

# 氣溫網格預測任務報告

## 一、研究目的與資料來源

本專案旨在處理並預測中央氣象署 (CWA) 所提供的氣溫網格資料，原始資料來自 XML 檔案 "O-A0038-003.xml"，內容為臺灣區域劃分為  $120 \times 67$  的網格，每個網格點包含一個氣溫值（若無資料則為 -999.0），並附有左下角座標與解析度資訊（每格 0.03 度）。

資料中同時含有：

- 有效資料點 (valid)：數值  $\neq$  -999.0
- 無效資料點 (invalid)：數值為 -999.0，視為缺漏

本實驗目標為訓練模型  $h(x)$ ，對每個位置預測其是否具有有效資料，若為有效則進一步預測其氣溫值。

## 二、資料前處理與張量構建

首先解析 XML 檔案並建構氣溫網格，再依據是否為有效資料構建：

- 二元標籤矩陣 (1=valid, 0=invalid)
- 將 -999.0 轉為 `np.nan` 以利視覺化

接著將經緯度轉為標準化座標 (0, 1)，建構訓練資料

## 三、混合模型架構設計 $h(x)$

設計一個混合神經網路  $h(x)$ ，結合分類器與回歸器：

(1) 分類器 Classifier：

- 輸入經緯度 (lon, lat) → 輸出預測為 valid 的機率
- 網路為多層感知機 (MLP) , 隱藏層配置為 [128 → 64 → 32 → 8]
- 輸出值經 sigmoid 解釋為 valid 機率 , 門檻設為 0.5

(2) 回歸器 Regressor :

- 輸入經緯度 → 預測氣溫值
- 使用 Tanh 啟動函數 , 損失函數為 MSELoss ( 均方誤差 )

(3) 混合模型 HxModel :

- 當 Classifier(x) > 0.5 , 回傳 Regressor(x)
- 否則輸出 -999.0 ( 代表無效值 )

## 四、GDA 模型設計與橢圓視覺化

設計 GDA ( Gaussian Discriminant Analysis ) 模型來建模 valid 資料的分布 :

- 對所有 valid 資料計算平均  $\mu$  與共變異數  $\Sigma$
- 使用  $\chi^2$  分佈 ( 自由度 2 , 95% 信賴區間 ) 繪製高斯橢圓

由於 GDA 在標準化空間計算 , 需經共變異數轉換後將橢圓畫至原始經緯度上 , 並正確計算主軸與角度 , 使用 Matplotlib 的 Ellipse 繪圖。

## 五、訓練與結果分析

(1) 回歸器訓練結果 :

模型能有效收斂 , 預測值穩定且不過擬合。

(2) 混合模型 h(x) :

- 成功在 valid 區域內預測合理氣溫，在 invalid 區域保持為 -999。
- 分類與回歸整合效果良好。

(3) GDA 橢圓可合理包覆有效資料，提供良好統計視覺輔助。

## 六、結論與心得

輸出結果成功整合資料解析、分類與回歸模型設計、混合預測架構與 GDA 統計建模，並有效進行氣溫網格預測任務。