

---

ディープニューラルネットワークを用いた  
顔画像による野球選手・サッカー選手の分類

日本大学 文理学部 情報科学科 谷聖一研究室

田名網 勇希

---

---

# 目次

---

1. はじめに
2. 演習方法
3. 演習結果
4. おわりに

---

# 1. はじめに

## 1.1 背景(1)

---

「〇〇は**野球顔**だよね」

「△△くんは**サッカー顔**っぽい」

---

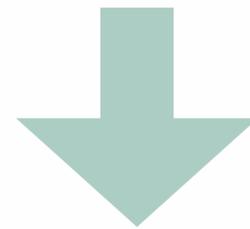
# 1. はじめに

## 1.1 背景(2)

---

野球顔： 丸い顔・丸い目

サッカー顔： 縦長な顔・細長い目



野球選手・サッカー選手の顔を見ても

**野球顔**なのか**サッカー顔**なのか分からない

---

# 1. はじめに

## 1.2 動機

---

### 顔画像から

### 野球選手とサッカー選手を分類できるのか

---

# 1. はじめに

## 1.3 先行研究(1)

---

Yilun Wang, Michal Kosinskiらが

**顔画像による**

**性的指向（同性愛者 or 異性愛者）の分類において**

**人間の分類精度を超えたという論文を発表**

(Wang, Y., & Kosinski, M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*. (in press))

6

# 1. はじめに

## 1.3 先行研究(2)

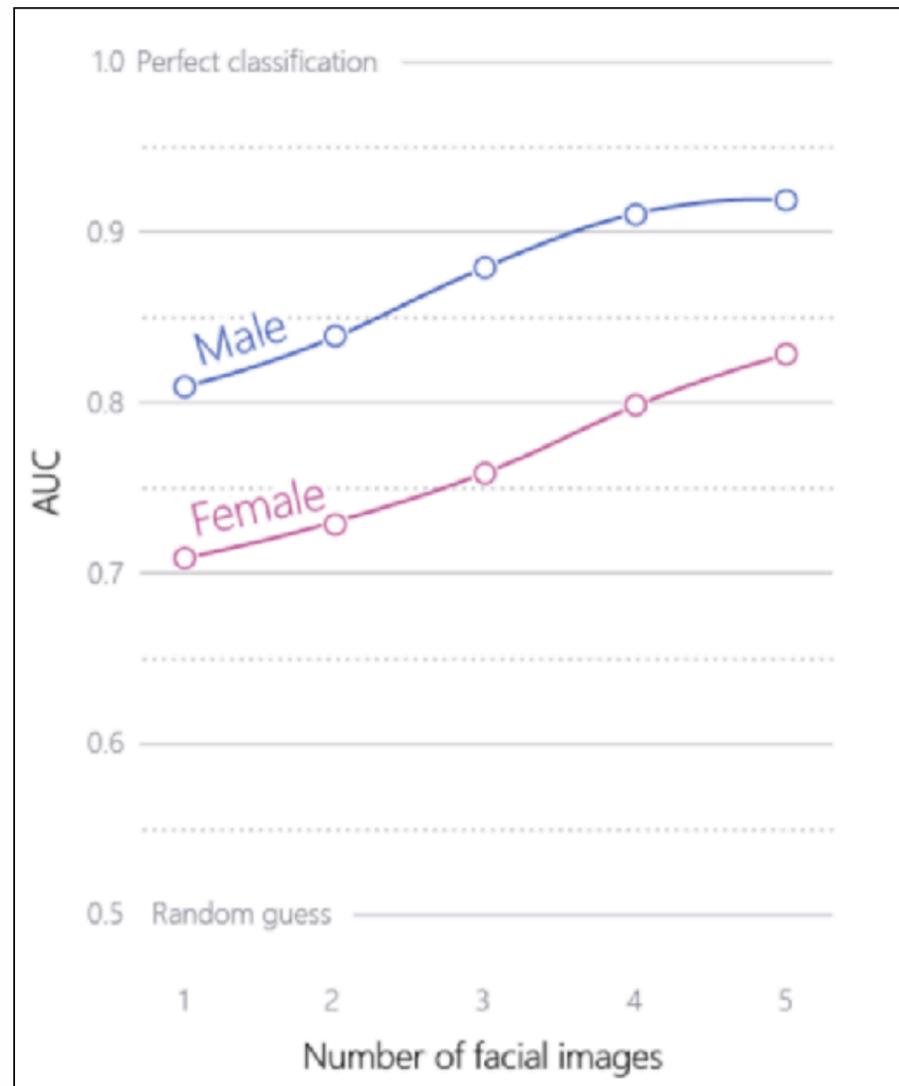
- 18~40歳の白人を対象
- アメリカの出会い系サイトのプロフィール画像を使用

