GDA運作原理

高斯判別分析 (Gaussian Discriminant Analysis, GDA) 是一種生成式學習演算法 (Generative Learning Algorithm)，主要用於解決分類問題。其核心思想是為每一個類別的資料分佈建立一個機率模型，而非直接學習類別之間的決策邊界。

GDA 的運作基於以下關鍵假設與步驟：

1. **核心假設**：GDA 假設在給定類別 $y=k$ 的條件下，特徵向量 $\vec{x}$ 的條件機率分佈 $p(\vec{x}|y=k)$ 服從一個多變量高斯分佈 (Multivariate Gaussian Distribution)。其數學表達式為：

   $p(\vec{x}|y=k) \sim \mathcal{N}(\vec{\mu}_k, \Sigma_k)$

   其中 $\vec{\mu}_k$ 是類別 k 的平均向量， Σ_k 是其協方差矩陣。在本作業中，我們採用的是 GDA 的一個特例，即線性判別分析 (Linear Discriminant Analysis, LDA)，它進一步假設所有類別共用同一個協方差矩陣 Σ。

2. **模型學習 (訓練階段)**：訓練 GDA 模型的目標是從訓練資料中估計出模型的參數。對於二元分類問題 (y=0, 1)，需要估計的參數包括：
   - **類別先驗機率 (Prior Probability)**: $\phi = p(y=1)$，通常由類別 1 樣本數在總樣本中的比例來估計。則 $p(y=0) = 1-\phi$。
   - **類別平均向量 (Mean Vectors)**: $\vec{\mu}_0$ 和 $\vec{\mu}_1$，分別透過計算屬於類別 0 和類別 1 的所有樣本特徵的平均值得到。
   - **共用協方差矩陣 (Shared Covariance Matrix)**: Σ，透過計算所有樣本點相對於其所屬類別均值的協方差，並進行加權平均來估計。

3. **預測階段**：對於一個新的輸入樣本 x_{new}，GDA 利用貝氏定理 (Bayes' Theorem) 計算其屬於每一個類別的後驗機率 (Posterior Probability)：

