112 年度第 1 次 巨量資料分析師能力鑑定 中級試題

科目 1：B23 資料分析與資料科學

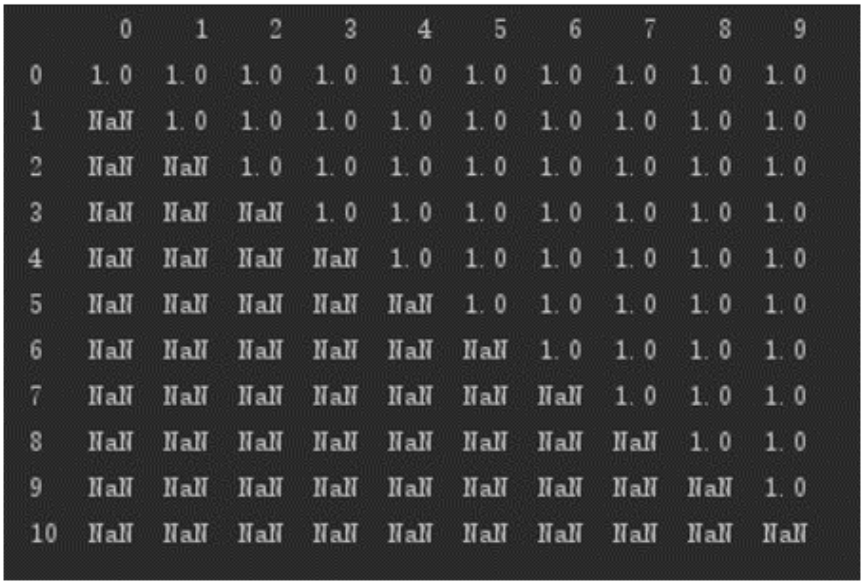
考試日期： 112 年 8 月 19 日

單選題 50 題（佔 **100%**）

A	1. 關於資料之遺缺值處理，下列敘述何者錯誤？ (A) 無須考慮遺缺值比例，全部刪除 (B) 類別資料補上眾數之值 (C) 利用模型補上估計產生之值 (D) 透過差值法（Interpolation Method）補上該值

C	2. 如附圖所示為 Python 語言 numpy 模組的使用，下列敘述何者正確？
	`import numpy as np`
	`x = np.array([[1,2,3,4], [4,5,6,7]])`
	(A) x.ndim 執行結果為：(2, 4)
	(B) x.size 執行結果為：6
	(C) x.reshape(-1, 2)執行結果：
	`array([[1, 2],`
	`[3, 4],`
	`[4, 5],`
	`[6, 7]])`
	(D) x.reshape(-1, -1) 執行結果：array([1, 2, 3, 4, 4, 5, 6, 7])

D	3. 如附圖所示，關於使用 Python 語言處理遺缺值的敘

	述，下列何者正確？ 假設透過 Python 語言的 pandas 套件，取得下列的 DataFrame d = pandas.DataFrame(data=a) print(d)  (A) print(d.dropna(axis=0, how='any')) 執行結果顯示 9 筆資料，每筆資料有 10 個列 (B) print(d.dropna(axis=0, how='all')) 執行結果顯示 1 筆資料，每筆資料有 10 個列 (C) print(d.dropna(axis='columns', thresh=5)) 執行結果顯示 10 筆資料，每筆資料有 5 個列 (D) print(d.dropna(axis=1, how='any', subset=[5,6,7])) 執行結果顯示 10 筆資料，每筆資料有 3 個列
B	4. 如附圖所示為 Python 語言 pandas 模組的使用，關於 sort_values 函數的敘述下列何者正確？

