

# CIFAR-100, MNIST, Fashion MNIST データセット

オープンデータ（人工知能関連）

URL: <https://www.kkaneko.jp/ai/data/index.html>

金子邦彦



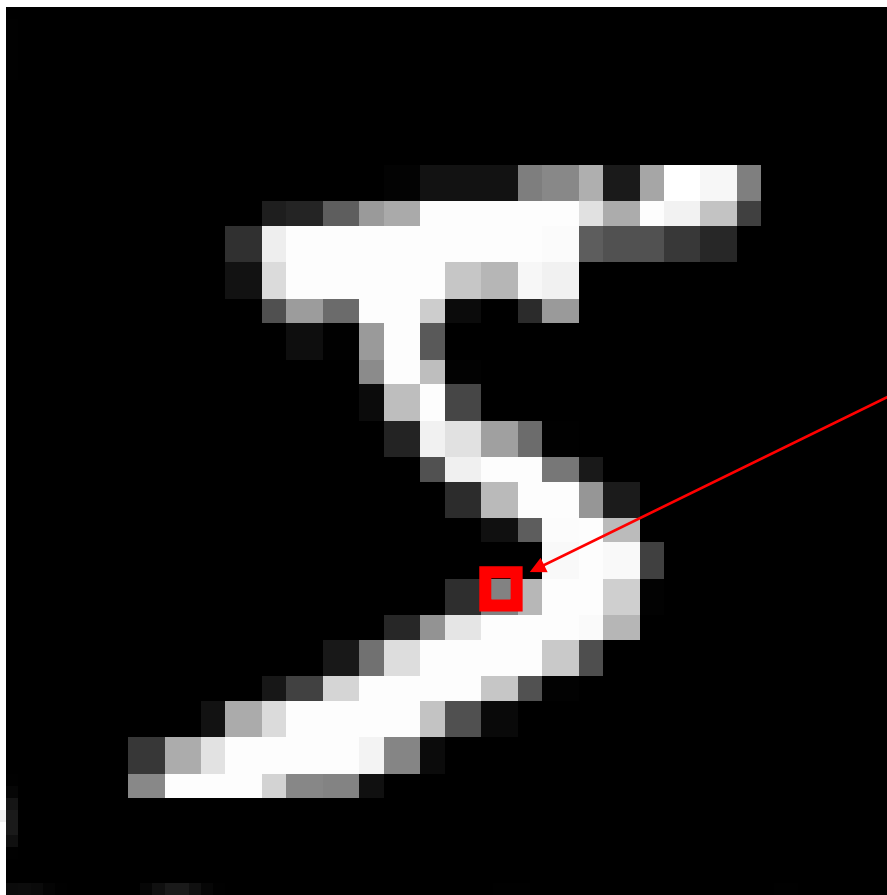
## アウトライン

- **画像と画素**
- **配列（アレイ）の形**
- Python で**配列（アレイ）の中身**を確認するには
- **Keras の小画像のデータセット**