   $$p(y=k|\vec{x}_{new}) = \frac{p(\vec{x}_{new}|y=k)p(y=k)}{p(\vec{x}_{new})}$$

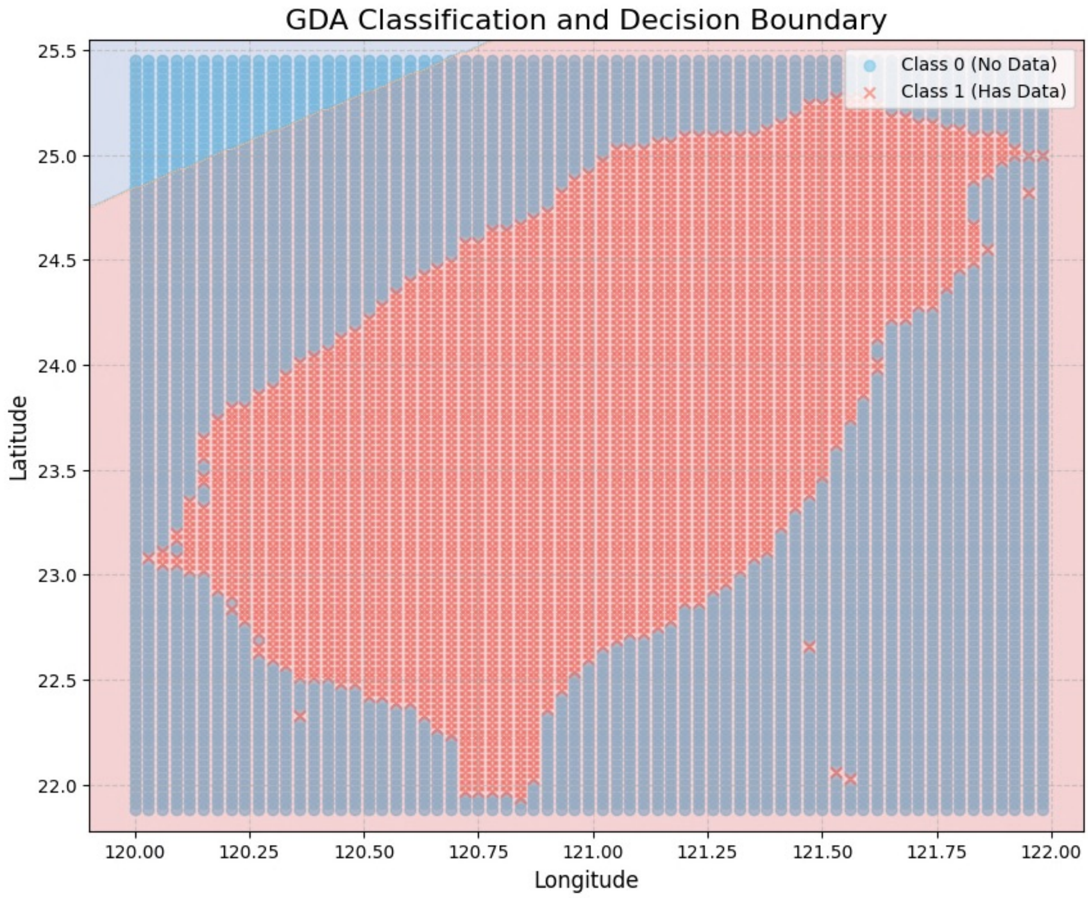
   模型最終會將 $\vec{x}_{new}$ 預測為後驗機率最大的那個類別。由於分母 $p(\vec{x}_{new})$ 對所有類別是常數，因此預測規則簡化為最大化分子 $p(\vec{x}_{new}|y=k)p(y=k)$。在 LDA 的假設下，這會產生一個線性的決策邊界。

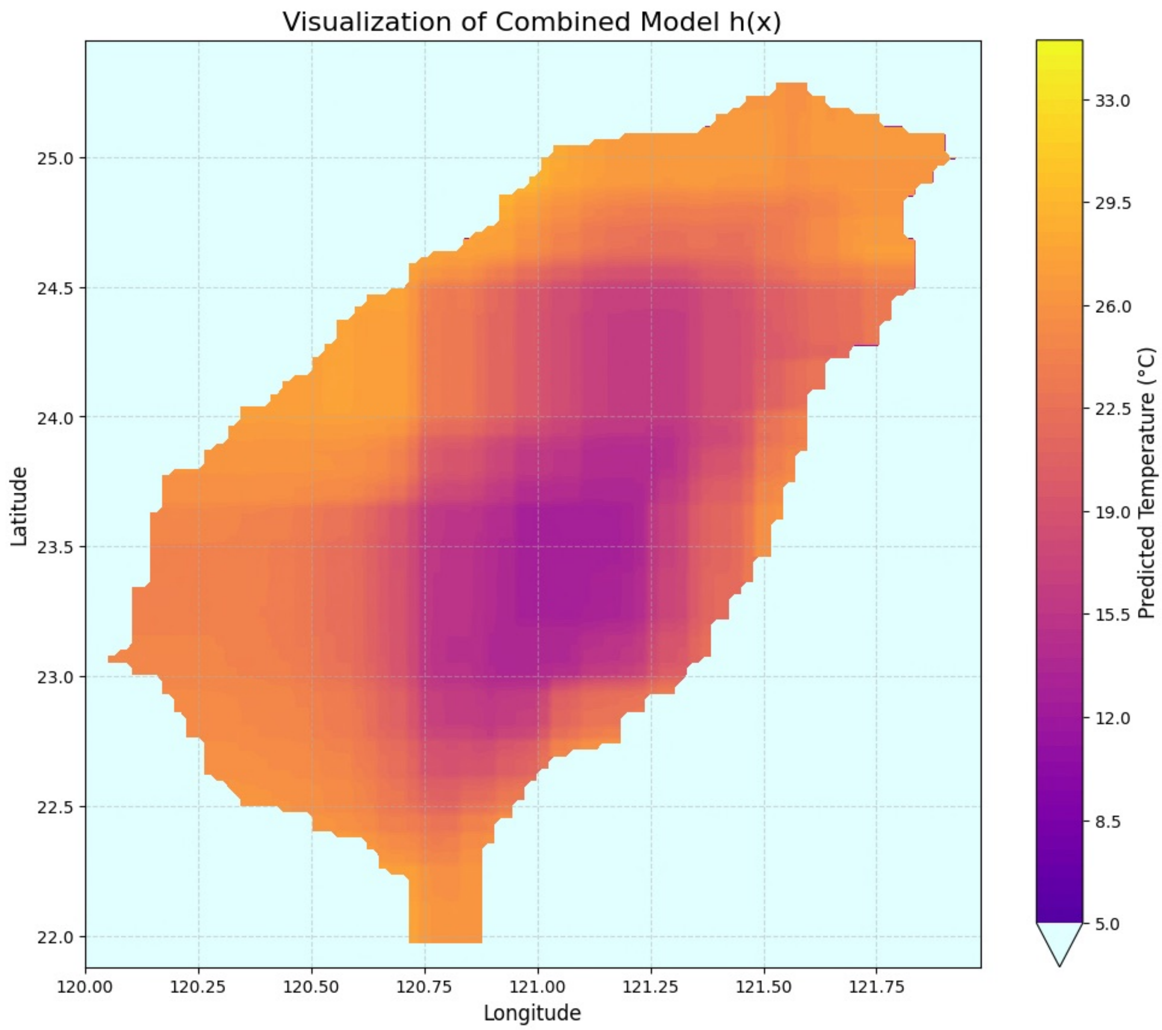
由於這個資料集（台灣本島）的資料點是類似一個橢圓的群集，所以GDA模型是可行的，準確性不至於太低。