	男性		女性		合計
	ゲイ	ストレート	レズ	ストレート	
人数	3,947人	3,947人	3,441人	3,441人	<b>14,776人</b>
枚数	8,996枚	8,645枚	7,457枚	10,228枚	<b>35,326枚</b>

(Wang, Y., & Kosinski, M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*. (in press))

# 1. はじめに

## 1.3 先行研究(3)



- 1人あたりの顔画像が**1枚**
  - 男性：81%
  - 女性：71%
- 1人あたりの顔画像が**5枚**
  - 男性：91%
  - 女性：83%

(Wang, Y., & Kosinski, M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*. (in press))

---

# 1. はじめに

## 1.4 目的

---

### 論文内の分類手法を用いて

### 顔画像から野球選手・サッカー選手を分類

## 2. 演習方法



---

## 2. 演習方法

### 2.1 画像収集(1)

---

#### **Bing Image Search API**を使用して画像を収集

- 野球選手

**NPB**に所属する日本人選手（2017/12/22）

- サッカー選手

**J1, J2, J3**に所属する日本人選手（2017/12/22）

## 2. 演習方法

### 2.1 画像収集(2)

収集した画像

年齢	人数		枚数	
	野球選手	サッカー選手	野球選手	サッカー選手
16~19歳	85人	259人	3,557枚	11,640枚
20~29歳	580人	1153人	27,787枚	53,561枚
30~39歳	183人	465人	8,917枚	22,581枚
40~49歳	5人	7人	244枚	338枚
50~59歳	-	2人	-	90枚
合計	853人	1,886人	40,505枚	88,210枚
	<b>2,739人</b>		<b>128,715枚</b>	

