

有使用 Chatgpt、gemini 生成 code

1.(a)在 code 裡

1.(b)

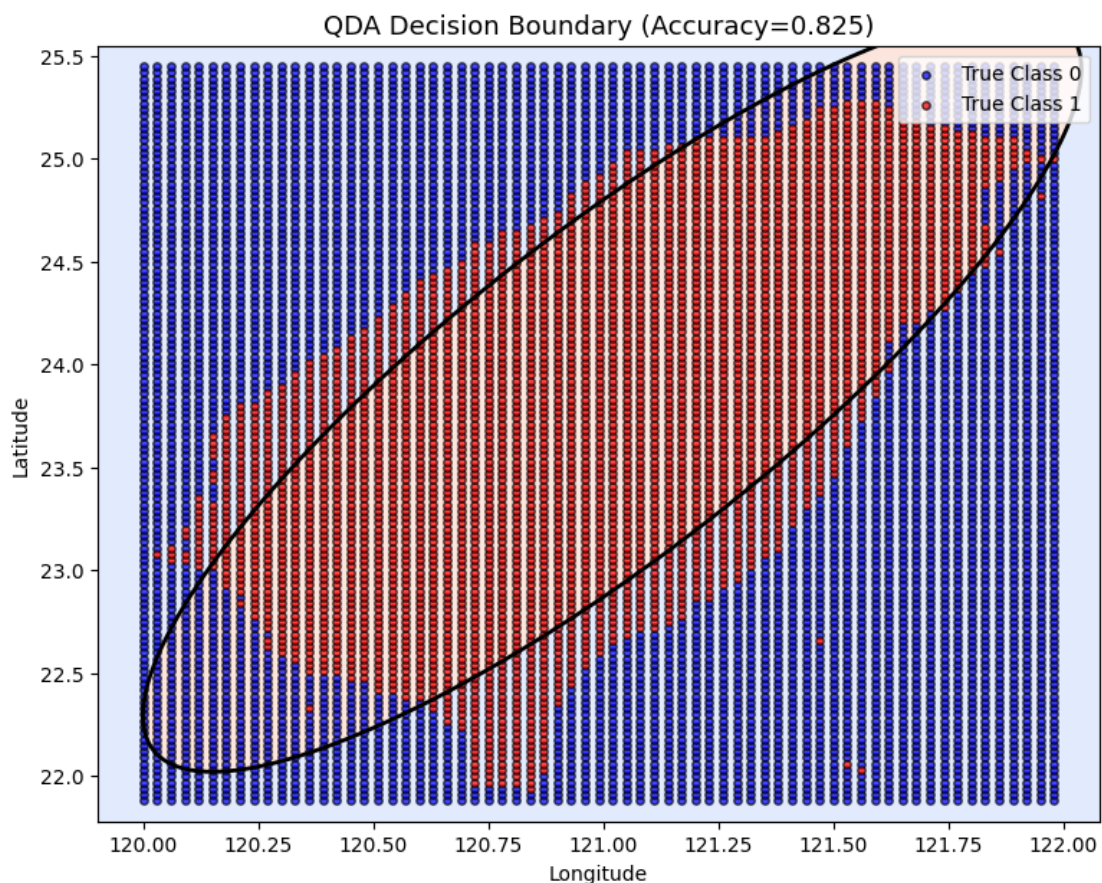
首先假設每個類別都是常態分佈，藉由上次作業找出的常態分佈的 **maximum likelihood estimators**，我們可以用樣本估計這些常態分佈的參數，從機器學習的角度來看，樣本就是訓練集，而估計的參數即是模型參數。接著我們就可以用這個參數的常態分佈算出某一資料點 X 屬於這個分佈的機率，之後預測 X 是屬於”機率最大的那個分佈的類別”。

這次的問題是二維平面上的分類問題，正確的 **decision boundary** 大致是一條簡單的封閉曲線(忽略外島的話)，由於二維的 **GDA** 模型可以跑出橢圓或是雙曲線的 **decision boundary**，所以還算是勉強強強能辦到。

1.(c)

QDA classification accuracy = 0.8252 (on test set)

1.(d)



2.(a)在 code 裡

2.(b)

取出 5 筆資料來看

Combined model sample predictions:

Input: [120. 21.88], C(x)=0, R(x)=33.94, h(x)=-999.00

Input: [120.03 21.88], C(x)=0, R(x)=33.78, h(x)=-999.00

Input: [120.06 21.88], C(x)=0, R(x)=33.62, h(x)=-999.00

Input: [120.09 21.88], C(x)=0, R(x)=33.48, h(x)=-999.00

Input: [120.12 21.88], C(x)=0, R(x)=33.39, h(x)=-999.00

當 C(x)是 0，h(x)都是-999

2.(c)

定義一個 combined_model(x)，把 clf_model(x)的值根據>0.5 與否替換成 1 或 0，再根據 clf_model(x)的值決定 h(x)要輸出 reg_model(x)還是-999

2.(d)