建構h(x)

組合函數的建構是基於一個條件化執行 (Conditional Execution) 的邏輯，其程式實現遵循以下步驟：

1. **建立函式框架**：首先，定義一個函式 predict_h_model，該函式接收輸入特徵 X、已訓練的分類模型 C_model 和迴歸模型 R_model 作為參數。
2. **執行分類作為閘門機制 (Gating Mechanism)**：對於所有輸入的特徵向量，首先通過分類模型 C_model 進行預測。此步驟的輸出是一個標籤陣列 (包含 0 和 1)，它將作為後續操作的判斷依據。
3. **初始化預設輸出**：建立一個與輸入樣本數相同大小的輸出陣列，並將所有元素初始化為 -999。此操作確保了所有被分類器判定為 0 的樣本，其最終輸出值都符合函數定義的第二個條件。
4. **選擇性執行迴歸**：利用上一步驟生成的標籤陣列，建立一個布林遮罩 (Boolean Mask) 以篩選出所有被判定為 1 的樣本。接著，僅將這部分樣本的特徵傳遞給迴歸模型 R_model 進行預測，從而得到對應的溫度值。
5. **整合預測結果**：最後，將迴歸模型產生的溫度預測結果，根據布林遮罩更新至初始化後的輸出陣列的相應位置。


GDA Classification and Decision Boundary
Class 0 (No Data)
Class 1 (Has Data)
Latitude
Longitude
25.5
25.0
24.5
24.0
23.5
23.0
22.5
22.0
120.00
120.25
120.50
120.75
121.00
121.25
121.50
121.75
122.00


Visualization of Combined Model h(x)
Latitude
Longitude
Predicted Temperature (°C)
25.0
24.5
24.0
23.5
23.0
22.5
22.0
120.00
120.25
120.50
120.75
121.00
121.25
121.50
121.75
33.0
29.5
26.0
22.5
19.0
15.5
12.0
8.5
5.0