

Day 42

機器學習

決策樹 - 程式碼撰寫



使用 Sklearn 建立決策樹模型

根據回歸/分類問題分別建立不同的 Classifier

- `from sklearn.tree_model import DecisionTreeRegressor`
- `from sklearn.tree_model import DecisionTreeClassifier`
- `clf = DecisionTreeClassifier()`

決策樹的超參數

- Criterion: 衡量資料相似程度的 metric
- Max_depth: 樹能生長的最深限制
- Min_samples_split: 至少要多少樣本以上才進行切分
- Min_samples_leaf: 最終的葉子 (節點) 上至少要有多少樣本

```
from sklearn.tree import DecisionTreeClassifier

clf = DecisionTreeClassifier(
    criterion = 'gini',
    max_depth = None,
    min_samples_split = 2,
    min_samples_leaf = 1,
)

# feature importance
clf.feature_importances_
```


解題時間 It's Your Turn

請跳出PDF至官網Sample Code & 作業
開始解題

