

期末小專題

說明：

1. 因應 COVID-19 實施遠距教學，本學期期末考取消，改以小專題方式綜合演練本學期所學內容。上傳截止時限是 **6/25(週六)24:00**，遲交恕不受理！
2. 小專題成績佔期末總成績的 20%，本學期作業與出席狀況調整為佔總成績的 50%，其餘部分的比重維持不變。
3. 每位同學可在下方題目中挑選一題作為小專題題目。評分項目如下：
 - 程式是否符合題目要求
 - 程式完整度
 - 是否提供註解增加可讀性
 - 程式架構是否充分運用本學期介紹的物件導向概念
 - 有無防呆設計 / 使用者操作的流程是否合理
 - 其他額外加分項目 (延伸超越原本題目的功能、完成不只一題或其他)
4. 程式結構與同學高度雷同者，不論原作是誰，滿分將以 60 分起評。

題目：

1. 請利用附件資料包中的「學生身高平均值(6 歲-15 歲).csv」檔案作為匯入資料集，設計一個身高預測系統滿足以下要求：
 - 將資料依據課堂範例，將所有資料藉由動態陣列讀入後存放在小數陣列。請注意資料集第一列是表頭要拿掉，另外資料集中的欄位分隔記號是逗號。
 - 使用者由鍵盤輸入年齡及性別，由程式找出符合該年齡與性別的 97 到 109 學年度身高資料。然後以這些資料計算其簡單線性迴歸模型。最後輸出這個迴歸方程式 (e.g. 身高 = 係數 * 學年度 + 截距項)，並輸出 110 學年度該性別年齡的預測平均身高、以及這個方程式的 R 平方。
 - 請在程式中計算前一小題中迴歸方程式的顯著性檢定(F test)以及其 p-value。輸出這兩個數據後額外輸出這個模型是否有預測能力的結論。
2. 請利用附件資料包中的「鮑魚.txt」檔案作為匯入資料集，實作一個 kNN (k-近鄰演算法) 分類模型。
 - 資料集共有 4177 筆資料，性別欄 M=雄性，F=雌性，I=幼兒。年齡環類似樹木年輪，數值+1.5 代表鮑魚年齡。請注意資料集第一列是表頭要拿掉，另外資料集中的欄位分隔記號是逗號。
 - 輸出該資料集的性別次數分配表，以及同一性別其他變數的平均值。輸出項目如下：
[性別] [資料筆數] [平均長度] [平均直徑] ... [平均殼重] [平均年齡環數]

- 程式將資料集的前 3800 筆資料作為 kNN 模型的已知分類資料點，然後將後 377 筆作為未知分類資料點。系統接受使用者輸入的 k 值後，程式利用迴圈逐筆找出 k 個距離未分類資料點最近的已知資料點，然後以多數決輸出這個未分類資料點的分類結果。程式要輸出該 377 筆資料分類結果的混淆矩陣 (confusion matrix)，如下表所示。

kNN 分類結果 \ 真實性別	M	F	I
M			
F			
I			

- 以課堂介紹的 jFugue 套件，設計一個「絕對音感記憶遊戲」：
 - 由程式以亂數決定題目，在 {Do, Re, Mi, Fa, So, La, Si, Do} 的八度音中抽一個音撥放出來，讓使用者以你所設計的方式輸入他要回答的音高 (例如以 1~8 或其他字元符號代表)。
 - 若第一題答對，第二題要再多抽一個音，連同前一題的答案連續播放。例如第一題是 So，第二題就可能是 So+Re。玩家答題時也要將音字符串完整答對才能繼續進行。
 - 請設計輸出的畫面以防玩家可以看到前面自己輸入的前幾題答案。
 - 設計最高分排行榜。若遊戲結束時該次得分能夠進入排行榜，則讓玩家輸入任意字串後輸出排行榜。
 - 可加入難度選項，將出題音程加大或加入半音以達到難度變化。但答題的輸入方式需考量易用性。
- 設計可以讓兩個玩家或一個人跟電腦玩的撲克牌「BlackJack 21 點」遊戲。
 - 假設電腦與玩家的規則完全相同，沒有莊家。程式從一組 52 張的撲克牌中輪流發牌給兩位玩家，發出的牌面花色點數，以及目前累計點數都需顯示在螢幕上。在詢問玩家是否要再拿牌後由程式再跑亂數發牌。擁有最高點數的玩家獲勝，其點數必須等於或低於 21 點；超過 21 點的玩家稱為爆牌 (Bust)。2 點至 10 點的牌以牌面的點數計算，J、Q、K 每張為 10 點。A 可記為 1 點或 11 點，而 2-10 則按牌面點數算，若玩家會因 A 而爆牌則 A 可算為 1 點 (規則說明部分取自 Wiki)。
 - 請在詢問玩家是否要再拿牌前，計算在當下拿牌會發生爆牌的機率 (P1)，以及當下拿牌可以得到 21 點(含)或以下的機率 (P2)，以提供玩家參考。若只有一人玩，電腦擔任二號玩家的部分就以 $P2 > P1$ 作為其是否加牌的決策。
 - 以迴圈詢問玩家，在一局結束後是否要再重開一局。可加入賭金的設計，讓遊戲更接近現實玩法。

5. 請利用附件資料包中的「鐵達尼.txt」檔案作為資料來源，完成以下要求。
- 按照課堂介紹，在電腦上架設 apache 以及 mySQL。將資料集直接匯入 mySQL 的 test 資料庫 titanic 資料表中。請注意原始資料集中第一列為英文表頭，第二列為中文表頭，匯入前請做必要調整。
 - 在主程式中提供使用者直接輸入 SQL 語法的查詢字串，然後直接以該語法查詢資料庫中符合條件的資料並計算資料筆數，輸出在螢幕上。例如若輸入「select * from titanic where Survived = 1 and Sex=' male' 」，則程式要輸出所有倖存男性的資料以及資料筆數為 109。
 - 承上，在螢幕輸出結果後，將結果另外匯出至指定路徑的 QueryOutput.txt 檔案，方便使用者後續處理。請留意該檔案內的格式需與原始資料檔相同，也就是以逗號作為欄位分隔記號。
 - 可額外嘗試將 apache 與 mySQL 服務放在 internet 可存取的遠端主機。若達成此要求，請將提供服務的主機在繳交程式到 6/28 24:00 前維持 online 狀態，以利老師查核操作。