## 2. 演習方法【再掲】



## 2. 演習方法

### 2.2 画像の前処理(1)

#### 収集した画像のイメージ

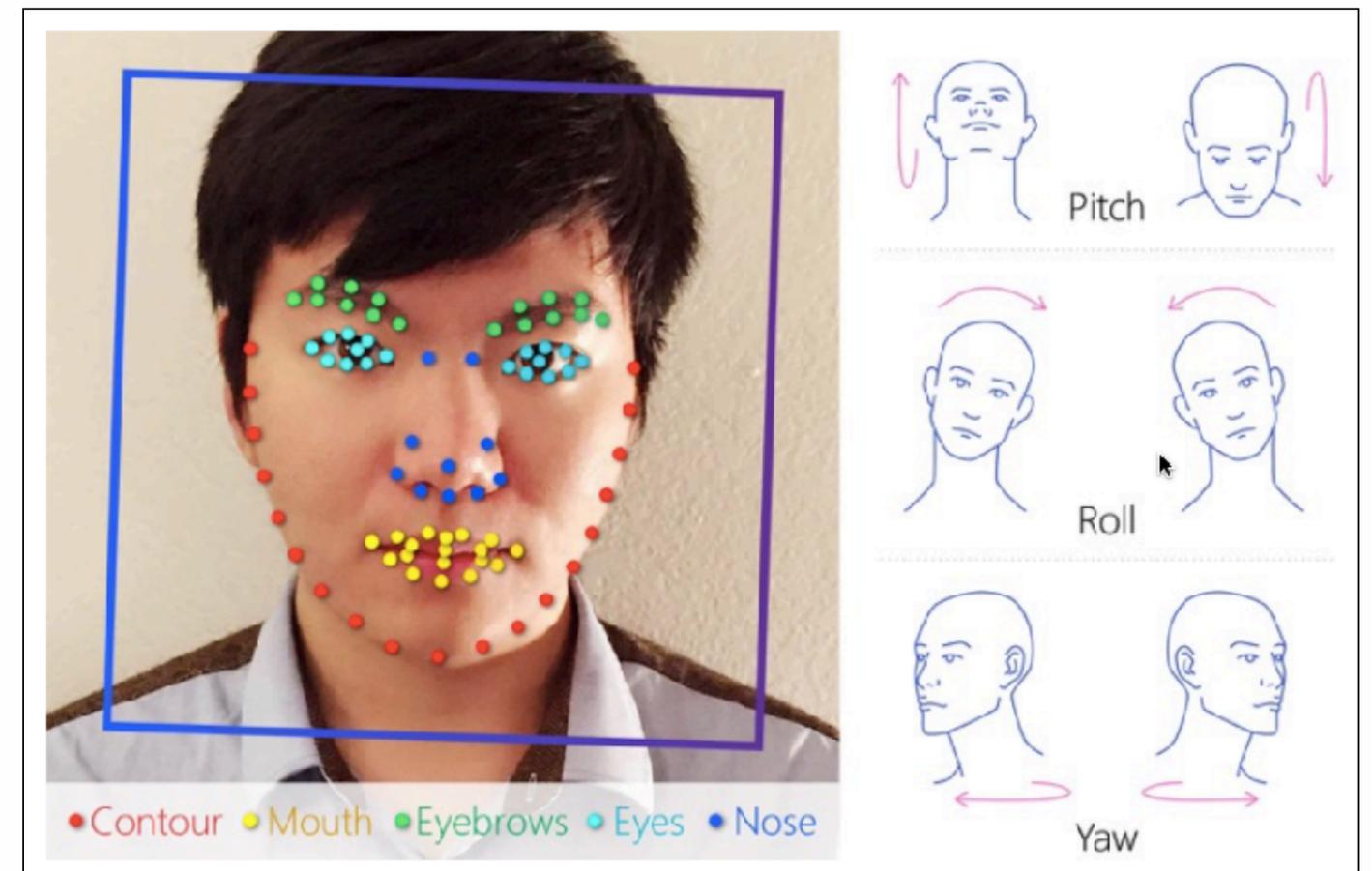


## 2. 演習方法

### 2.2 画像の前処理(2)

#### Face++

- 顔認識サービス
- 画像から顔に関する情報を取得



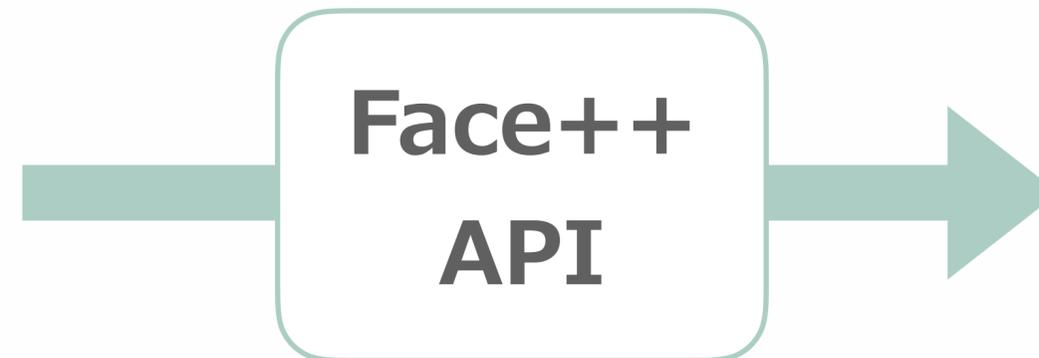
(Face++ : <https://www.faceplusplus.com/face-detection/>)

(図 : Wang, Y., & Kosinski, M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*. (in press))

