智慧型控制及應用實驗室
ICALAB

Machine Learning HW1

Advisor : Ching-Hung Lee
TA : 陳健瑋、施雨
2021/03/22

Question 2

- Library required for HW1:
  - Pandas
  - Numpy
  - Scikit-learn (sklearn)

- Example:

Import library

```
In [1]:
import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
from sklearn.metrics import  mean_squared_error
```

Define RMSE

```
In [2]:
def rmse(y_true, y_pred):
    return np.sqrt(mean_squared_error(y_true, y_pred))
```

Question 2.(b)

input: feature_0, output: label

In [3]:

```
data = pd.read_csv('dataset_v2.csv', index_col=False)        # load data

feature_0 = data.loc[:, 'feature_0'].values.reshape(-1, 1)  # select feature_0 as input
label = data.label                                           # select label as output

model = LinearRegression(n_jobs=-1, normalize=True)          # set model
model.fit(feature_0, label)                                  # fit input & output

RMSE = rmse(label, model.predict(feature_0))                 # calculate error

result = pd.DataFrame(np.zeros(119), columns=['RMSE'], index=data.columns[:-1])
result.iloc[0, 0] = RMSE
print(result)

result.to_csv('score.csv')                                   # export result to csv
```

```
                RMSE
feature_0    6.198733
feature_1    0.000000
feature_2    0.000000
feature_3    0.000000
feature_4    0.000000
...               ...
feature_114  0.000000
feature_115  0.000000
feature_116  0.000000
feature_117  0.000000
feature_118  0.000000

[119 rows x 1 columns]
```

Question 2.(b)

- After calculating the RMSE of all features, sort the features according to the magnitude of the RMSE from low to high.

越有用的放越上面

Finally, export the result to .csv

	A	B
1		RMSE
2	feature_0	6.1987327
3	feature_1	0
4	feature_2	0
5	feature_3	0
6	feature_4	0
7	feature_5	0
8	feature_6	0
9	feature_7	0
10	feature_8	0
11	feature_9	0
12	feature_10	0
13	feature_11	0
14	feature_12	0
15	feature_13	0
16	feature_14	0

Low

High

Feature name

Submission Format

- Upload .rar (or .zip) include:
  - Q.1 and Q.2 (d) answer in PDF
  - Q.2 code (.py or .ipynb)
  - Q.2 (a)(b)(c)(e) output (.csv)