

# 総合添削問題

教師あり学習（分類）の総合添削問題はこれまでの手法の特性とハイパーパラメーターのチューニングの重要性、そしてハイパーパラメーターのサーチ方法を理解しているかを問う問題となっています。

使用するデータセットは手書きの数字の画像であり、判別しにくい数字もあることからパラメーターサーチや手法選択がより重要になります。

## 問題

- 手書きの数字の認識・分類をするための学習器をより高い精度で作成したいと思います。
  - 与えられるデータに対して手法を選び、ハイパーパラメーターを調整して学習能力の高い学習器を作ってください。
  - また、一番評価の高い手法の名前と、調整したパラメーターとその値を出力してください。
- 
- 採点基準は問題文の条件を満たしている状態で、以下の事項を考慮して総合的に採点いたします。
    - 評価値の高さ
    - パラメーターの調整方法
    - プログラムの実行時間
  - また、プログラムの実行時間は3分以内とします。

In [ ]:

```
1  # 必要なモジュールがあれば追記してください
2  from sklearn.datasets import load_digits
3  from sklearn.model_selection import train_test_split
4
5  data = load_digits()
6  train_X, test_X, train_y, test_y = train_test_split(
7      data.data, data.target, random_state=42)
8
9  # 以下にコードを記述してください
10
11
12  print("学習モデル:{}, \nパラメーター:{}".format(best_model, best_param))
13  # 最も成績のいいスコアを出力してください。
```

## ヒント

- プログラムの実行時間を計測するにはプログラムの先頭に %%time と記述してください。
  - jupyter notebook上でのみ実行できるおまじないです。Pythonのプログラムでは実行できないので注意してください。
- モデルの手法の名前をプログラムで取得するには model\_name = model.\_\_class\_\_.\_\_name\_\_ とします。
- グリッドサーチ、ランダムサーチの結果得られるパラメーターセットを取得するには best\_params = clf.best\_params\_ とします。
- 最高評価を得たモデルが複数ある場合、そのうちの一つだけ出力すれば良いです。
- 採点時、モデルの評価にはF値という値を用います。

- F値はprecision(精度)とrecall(再現率)という二つの評価値の調和平均です。

$$\frac{2(\text{precision} \times \text{recall})}{\text{precision} + \text{recall}}$$

- プログラムを作る際には `model.score(test_X, test_y)` で構いませんが、自身で評価を確認したい場合はモデルに教師データを学習させたのち以下のようにします。

```
from sklearn.metrics import f1_score
# モデルにデータを予測させる
pred_y = clf.predict(testX)

# モデルのF値を計算する
score = f1_score(test_y, pred_y, average="micro")
```

- 今回のデータセットに関する情報は[UCI Machine Learning Repository\(英語版サイト\)](http://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits) (<http://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits>)を参照してください。
- 講座でいくつかのパラメーターを手法ごとに紹介しましたが、他にも調整可能なパラメーターが存在します。  
[scikit-learnのドキュメント\(英語版サイト\)](http://scikit-learn.org/stable/modules/classes.html) (<http://scikit-learn.org/stable/modules/classes.html>)を参照してください。  
また、パラメーターに関してわからないことがあればチューターまでご質問ください。

## 解答例

添削課題の提出は以下のアドレスから提出いただきますようお願いいたします。

<https://goo.gl/forms/fW7CAspZMwHuWuqk2> (<https://goo.gl/forms/fW7CAspZMwHuWuqk2>)

以下のアドレスからアンケートにご協力頂きたく存じます。

ご回答のほど、よろしくお願いいたします。

<https://goo.gl/forms/WHjJQYeodIndRvyz2> (<https://goo.gl/forms/WHjJQYeodIndRvyz2>)