## 2. 演習方法

### 2.2 画像の前処理(3)

#### Face++



```
{
  "image_id": "BZB9LxfBpWJ7Ovph8mNSSA==",
  "request_id": "1515174866,8f253f92-e8de-4b90-b7c2-83623a56c754",
  "time_used": 211,
  "faces": [
    {
      "landmark": {
        "mouth_upper_lip_left_contour2": {"y": 135, "x": 70},
        "mouth_upper_lip_left_contour3": {"y": 138, "x": 73},
        "mouth_lower_lip_right_contour3": {"y": 144, "x": 89},
        "mouth_upper_lip_left_contour1": {"y": 132, "x": 77},
        "left_eye_upper_left_quarter": {"y": 86, "x": 58},
        "left_eyebrow_lower_middle": {"y": 74, "x": 62},
        "contour_chin": {"y": 172, "x": 81},
        "left_eyebrow_lower_left_quarter": {"y": 75, "x": 55},
        "right_eyebrow_lower_left_quarter": {"y": 76, "x": 104},
        "mouth_lower_lip_right_contour1": {"y": 139, "x": 91},
        "mouth_lower_lip_left_contour2": {"y": 142, "x": 69},
        "left_eye_bottom": {"y": 91, "x": 63},
        "mouth_lower_lip_bottom": {"y": 145, "x": 82},
        "contour_left9": {"y": 169, "x": 68},
        "mouth_lower_lip_top": {"y": 139, "x": 82},
        "right_eyebrow_upper_middle": {"y": 72, "x": 111},
        "right_eyebrow_left_corner": {"y": 75, "x": 97},
        "right_eye_bottom": {"y": 95, "x": 106},
        "contour_left7": {"y": 135, "x": 70}
      }
    }
  ]
}
```

(Face++ : <https://www.faceplusplus.com/face-detection/>)

## 2. 演習方法

### 2.2 画像の前処理(4)

検出した顔の数が, 0または複数の画像を除去



## 2. 演習方法

### 2.2 画像の前処理(4)

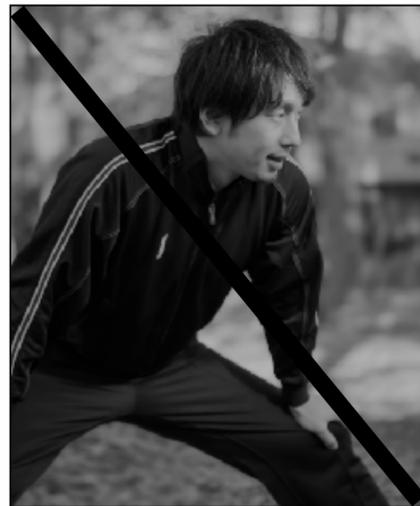
顔が一定サイズ以下の画像を除去



## 2. 演習方法

### 2.2 画像の前処理(4)

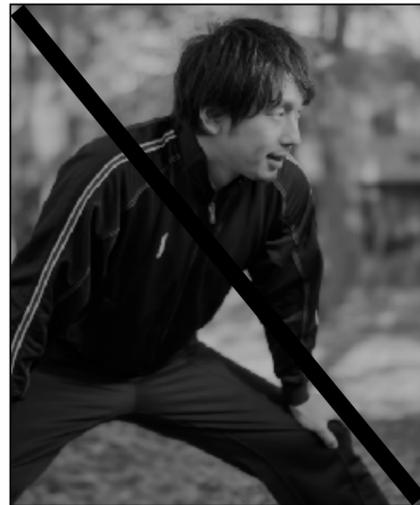
顔が一定角度以上, 横または上下を向いている画像を除去



## 2. 演習方法

### 2.2 画像の前処理(4)

#### 顔の切り取り



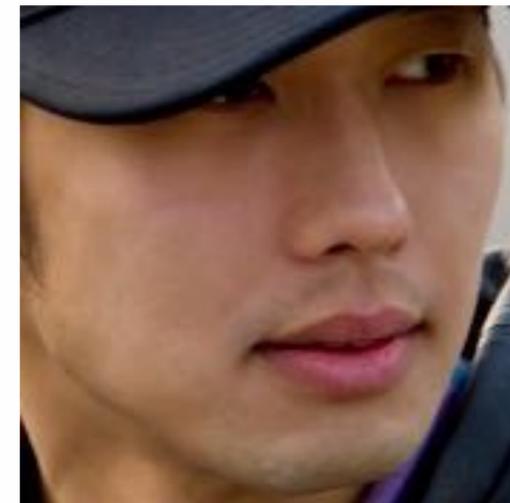
---

## 2. 演習方法

### 2.2 画像の前処理(6)

---

#### 顔の切り取り



---

## 2. 演習方法

### 2.2 画像の前処理(5)

---

#### 手作業で以下を満たす画像を除去

- ・ 収集した人物と一致していない
- ・ 顔の輪郭, 眉毛が隠れている
- ・ 顔の領域内に顔以外のものが写っている  
(眼鏡・手・飲み物・食べ物等)
- ・ 野球選手, サッカー選手であることがわかる  
(ボール・グラブ・ヘルメット・帽子・帽子の影等)

