

*入場通知書編號：_____

注意：①作答前先檢查答案卷，測驗入場通知書編號、座位標籤、應試科目等是否相符，如有不同應立即請監試人員處理。使用非本人答案卷作答者，該節不予計分。
②本試卷為一張單面，非選擇題共 4 大題，每題 25 分，共 100 分。
③非選擇題限以藍、黑色鋼筆或原子筆於答案卷上採橫式作答，並請依標題指示之題號於各題指定作答區內作答。
④請勿於答案卷上書寫姓名、入場通知書編號或與答案無關之任何文字或符號。
⑤答案卷務必繳回，未繳回者該節以零分計算。

第一題：

有關機器學習中監督式學習常應用在金融相關的服務，請回答下列問題：

- (一) 請列舉五個監督式學習的演算法名稱，例：決策樹。【5 分】
- (二) 何謂 k 折交叉驗證(k-fold cross-validation)? 【6 分】
- (三) 混淆矩陣(confusion matrix)是用來評估監督式學習的模型效果，請用表格方式解釋並定義何謂混淆矩陣(confusion matrix)? 【8 分】
- (四) 請根據第(三)小題的混淆矩陣分別寫出精準率(precision)、召回率(recall)、正確率(Accuracy)的公式。【6 分】

第二題：

請回答下列問題：

- (一) 金融業需要維護顧客關係，希望建立機器學習模型找出客戶流失的原因，請提出一種機器學習的方法，並說明理由跟做法。【10 分】
- (二) 深度學習技術近年蓬勃發展，請比較深度學習與類神經網路的差異。【5 分】
- (三) 何謂卷積神經網路(Convolutional Neural Network, CNN)? 常用在哪種應用? 【10 分】

第三題：

「插入排序法(Insertion Sort)」的執行方式描述如下：在第 i 回合執行前，陣列的前 i 個數值已經從小排到大。在第 i 回合執行中，陣列的第(i+1)個值，會與前 i 個排序過後的值比較大小，然後放入適當的位置，以便將陣列前(i+1)個值從小排到大。如此完成第 i 回合的計算。請根據上列敘述，回答下列問題：

- (一) 假設給定的輸入陣列長度為 100，「插入排序法(Insertion Sort)」需執行幾回合，才能完成排序? 【5 分】(請寫出解答說明，否則不予計分)
- (二) $B = \{ b[0], b[1], b[2], b[3], b[4], b[5] \}$ 是一長度為 6 的整數陣列，內含不同整數。請針對陣列 B 執行「插入排序法(Insertion Sort)」，將 B 中的整數，由小到大排序。請依照【表三】之格式，將執行結果依序列出。【20 分】(未依格式寫出，不予計分)

| | b[0] | b[1] | b[2] | b[3] | b[4] | b[5] |
|--------|------|------|------|------|------|------|
| 原始輸入 | 13 | 7 | 86 | 5 | 21 | 15 |
| 第 1 回合 | | | | | | |
| 第 2 回合 | | | | | | |
| | | | | | | |

【表三】

第四題：

有一文字檔由 8 個字母 a,b,c,d,e,f,g,h 組成，已知在該文字檔中，a 出現 6 次，b 出現 10 次，c 出現 2 次，d 出現 7 次，e 出現 3 次，f 出現 8 次，g 出現 12 次，h 出現 14 次。小明想要將這 8 種字母進行二元編碼，請回答下列問題：

- (一) 小明要為每一個字母進行編碼時，要求編碼後字母的編碼長度要一樣，而且要能正確解碼與編碼後的文字檔的位元數為最少，請問每個字母經此固定長度二元編碼後，長度為何? 編碼後的文字檔總位元數為多少? 【10 分】(請寫出解答說明，否則不予計分)
- (二) 請利用霍夫曼編碼法(Huffman Coding)進行上述文字檔的不固定位元長度之二元編碼，請將其霍夫曼編碼樹畫出。【15 分】