	0	1	2	3	4	5	6	7	8	9
0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	NaN	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2	NaN	NaN	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
3	NaN	NaN	NaN	1.0	1.0	1.0	1.0	1.0	1.0	1.0
4	NaN	NaN	NaN	NaN	1.0	1.0	1.0	1.0	1.0	1.0
5	NaN	NaN	NaN	NaN	NaN	1.0	1.0	1.0	1.0	1.0
6	NaN	NaN	NaN	NaN	NaN	NaN	1.0	1.0	1.0	1.0
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	1.0	1.0
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	1.0
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0
10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

```
In: import pandas as pd
...: mydf = pd.DataFrame(
...:     {'A': [1,2,2,4,2],
...:      'B': [10,20,26,8,29]})
...: mydf
Out:
   A   B
0  1  10
1  2  20
2  2  26
3  4   8
4  2  29

In: mydf.sort_values(by=['A', 'B'], ascending = [True, False])
```

(A)

```
Out:
   A   B
0  1  10
1  2  20
2  2  26
4  2  29
3  4   8
```

(B)

```
Out:
   A   B
0  1  10
4  2  29
2  2  26
1  2  20
3  4   8
```

(C)

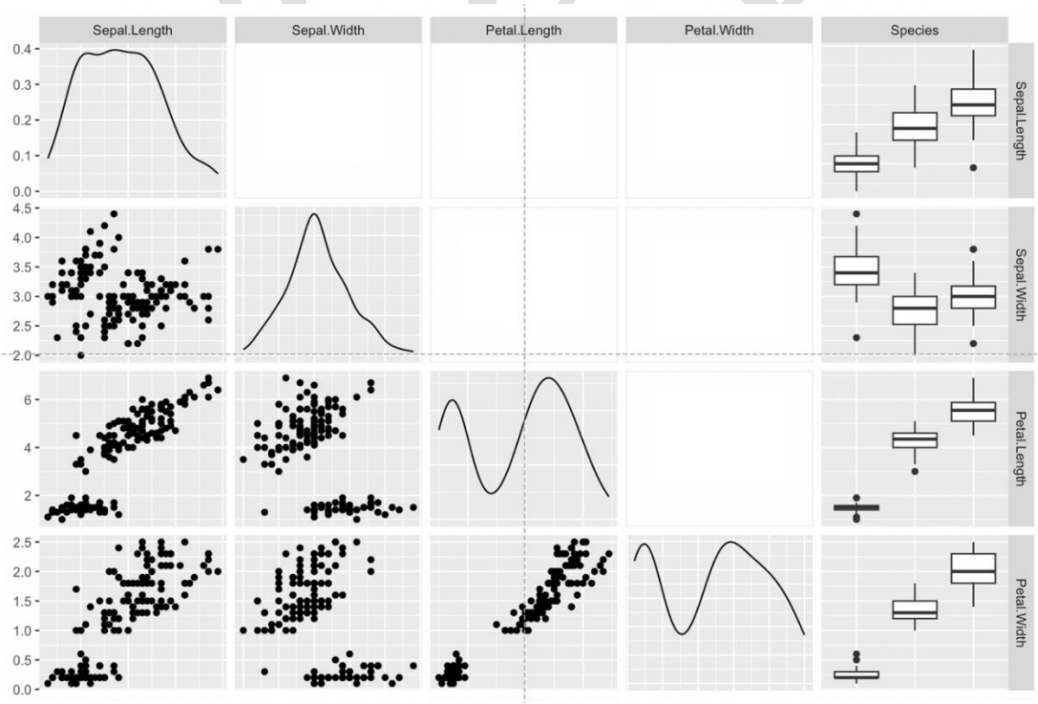
```
Out:
   A   B
3  4   8
4  2  29
2  2  26
1  2  20
0  1  10
```

(D)