---

## 2. 演習方法

### 2.2 画像の前処理(6)

---

#### 手作業で除去



## 2. 演習方法

### 2.2 画像の前処理(4)

#### 前処理後の画像

年齢	人数		枚数	
	野球選手	サッカー選手	野球選手	サッカー選手
16~19歳	68人	68人	156枚	144枚
20~29歳	437人	437人	1,101枚	1,128枚
30~39歳	116人	116人	297枚	384枚
40~49歳	5人	5人	16枚	16枚
50~59歳	-	-	-	-
合計	626人	626人	1,570枚	1,672枚
	<b>1,252人</b>		<b>3,242枚</b>	

## 2. 演習方法【再掲】



---

## 2. 演習方法

### 2.3 VGG FACEによる特徴抽出(1)

---

#### VGG

- 全16層のディープニューラルネットワーク
- オックスフォード大学のVisual Geometry Groupが開発
- ILSVRC2014の分類問題で2位

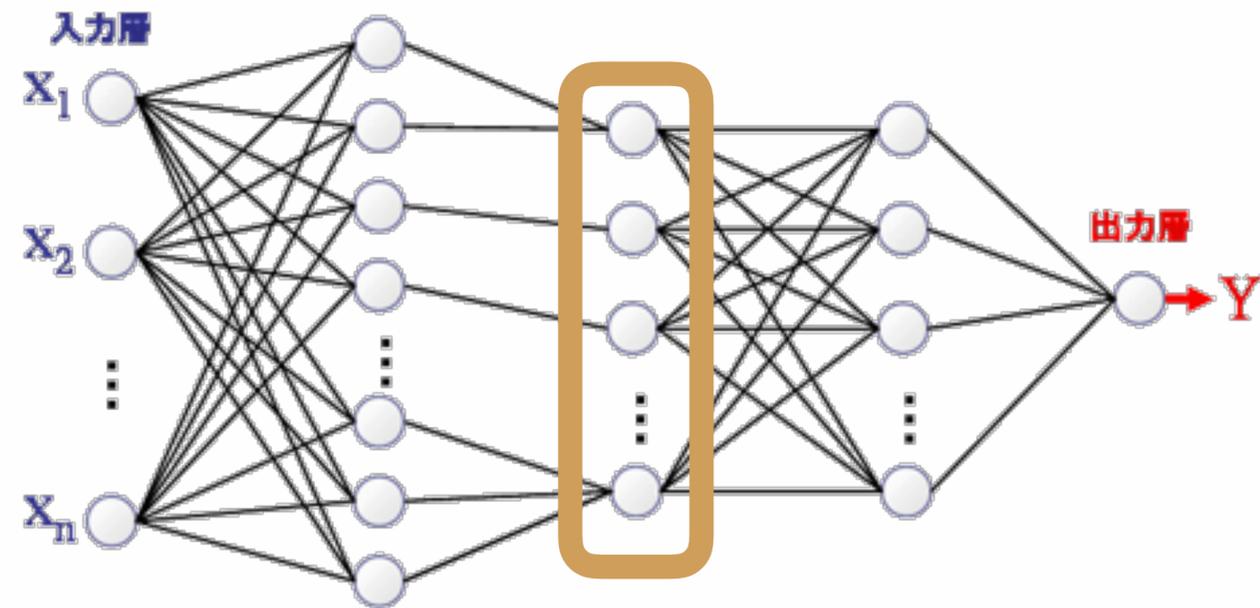
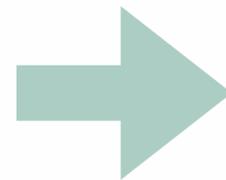
#### VGG Face