演習の資料を準備している

<https://www.kkaneko.jp/ai/classify/tutorials.html>

# 画像と画素



**MNISTデータセット** (手書き文字のデータセットで, 濃淡画像)

画像サイズ: 28 × 28

**画素**

**画素値**



**白**

**255**



**黒**

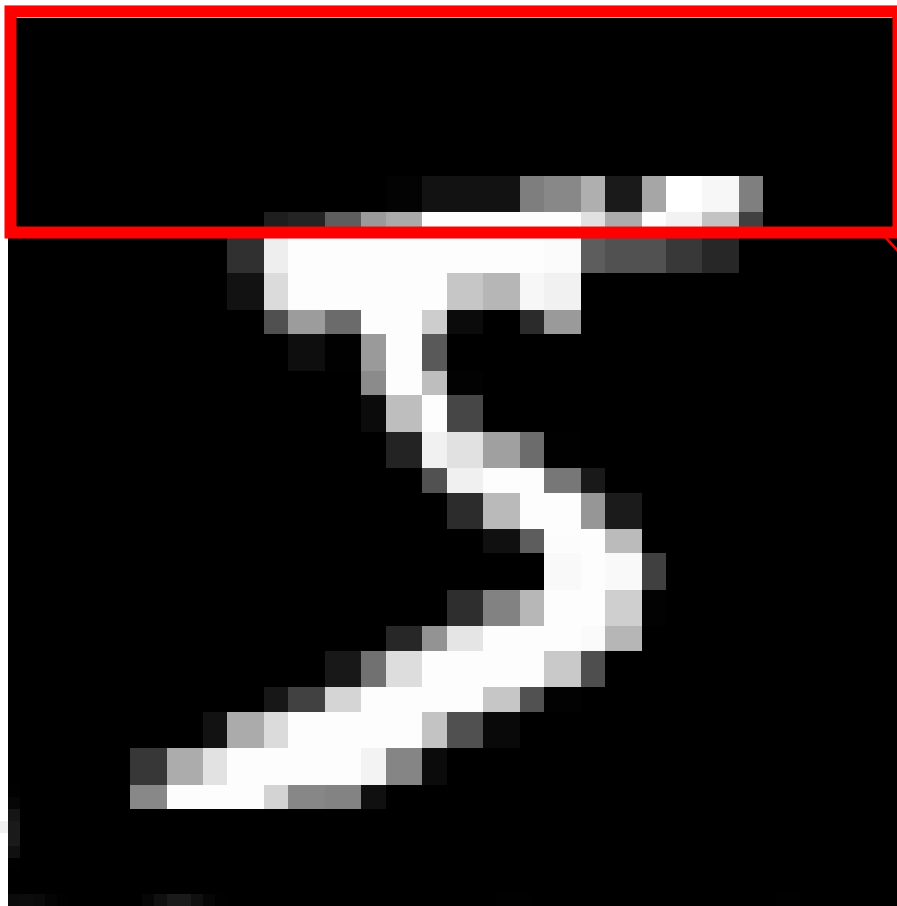
**0**

**画素値**は, **画素**の明るさに応じた **0 から 255 の数値**



# 画像と配列 (アレイ)

画像全体は、サイズ  
28 × 28 の配列  
(データの並び)



```
[ [ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  [ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  [ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  [ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  [ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 18 18 18 126 136
  175 26 166 255 247 127 0 0 0 0]
  [ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 30 36 94 154 170 253 253 253 253
  225 172 253 242 195 64 0 0 0 0]
```

MNISTデータセット (手書き文字のデータセットで、濃淡画像)

画像サイズ: 28 × 28

画像の上 7行分の画素値を表示したところ (28 × 7分)

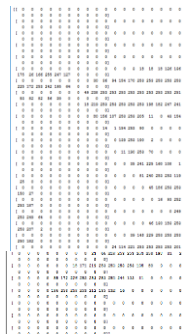
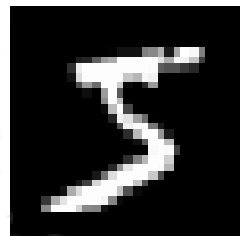
# 配列（アレイ）の形と次元



[8 5 4 1 3]

5

次元数は 1



28 × 28

次元数は 2

1枚の画像

60000枚の  
画像（同じ  
大きさ）



60000 ×  
28 × 28

次元数は 3

データ

配列（アレイ）の形

次元数

# CIFAR 100

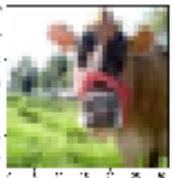


- カラー画像 **60000**枚
  - うち学習用 **50000**枚, テスト用 **10000**枚
- 学習用のカラー画像の**配列 (アレイ) の形** :  
**50000 × 32 × 32 × 3**  
枚数が **50000**枚, 高さと幅が **32×32**, RとGとB の **3成分**

```
In [80]: from keras.datasets import cifar100
...: (X_train, y_train), (X_test, y_test) = cifar100.load_data(label_mode='fine')
...: print( X_train.shape )
...: print( X_train.ndim )
...: import matplotlib.pyplot as plt
...: plt.imshow( X_train[0] )
```

```
Downloading data from https://www.cs.toronto.edu/~kriz/cifar-100-python.tar.gz
169009152/169001437 [=====] - 39s 0us/step
(50000, 32, 32, 3)
```

```
4
Out[80]: <matplotlib.image.AxesImage at 0x2e3ddcafef0>
```



- 濃淡画像 70000枚

うち学習用 60000枚, テスト用 10000枚

- 学習用の濃淡画像の配列 (アレイ) の形 :

**60000 × 28 × 28**

枚数が 60000枚, 高さと幅が 28×28

MNIST では  
0 は白, 255 は黒

```
In [87]: from keras.datasets import mnist
...: (X_train, y_train), (X_test, y_test) = mnist.load_data()
...: print( X_train.shape )
...: print( X_train.ndim )
...: import matplotlib.pyplot as plt
...: plt.imshow( X_train[0], cmap='gray' )
```

```
(60000, 28, 28)
```

```
3
```


```
Out[87]: <matplotlib.image.AxesImage at 0x2e3de3769e8>
```



# Fashion MNIST



- 濃淡画像 70000枚
  - うち学習用 60000枚, テスト用 10000枚
- 学習用の濃淡画像の配列 (アレイ) の形 :  
**60000 × 28 × 28**  
枚数が 60000枚, 高さと幅が 28×28

```
In [88]: from keras.datasets import fashion_mnist
...: (X_train, y_train), (X_test, y_test) = fashion_mnist.load_data()
...: print( X_train.shape )
...: print( X_train.ndim )
...: import matplotlib.pyplot as plt
...: plt.imshow( X_train[0], cmap='gray' )
Downloading data from http://fashion-mnist.s3-website.eu-central-1.amazonaws
32768/29515 [=====] - 1s 25us/step
Downloading data from http://fashion-mnist.s3-website.eu-central-1.amazonaws
26427392/26421880 [=====] - 11s 0us/step
Downloading data from http://fashion-mnist.s3-website.eu-central-1.amazonaws
8192/5148 [=====] - 0s 0us/step
Downloading data from http://fashion-mnist.s3-website.eu-central-1.amazonaws
4423680/4422102 [=====] - 7s 1us/step
(60000, 28, 28)
3
Out[88]: <matplotlib.image.AxesImage at 0x2e3f0574a90>

In [89]:
```