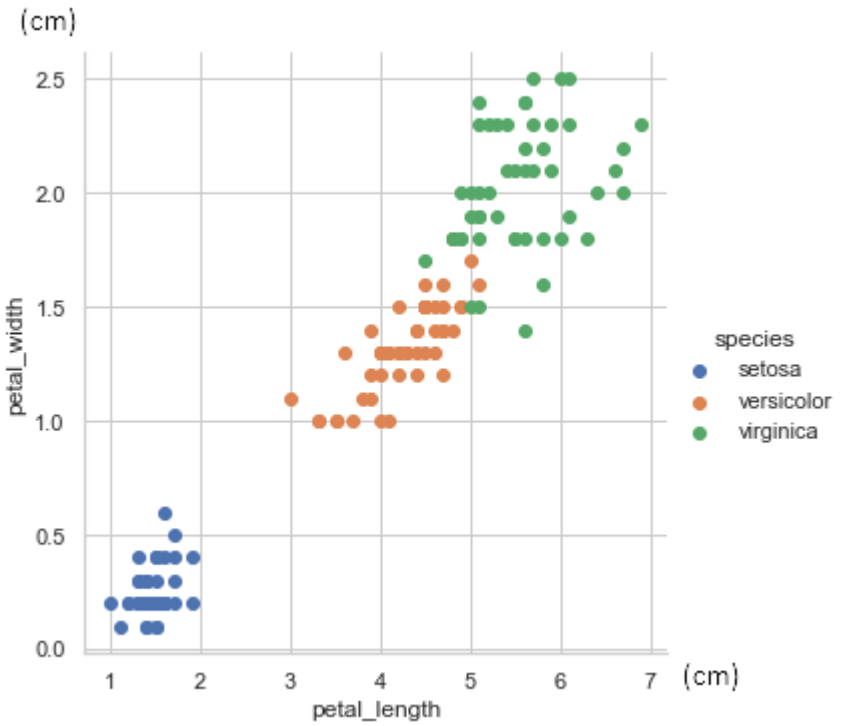
```
Out:
   A   B
3  4   8
1  2  20
2  2  26
4  2  29
0  1  10
```

C	5. 如附圖所示，使用 Python 語言處理布林陣列索引值（Boolean array indexing），下列敘述何者正確？ import numpy as np x = np.array([[[-0.26, 0.49, 0.18], [0.43, 0.3 , 0.29]], [[-0.44, 0.3 , 0.28], [0.27, -0.09, -0.13]]]) bool_ind = x > 0 print(bool_ind) (A) [0.49, 0.18, 0.43, 0.3, 0.29, 0.3, 0.28, 0.27] (B) [-0.26, -0.44, -0.13] (C) [[[False,True,True],[True,True,True]], [[False,True,True],[True,False,False]]] (D) [[[True, False, False], [False, False, False]], [[True, False, False], [False, True, True]]]
C	6. 關於資料標準化（Standardization）和正規化（Normalization）的敘述，下列何者錯誤？ (A) 標準化適用於數值屬性資料 (B) 正規化適用於數值屬性資料 (C) 標準化範圍值 [0, 1] (D) 正規化範圍值可以是 [0, 1] 或 [-1, 1]
D	7. 資料進行屬性轉換通常可以降低量綱尺度（Scale）對模型的影響。下列哪一種類型的模型方法，「不」需要做屬性轉換？ (A) k 平均數（k-means）集群 (B) 支援向量機 (C) 類神經網路 (D) 樹狀模型
C	8. 關於主成份分析（Principal Component Analysis, PCA）的敘述，下列何者錯誤？ (A) 主成份分析的過程不需考慮目標變數 (B) 主成份分析是屬於非監督式學習方法 (C) 主成份分析中各主成份的重要度依序遞增