- **VGGを顔画像に特化させて学習**させたもの

## 2. 演習方法

### 2.3 VGG FACEによる特徴抽出(2)

#### 特徴抽出



**中間層の出力を取り出す**

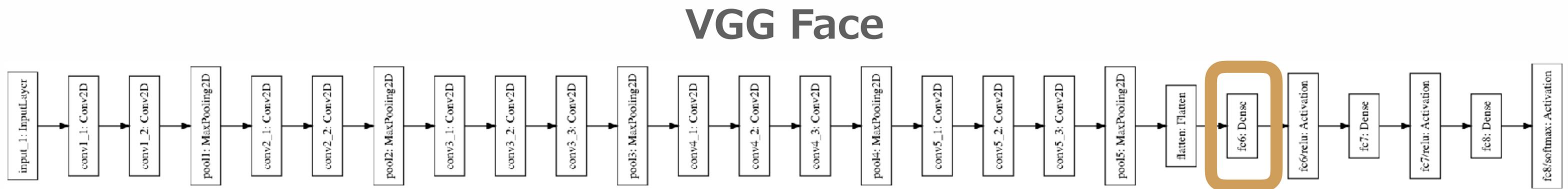
## 2. 演習方法

### 2.3 VGG FACEによる特徴抽出(2)

#### 特徴抽出

画像サイズ：224 x 224

VGG Faceの中間層から、**4096次元**の特徴を抽出



## 2. 演習方法【再掲】



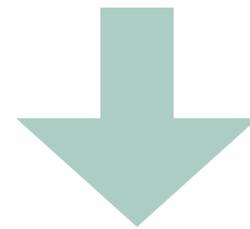
---

## 2. 演習方法

### 2.4 データの次元削減(1)

---

4096次元は次元数が大きい



**次元削減**

## 2. 演習方法

### 2.4 データの次元削減(2)

#### 特異値分解 (Singular Value Decomposition: SVD)

データ<sub>0</sub>

0	1	...	4095
---	---	-----	------

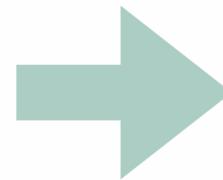
データ<sub>1</sub>

0	1	...	4095
---	---	-----	------

⋮

データ<sub>n-1</sub>

0	1	...	4095
---	---	-----	------



**n × 4096**

0	1	...	4095
0	1	...	4095
0	1	...	4095
⋮			
0	1	...	4095

## 2. 演習方法

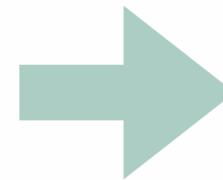
### 2.4 データの次元削減(2)

#### 特異値分解 (Singular Value Decomposition: SVD)

**n × 4096**

0	1	...	4095
0	1	...	4095
0	1	...	4095
⋮			
0	1	...	4095

**500次元に次元削減**



**n × 500**

0	1	...	499
0	1	...	499
0	1	...	499
⋮			
0	1	...	499

## 2. 演習方法【再掲】



---

## 2. 演習方法

### 2.5 分類

---

**ラベル：** 野球選手 or サッカー選手

**出力：** ラベルの発生確率

- ロジスティック回帰 (Logistic Regression: LR)
- サポートベクターマシン (SVC)

## 2. 演習方法【再掲】



---

## 2. 演習方法

### 2.6 評価(1)

---

**評価方法：K-分割交差検証法**

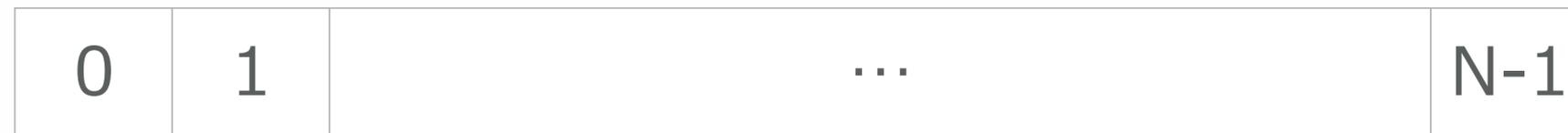
**評価指標：AUC**

## 2. 演習方法

### 2.6 評価(2)

#### K-分割交差検証法 (K-Fold Cross Validation)

データセット (N : データ数)



