

D42 : tree based model - 決策樹程式碼撰寫



簡報閱讀



範例與作業



問題討論

決策樹 - 程式碼撰寫
Coding 練習日

決策樹 - 程式碼撰寫 Coding 練習日

知識地圖



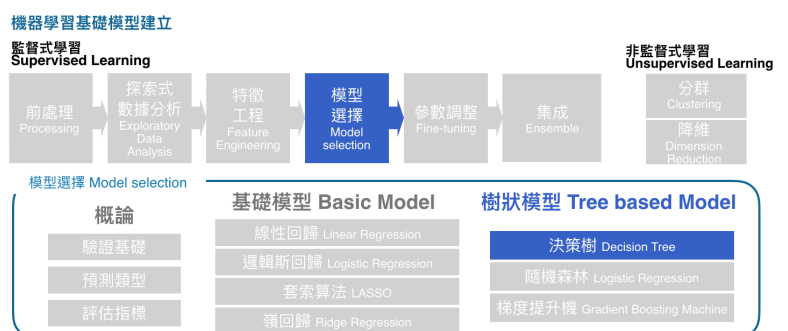
本日知識點目標

使用 Sklearn 建立決策樹
模型

決策樹的超參數

知識地圖

延伸閱讀



推薦延伸閱讀

解題時間

本日知識點目標

本日知識點目標

- 了解決策樹的程式碼應用
- 如何用 Scikit-learn 撰寫決策樹的程式碼
- 使用內建的特徵重要性進行特徵選取

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,
)

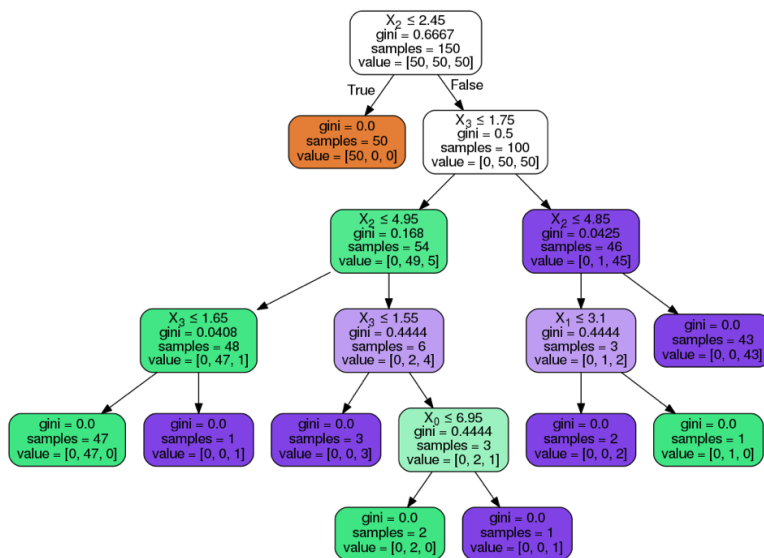
# feature importance
clf.feature_importances_
```



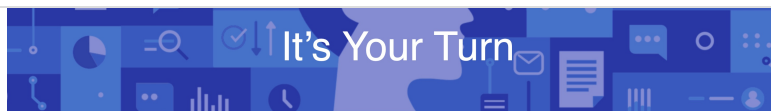
- 除了每日知識點的基礎之外，推薦的延伸閱讀能補足學員們對該知識點的了解程度
- 建議您解完每日題目後，若有多餘時間，可再補充延伸閱讀文章內容

推薦延伸閱讀

- 可安裝額外的套件 graphviz，畫出決策樹的圖形幫助理解模型分類的準則
- Creating and Visualizing Decision Trees with Python [網頁連結 \(英文\)](#)



解題時間



Sample Code & 作業
開始解題



[下一步：閱讀範例與完成作業](#)