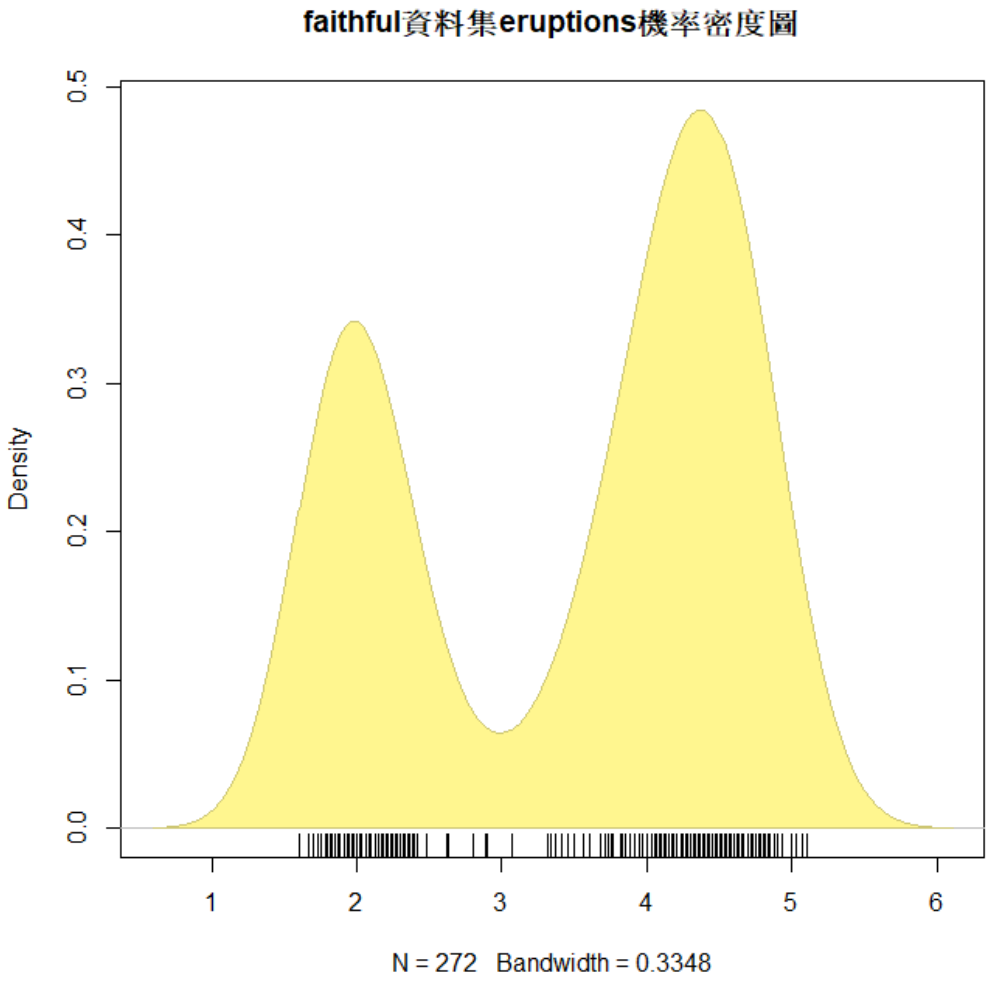
	(D) 主成份分析是一種特徵萃取技術
C	9. R 語言進行變數篩選時，有下列何種情形的變數「不」建議刪除？ (A) 分佈退化（Degenerate Distribution） (B) 高度偏斜變數 (C) 獨一無二的類別值數量與樣本大小的比例低於 10% (D) 某變數只有兩個類別值，最頻繁的類別值頻次是次頻繁的類別值頻次 19 倍以上
D	10. 請排出特徵工程（Feature engineering）四步驟最可能的正確順序？(a) 特徵優化（Feature optimization）。(b) 特徵理解（Feature understanding）。(c) 特徵結構化（Feature structuring）。(d) 特徵評估（Feature evaluation）。 (A) b a c d (B) a b c d (C) c b a d (D) b c a d
A	11. 下列何種類別變量編碼方式，最容易產生過度配適（Overfitting）的結果？ (A) 均值編碼（Mean encoding） (B) 虛擬編碼（Dummy encoding） (C) 標籤編碼（Label encoding） (D) 獨熱編碼（One-hot encoding）
D	12. 過濾法（Filtering）的特點是不涉及模型。下列何項「不」屬於特徵挑選（Feature Selection）三大類方法中的過濾法？ (A) 低變異過濾（Low variance filtering） (B) 卡方檢定（Chi-squared test） (C) 訊息增益（Information gain） (D) 遞迴特徵刪除（Recursive Feature Elimination, RFE）