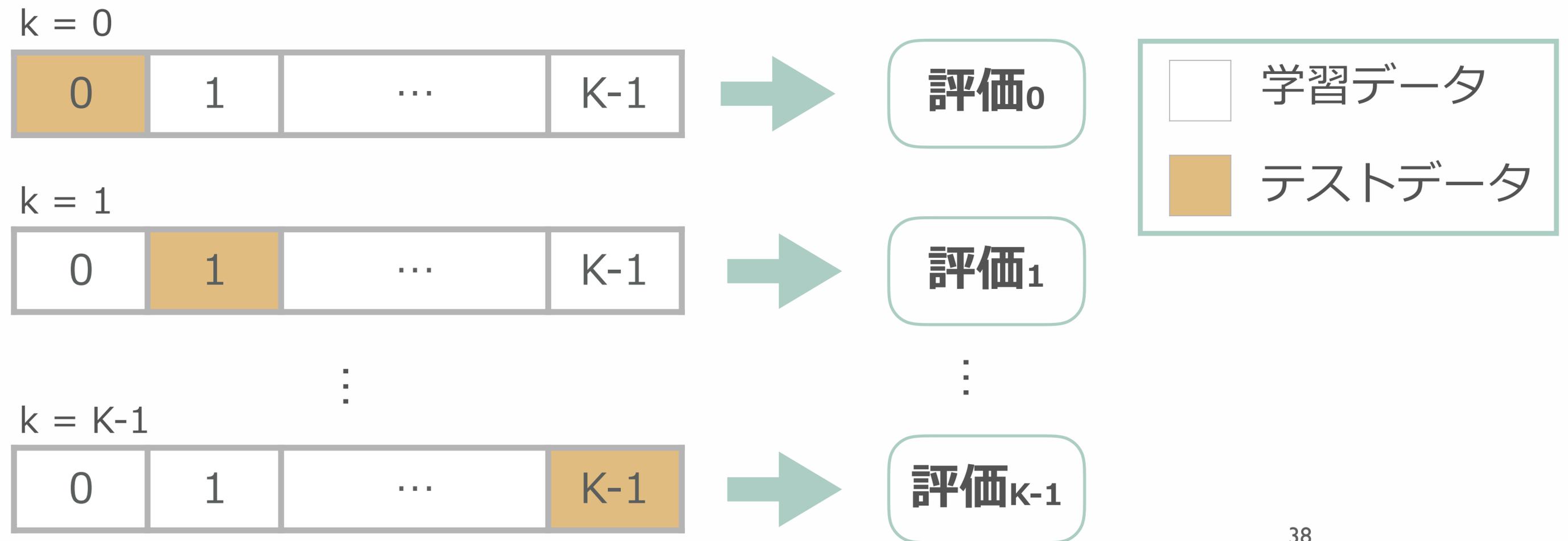
**K個に分割** ( $K \leq N$ )



