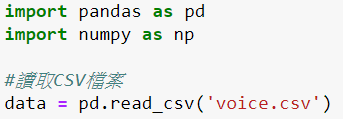
**2019 ECT作業五**

**1. 請用python依照步驟對voice.csv 進行SVM分析，過程中對所有重要程式步驟進行截圖並加以說明，越詳盡越好。**

首先，引用所需套件，並讀取csv檔，



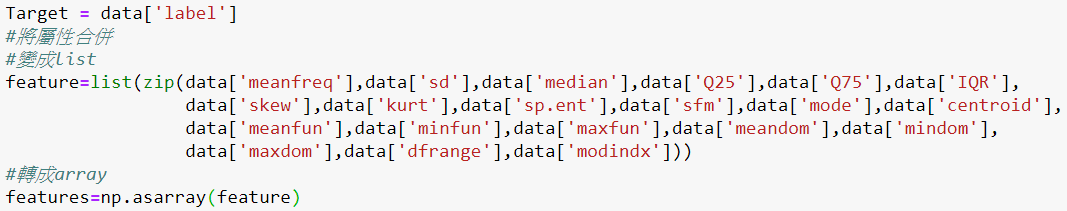
**使用 Python**：(60%)

**(a) 請檢查資料集是否有空值，如有空值即去掉該筆資料**

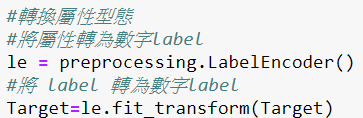
運用 dropna()函數去掉資料的空值



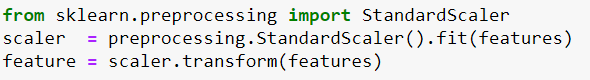
**(b) 將最後一個屬性值”label”切分為Target，其餘屬型切分為Feature**



**(c) 將Target進行encoded，用LabelEncoder將male轉為1，female轉為0**



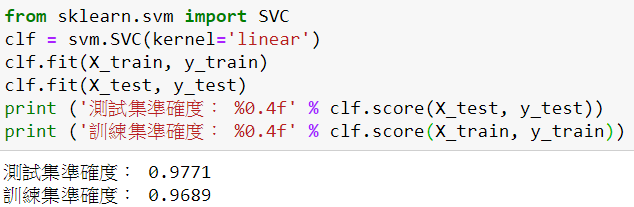
**(d) 將Feature用sklearn.preprocessing 的StandardScaler進行標準化**



**(e) 切分資料集與測試集，設test\_size=0.33，random\_state=1**



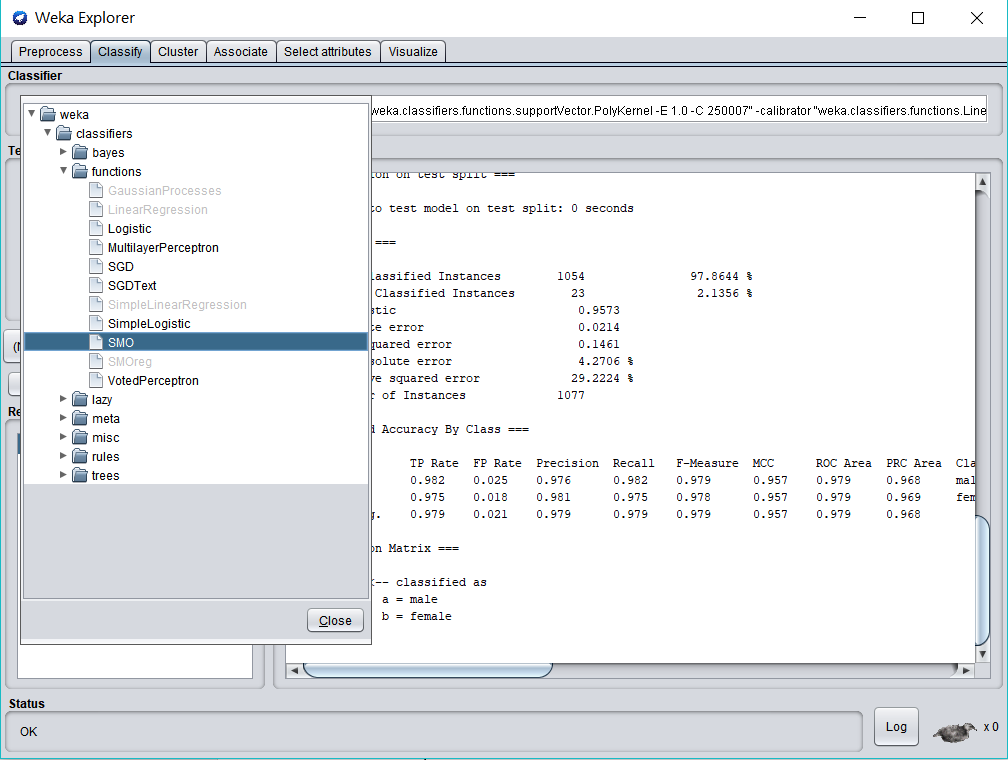
**(f) 最後，使用sklearn.svm裡的SVC進行分析，kernel設為' linear '，並印出模型最終的準確度**