A	13. 關於圖表功能的說明，下列何者錯誤？ (A) 散佈圖可以知道變數的因果關係 (B) 機率密度圖可以知道資料偏斜的可能性 (C) 長條圖（Bar chart）可以比較變數的差異 (D) 直方圖（Histogram）的橫軸為「連續型數值變數」
D	14. 請依據以下三點資訊，利用盒鬚圖（Box plot）判斷下列何者是極端值？(1) 第一四分位數：70。(2) 中位數：80。(3) 第三四分位數：110。 (A) 12 (B) 75 (C) 150 (D) 171
D	15. 如附圖所示為視覺化圖表，請依據圖表（從左至右）判斷下列敘述何者正確？  (A) Petal.Length 和 Petal.Width 有高度負相關 (B) Sepal.Width 和 Sepal.Length 線性相關性可能很高 (C) Petal.Length 三種類型都有極端值 (D) Sepal.Width 第一種類型的中位數最大

B	16. 下列何種圖形「無法」看出數據中的多峰分佈（Multimodal Distribution）特性？ (A) 長條圖（Bar plot） (B) 盒鬚圖（Box and whisker plot） (C) 直方圖（Histogram） (D) 散佈圖（Scatter plot）
D	17. 下列何種圖形最適合呈現相關係數（Correlation Coefficients）矩陣？ (A) 馬賽克圖（Mosaic plot） (B) 四重圖（Fourfold display） (C) 濾網圖（Sieve plot） (D) 熱圖（Heat map）
C	18. 下列何種圖形運用視覺化方式進行類別變數之間的關聯檢驗（Association Test）？ (A) 克里夫蘭點圖（Cleveland dot plot） (B) 盒鬚圖（Box and whisker plot） (C) 濾網圖（Sieve plot） (D) 熱圖（Heat map）
B	19. 如附圖所示為鳶尾花資料集（Iris dataset）所繪製而成的分佈圖。關於對該數據與圖表的敘述，下列何者較「不」正確？（x 軸為為花瓣長度；y 軸為花瓣寬度）

<table>
<tr><td></td><td> 

(A) 整體不分花種來看，花瓣長度與花瓣寬度之間呈現正相關（Positive correlation）
(B) versicolor 花種之平均花瓣寬度，是三個花種最大
(C) setosa 花種的花瓣長度，大部分在 1-2 公分之間
(D) virginica 花種的花瓣寬度分佈較廣，其最大最小之差距可以達到接近 1 公分</td></tr>
<tr><td>D</td><td>20. 如附圖所示為 faithful 資料集，請問 eruptions 機率密度圖敘述下列何者正確？</td></tr>
</table>



