目標: 根據葡萄酒的化學特性將葡萄酒分類成三種不一樣的葡萄酒類型

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

(將資料檔讀入並設立每個欄位的名稱

第一欄為預測的類別

後十三欄為特徵)

特徵介紹:

Alcohol（酒精）： 葡萄酒中的酒精含量。

Malic acid（蘋果酸）： 葡萄酒中的蘋果酸含量。

Ash（灰分）： 葡萄酒中的灰分含量。

Alcalinity of ash（灰的鹼性）： 葡萄酒中灰的鹼性。

Magnesium（鎂）： 葡萄酒中的鎂含量。

Total phenols（總酚）： 葡萄酒中的總酚含量。

Flavanoids（黃酮）： 葡萄酒中的黃酮類化合物含量。

Nonflavanoid phenols（非黃酮酚）： 葡萄酒中的非黃酮類酚含量。

Proanthocyanins（原花青素）： 葡萄酒中的原花青素含量。

Color intensity（顏色強度）： 葡萄酒的顏色強度。

Hue（色調）： 葡萄酒的色調。

OD280/OD315 of diluted wines（稀釋葡萄酒的 OD280/OD315）： 透過光學密度測量的葡萄酒的稀釋度。

Proline（脯氨酸）： 葡萄酒中的脯氨酸含量。

一張含有 文字, 字型, 螢幕擷取畫面, 圖形 的圖片

自動產生的描述

一張含有 螢幕擷取畫面, 對稱, 建築 的圖片

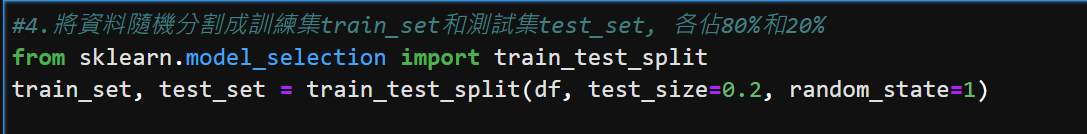
自動產生的描述

(使用.head(20)印出資料集中的前20筆資料)

一張含有 文字, 螢幕擷取畫面, 字型, 行 的圖片

自動產生的描述

(印出資料集的大小，得知一共有178筆資料，以及14個欄位)



(將資料進行分割，train\_data佔80% 而 test\_data佔20%)

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

(提取出train\_data和test\_data各自的特徵和標籤)

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

(匯入不同機器學習分類器的函式庫)

一張含有 文字, 螢幕擷取畫面 的圖片

自動產生的描述

(首先先建立一個包含不同機器學習分類器的列表，然後以迴圈的方式訓練每個分類器(.fit)並印出它們各自訓練集的辨識率和測試集的辨識率)

最後建立了一個results\_df來存取每個分類器的辨識率方便於後面視覺化

ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.

我們發現在我們使用MLPClassifier()和LogisticRegression時因為迭代次數太少

所以導致優化過程尚未達到收斂，因此我們可能需要增加他它們的max\_iter

一張含有 文字, 螢幕擷取畫面, 軟體, 字型 的圖片

自動產生的描述

(印出results\_df)

一張含有 螢幕擷取畫面, 文字, 字型 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 平行, 行 的圖片

自動產生的描述

(將每個分類器訓練集的辨識率和測試集的辨識率進行比較)

從結果可以看出，不同的分類器在處理同一個資料集時的表現會有所差異

我們必須要依據不同的情況選擇表現相對較好的分類器

又或說可能會需要對一些辨識率較低的模型進行進一步參數的調整來達到更好的訓練效果