## 2. 演習方法

### 2.6 評価(3)

#### K-分割交差検証法 (K-Fold Cross Validation)



## 2. 演習方法

### 2.6 評価(4)

#### K-分割交差検証法 (K-Fold Cross Validation)



**K個の評価の平均**

---

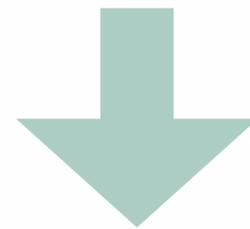
## 2. 演習方法

### 2.6 評価(5)

---

#### K-分割交差検証法 (K-Fold Cross Validation)

**人単位**でのデータの分割を行なった



**学習データとテストデータに  
同一の人物が含まれないようにするため**

---

## 2. 演習方法

### 2.6 評価(1)【再掲】

---

評価方法：K-分割交差検証法

**評価指標：AUC**

---

## 2. 演習方法

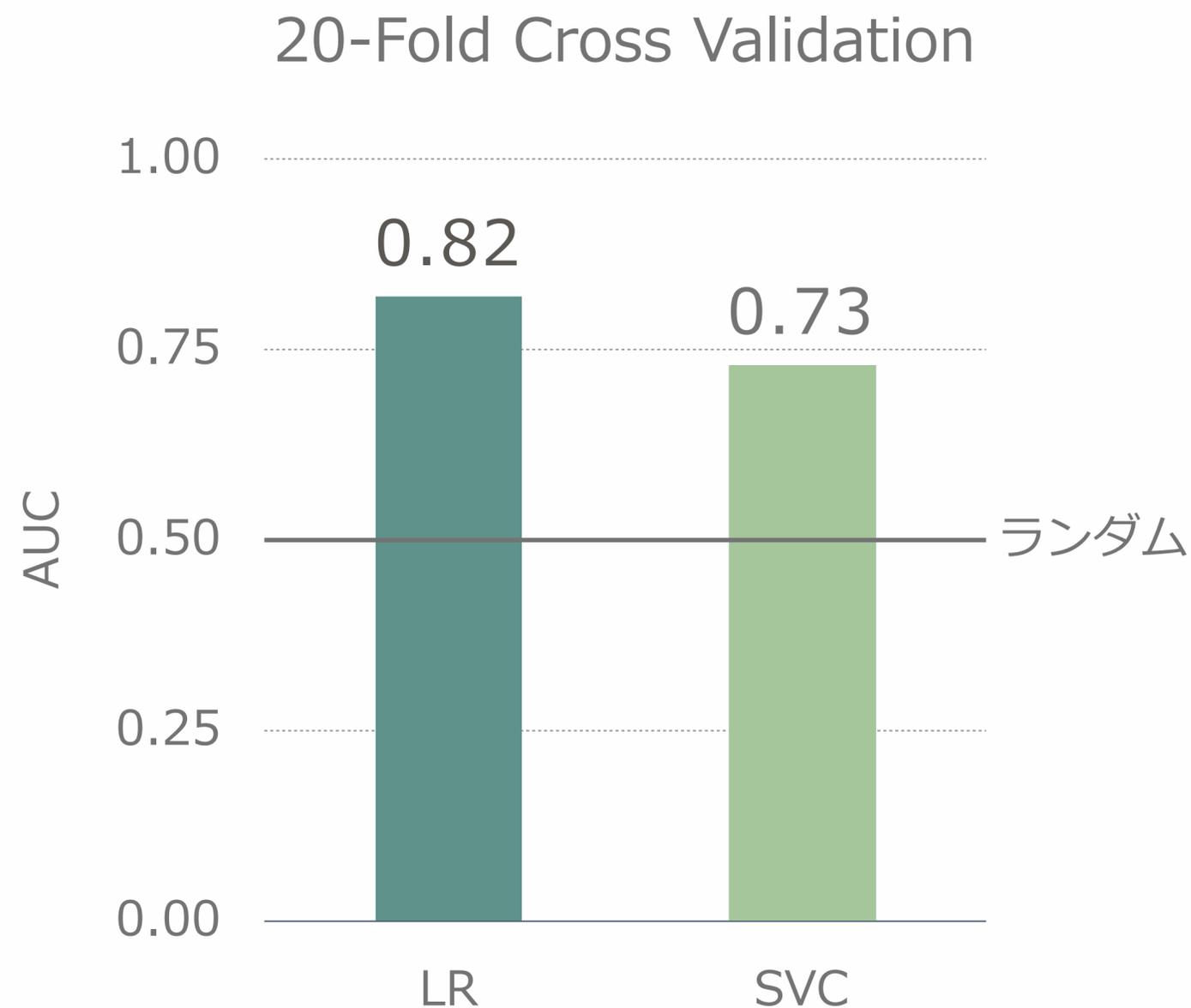
### 2.5 評価(6)

---

#### AUC (Area Under the Curve)

- $0.0 \leq \text{AUC} \leq 1.0$
- **AUC = 1.0 : 完全な分類**
- **AUC = 0.5 : ランダムな分類**
- **AUC < 0.5 : ランダムな分類より劣っている**

# 3. 演習結果



	AUC
LR	0.82
SVC	0.73

---

# 4.おわりに

---

## 今後の課題

- 画像の手作業による除去を複数人で行う
- 画像の背景, 顔の表情, 写真撮影時の年齢と現在の年齢との違い等による影響があるのか検証
- 今回使用した画像を, 人間が分類した場合の精度と比較
- 他の手法で分類した場合の精度と比較