(A) 資料呈現單峰分佈

(B) 資料出現最多的數值約為 3

(C) 資料最大值約 0.5

(D) 位於 [4, 5] 範圍的資料筆數較位於 [2, 3] 範圍多

D

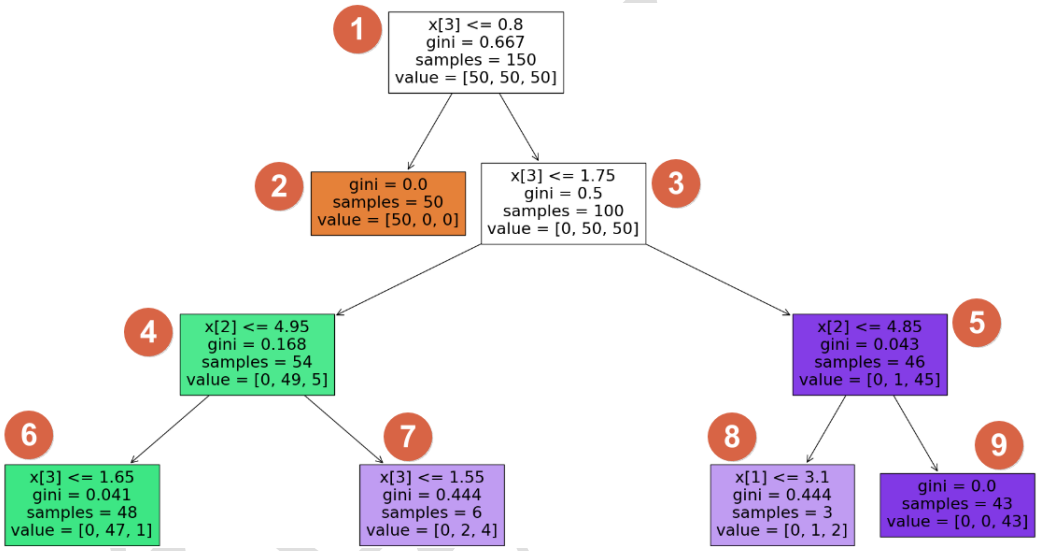
21. 關於分類方法的敘述，下列何者錯誤？

(A) 單純貝式（Naïve Bayes）分類是一種機率方法

(B) 支援向量機（Support Vector Machines, SVM）可以用來進行異常偵測

(C) 單純貝式（Naïve Bayes）是基於貝式定理而發展出來的方法

(D) 支援向量機（Support Vector Machines, SVM）是非監督式學習的一種

A	22. 關於 k 近鄰法（KNN）的敘述，下列何者正確？ (A) 基本運作是基於樣本間的距離為基礎 (B) 是非監督式學習的一種 (C) k 值小，容易配適不足 (D) k 近鄰法需要有適配模型才能進行
C	23. 如附圖所示，Python 語言使用 Iris 資料集與「from sklearn import tree」方式建立決策樹的結果，下列敘述何者正確？  (A) 樹根的吉尼係數（Gini coefficient）值為 0.041 (B) 節點 8 與節點 9 分類結果的資料點個數為 54 個 (C) 考量 x[1]=3.1, x[2]=4.96, x[3]=1.45，則決策樹分類結果會位於節點 7 (D) 決策樹使用吉尼集中度（Gini concentration）係數決定分裂條件
B	24. 如附圖所示，假設有一個 Python 名稱為 df 的 pandas DataFrame。關於使用 Python 語言「購物籃分析」（Basket Analysis）的敘述，下列何者正確？

```
from mlxtend.frequent_patterns import apriori
freq_itemsets = apriori(df, min_support=0.8, use_colnames=True)
print(freq_itemsets)
```

	Apples	Bread	Eggs	Fish	Ice cream	Kidney Beans	Milk	Onion	Sugar	Tea Leaves	Yoghurt
0	False	True	True	False	False	True	True	True	False	False	True
1	False	True	True	True	False	True	False	True	False	False	True
2	True	False	True	False	False	True	True	False	False	False	False
3	False	False	False	False	False	True	True	False	True	True	True
4	False	False	True	False	True	True	False	True	False	True	False

(A)

	support	itemsets
0	0.8	(2)
1	1.0	(5)
2	0.8	(2, 5)

(B)

	support	itemsets
0	0.8	(Eggs)
1	1.0	(Kidney Beans)
2	0.8	(Kidney Beans, Eggs)

(C) Error

(D) NaN

B

25. 如附圖所示，假設有一個 Python 名稱為 df 的 pandas DataFrame。關於使用 Python 語言「關聯規則學習」（Association Rule Learning）的敘述，下列何者正確？

```
from mlxtend.frequent_patterns import apriori
from mlxtend. frequent_patterns import association_rules
freq_itemsets = apriori(df, min_support=0.8, use_colnames=True)
association = association_rules(freq_itemsets, metric='confidence',
min_threshold=0.7)
print(association)
```

	Apples	Bread	Eggs	Fish	Ice cream	Kidney Beans	Milk	Onion	Sugar	Tea Leaves	Yoghurt
0	False	True	True	False	False	True	True	True	False	False	True
1	False	True	True	True	False	True	False	True	False	False	True
2	True	False	True	False	False	True	True	False	False	False	False
3	False	False	False	False	False	True	True	False	True	True	True
4	False	False	True	False	True	True	False	True	False	True	False