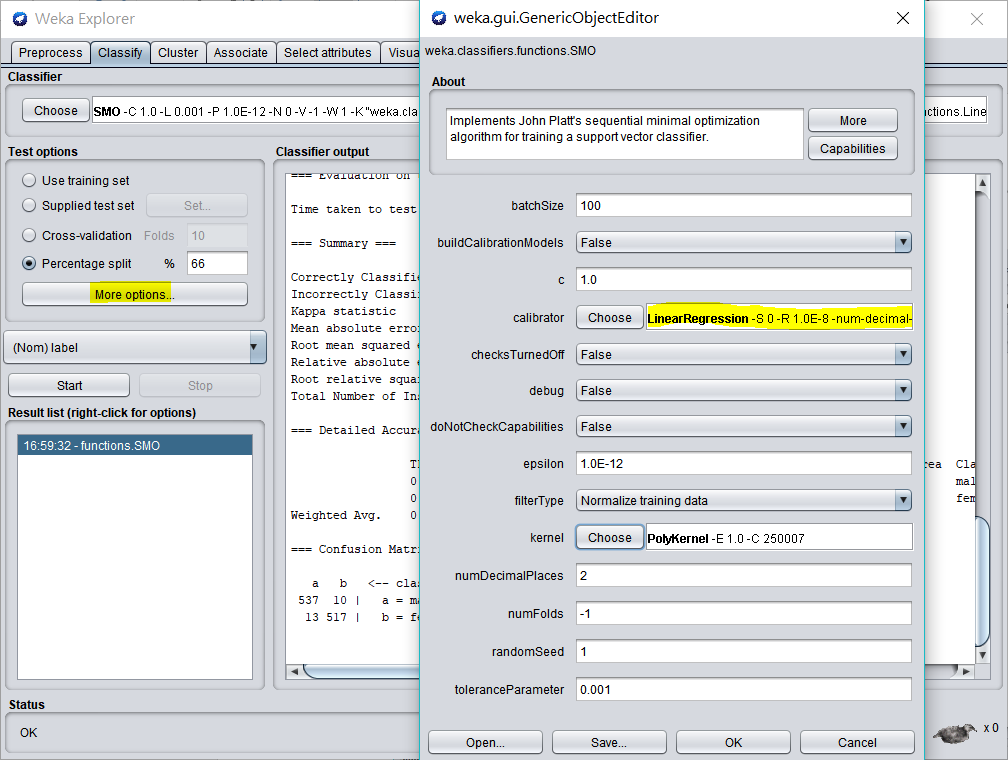
**使用 Weka 軟體：(40%)**

**(a) 使用weka裡的SMO function對voice.csv進行SVM分析，kernel設為' linear '，Percentage spilt設為66%，截圖並附上過程及準確率**

1. 首先在Weka中開啟voice.csv。
2. 在Classify面板中，在Classifier選擇「weka / classifiers / functions / SMO」。



1. 在「More options」中的「calibrator」選擇「LinearRegression」。



1. 在「Test options」中選取「Percentage split」，並設定為66%；選擇預測「(Nom)label」，並點選「Start」。

可得準確度為 97.8664%

