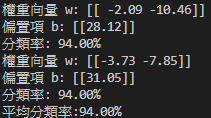
機器學習HW2

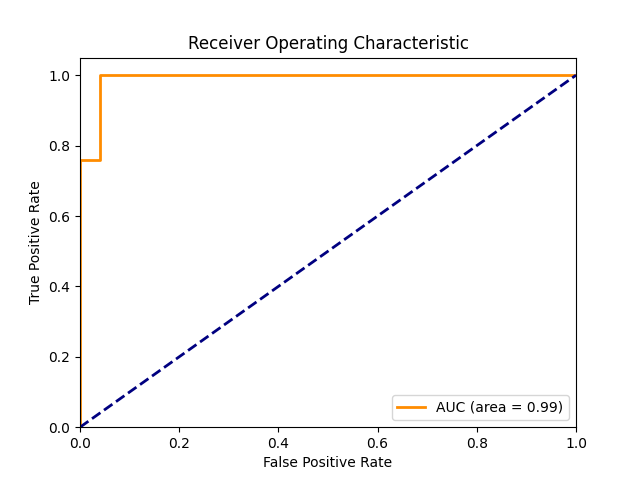
**Part2**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **W**T | b | CR |
| 第一次分類 | [-2.09, -10.46] | 28.12 | 94.00% |
| 第二次分類 | [-3.73, -7.85] | 31.05 | 94.00% |

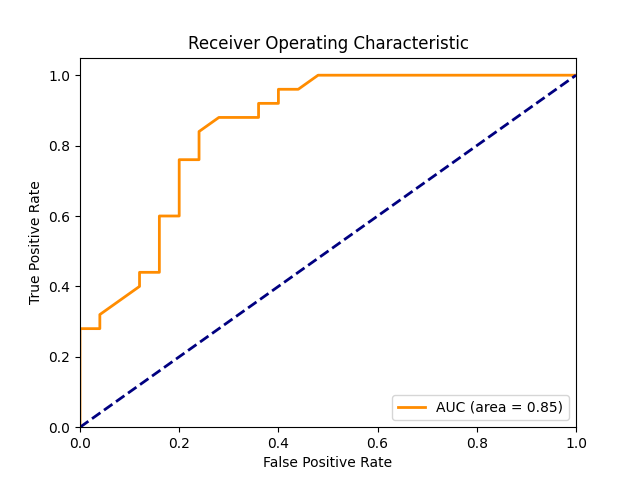
Average CR = 94.00%



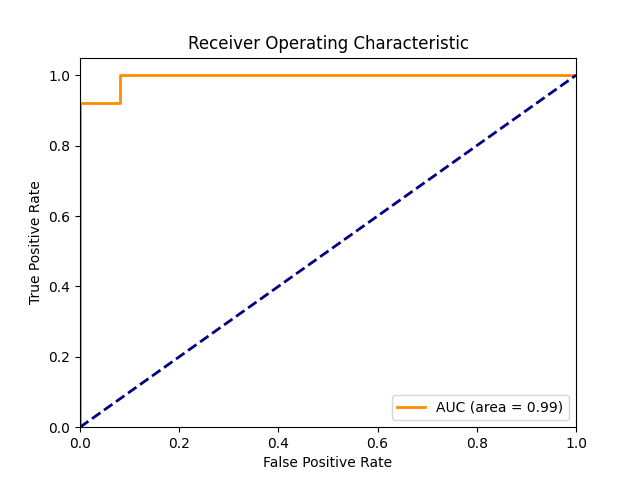
**Part3**



圖(一): 使用特徵1~4的ROC曲線



圖(二): 使用特徵1~2的ROC曲線



圖(三): 使用特徵3~4的ROC曲線

|  |  |  |
| --- | --- | --- |
|  | CR | AUC |
| 使用特徵1~4 | 97.00% | 0.99 |
| 使用特徵1~2 | 72.00% | 0.85 |
| 使用特徵3~4 | 94.00% | 0.99 |

ROC曲線代表模型在不同分類閾值下的分類表現，展示了TPR與FPR之間的權衡。我們可以通過分析這條曲線來選擇最合適的閾值，並根據應用需求中的正負類分類成本比率（Cp/Cn），來優化模型的性能。從圖中也能看出橘線下包含的面積越大AUC越大。

從圖(一)和圖(三)可以看到，FPR接近0，TPR接近1，這說明模型幾乎沒有將負例錯誤分類為正例的情況，同時能夠識別出大部分正例。然而，圖(二)的ROC曲線顯示出明顯的階梯形狀，反映了模型的分類效果有波動。這表示在不同的閾值下，FPR和TPR之間存在較大的折衷，導致模型在一些閾值下會產生一定的誤判和準確率的變化。這些差異也反映在它們的AUC值上，圖(一)和圖(三)的AUC值較高，說明分類效果接近完美，而圖(二)的AUC為0.85，相對較低，說明該模型的分類性能雖然不錯，但仍然存在改進的空間，可能在某些情況下更容易將負例錯誤分類為正例。

他們的CR分別為97%、72%和94%，這與他們的AUC高度相關。因此，AUC是衡量模型整體分類能力的重要指標，而CR則反映了模型在特定閾值下的表現。將兩者結合起來，可以更全面地評估模型的分類性能。

**Part4**

CR1 = 96.00%

CR2 = 96.00%

Average CR = 96.00%



**Part5**

本週作業使用LDA分類器的結果顯示，使用特徵3和4進行分類能夠達到較高的分類準確率（CR），這與上週的散佈圖（scatter plot）結果高度一致。下圖為特徵3和4的散佈圖，顯示了Class 2和Class 3之間的明顯區分。在Part 2中，針對Class 2和Class 3的分類，LDA分類器達到了94%的準確率。當在Part 4中加入Class 1後，分類準確率進一步提升，顯示出LDA分類器在多類別情境下依然具備優異的分類能力

從ROC曲線與AUC的角度來看，使用特徵3和4時的AUC值較高，顯示這兩個特徵在不同閾值下均具備良好的區分能力，這與分類結果相符。而且即使LDA只能進行二元分類，我們仍然可以通過OAO方法（One against One）來應對多類別問題。

