# -\*- coding: utf-8 -\*-

from sklearn import datasets

import numpy as np

# 手書き数字のデータをロードし、変数digitsに格納

digits = datasets.load\_digits()

# 特徴量のセットを変数Xに、ターゲットを変数yに格納

X = digits.data

y = digits.target

# Xとyの次元を表示。それぞれ(1797, 64)と(1797,)となる。

# サンプル数1797, 特徴量の次元64, という意味

print(X.shape)

print(y.shape)

# サンプル数と特徴量の次元の取り出し方法

(n\_samples, n\_features) = X.shape

print('サンプル数: {0}'.format(n\_samples))

print('特徴量の次元: {0}'.format(n\_features))

# クラス数の取り出し方法

n\_classes = len(np.unique(y))

print('クラス数: {0}'.format(n\_classes))

# 0～9の全ての数字に対するそれぞれのサンプル数を表示

for i in range(n\_classes):

print('{0}のサンプル数:{1}'.format(i, len(X[y==i])))