① 音楽構成要素の特徴量を標準化
- ② 図7の相関値を用いて重みづけ
- ③ 重みづけした構成要素を足し合わせる  
(印象値のデータフレームになる)

楽しい印象	落ち着いた印象
- 8.93	1.78
4.22	-3.76
- 6.11	-8.87
1.09	2.22

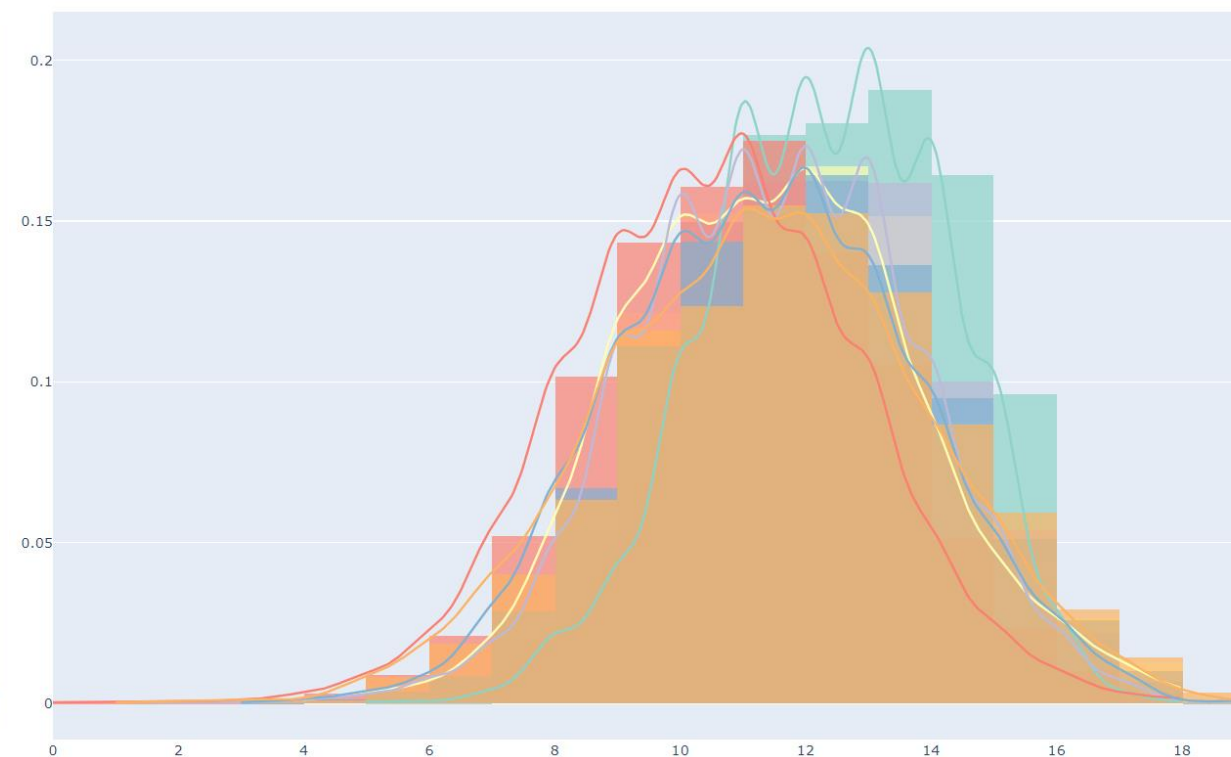
- ④ 20個にbin分割し、得点化する

# Step3 : 検証の実施

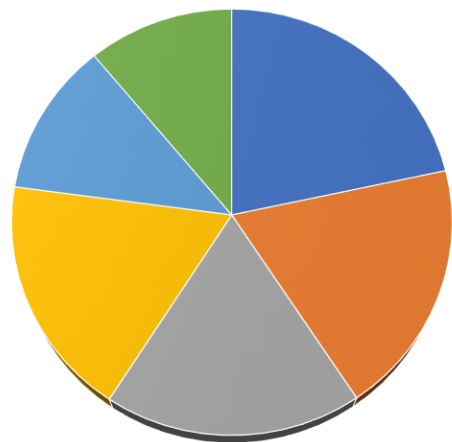
- 2つの図で分布が異なる。各ジャンルの構成比率に着目し分析する



※ 出現率 = 全データに対するデータ割合



# Step3 : 検証の実施



■ pop ■ rock ■ rap ■ R&B ■ latin ■ edm

図10 全体の構成割合

以上の結果より楽しいイメージのあるPOPやEDMが増加している。  
これにより印象による分類ができていると判断できる。

増減率

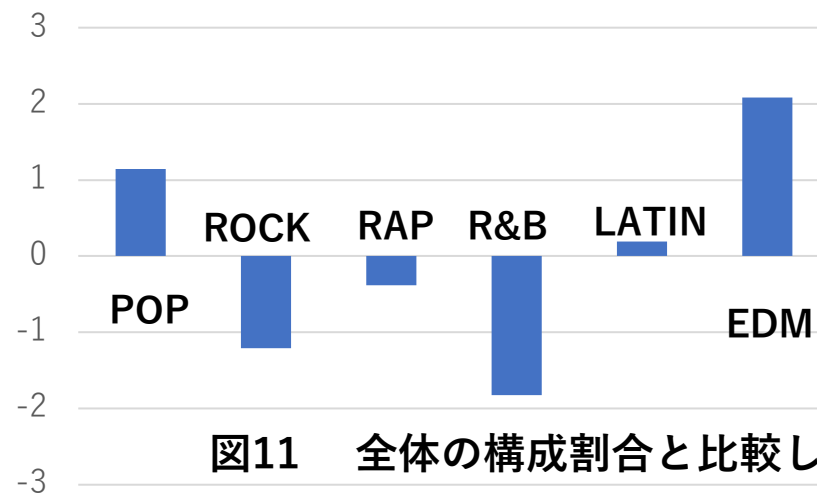


図11 全体の構成割合と比較した増減率

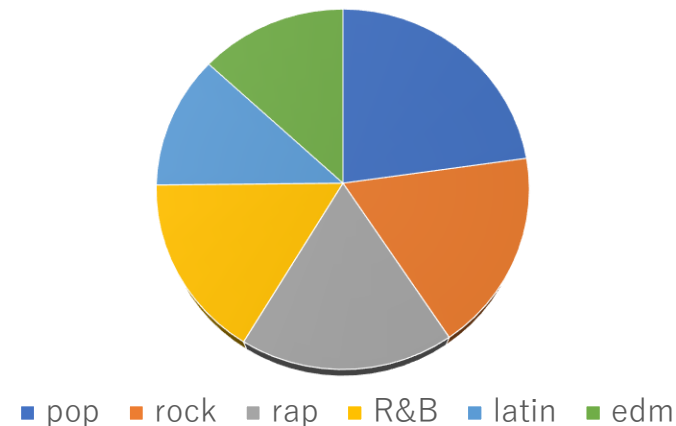


図12 出現率上位の楽しい印象の構成割合



# まとめ

- ・ 音楽のあいまい性により分類が難しい。
- ・ 楽曲の構成要素に印象という特徴量を追加した。
- ・ 出現率上位の構成割合が印象のイメージにあうものになり正しく分類できていると判断した。