(A) 得到有 0 條關聯規則

	(B) 得到有 2 條關聯規則 (C) 得到有 3 條關聯規則 (D) 得到有 4 條關聯規則
D	26. 監督式學習是指在反應變數 y 引導下的學習;而非監督式學習則是指沒有或暫不考慮具體的學習目標 y 的學習。關於巨量資料分析方法論，下列何者是非監督式學習？ (A) 偏最小平方法（Partial Least Squares, PLS） (B) 單純貝氏分類（Naïve Bayes classifier） (C) 支援向量機（Support Vector Machines） (D) 集群分析（Cluster analysis）
D	27. 關於效能提升法（Boosting）的敘述，下列何項正確？ (A) 英文全名為 Bootstrap aggregating (B) 產生的模型集合俗稱裝袋樹（Bagged trees） (C) 在集成模型中融入屬性隨機挑選的機制 (D) 效能提升之意是建立多個互補的弱模型（Weak learner），將之集成後發揮團結力量大的綜效
A	28. 請簡述在隨機森林進行離散型分類評估演算法中，如何進行森林投票？ (A) 根據森林中每一棵樹評估結果，取眾數做為最後投票結果依據 (B) 根據森林中每一棵樹評估結果，取平均做為最後投票結果依據 (C) 根據森林中每一棵樹評估結果，隨機取 1 棵樹的結果做為最後投票結果依據 (D) 根據森林中每一棵樹評估結果，隨機取 k 棵樹的結果再取中位數做為最後投票結果依據
C	29. 關於訓練機器學習（Machine Learning）模型的敘述，下列哪一項錯誤？ (A) 資料清理、特徵萃取、特徵選擇都是重要的過程

	(B) 機器學習是實現人工智慧的其中一種方式 (C) 為資料貼標是機器學習的必要方法 (D) 特徵萃取（Feature Extraction）是從資料中挖出可以用的特徵
A	30. 關於非監督式學習（Unsupervised Learning），下列敘述何者正確？ (A) 在訓練時僅須對機器提供輸入無標籤的範例，非監督式學習的方法會自動從這些範例中找出潛在的規則 (B) KNN（K Nearest Neighbor）演算法屬於非監督式學習方法 (C) 針對網站上線後進行 A/B Test 是屬於非監督式學習的一種實務應用 (D) 因為對大量資料進行標籤相當費時，所以非監督式學習只需要對少部分資料進行標籤即可
D	31. 下列何者「不」是支援向量機的特點？ (A) 背後的最佳化問題為凸性的（Convex），因此只有一個最佳解存在 (B) 可用於分類或迴歸問題，且其績效通常十分卓著 (C) 不容易受雜訊資料過度的影響，也較不易過度配適 (D) 屬於白盒模型，結果容易詮釋
C	32. 關於決策樹下列何者敘述錯誤？ (A) 可用於分類問題 (B) 可用於迴歸問題 (C) 屬於黑盒模型的一種 (D) 屬於白盒模型的一種
C	33. 當訓練資料有遺缺值（Missing Value），某些樣本缺少某些維度的資料時，下列何者通常較「不適合」處理方式？ (A) 將遺缺資料的樣本移除 (B) 以原始資料的平均值補上遺缺資料

(C) 將遺缺資料全部補成 0

(D) 可以用 K-Nearest Neighbours 演算法找出合理的數字補上

A

34. 如附圖所示，當身體肌肉收縮時，肌電訊號（Electromyography, EMG）肌電信號可用於控制計算機，做為一種用戶界面。EMG 訊號一般可由：頻率（Frequency, F）、強度（Strength, S）與時間（Time, T）來表示，下表為 EMG 的實驗資料（F, S, T）和相應的動作分類（Action, A），若用 Gini 係數來建立決策樹模型，第一個分類屬性為下列哪一項？

Frequency (F)	Strength (S)	Time (T)	Action (A)
810	1	1	A1
864	1	0.5	A2
485	1	1	A3
950	1	0.5	A2
1003	1	0.5	A2
524	1	1	A3
736	1	0.5	A4
661	1	0.5	A4
*	2	*	A5

(A) 頻率（Frequency, F）

(B) 強度（Strength, S）

(C) 時間（Time, T）

(D) 動作分類（Action, A）

C

35. 與隨機誤差建模相關的參數（Parameters）有兩種：一種是直接利用資料估計其值的模型參數；另一種是不易從資料中估計的超參數（Hyperparameters），下列何者「不是」超參數？

(A) 迴歸方程式的預測變數集

(B) 人工神經網路的隱藏層節點數

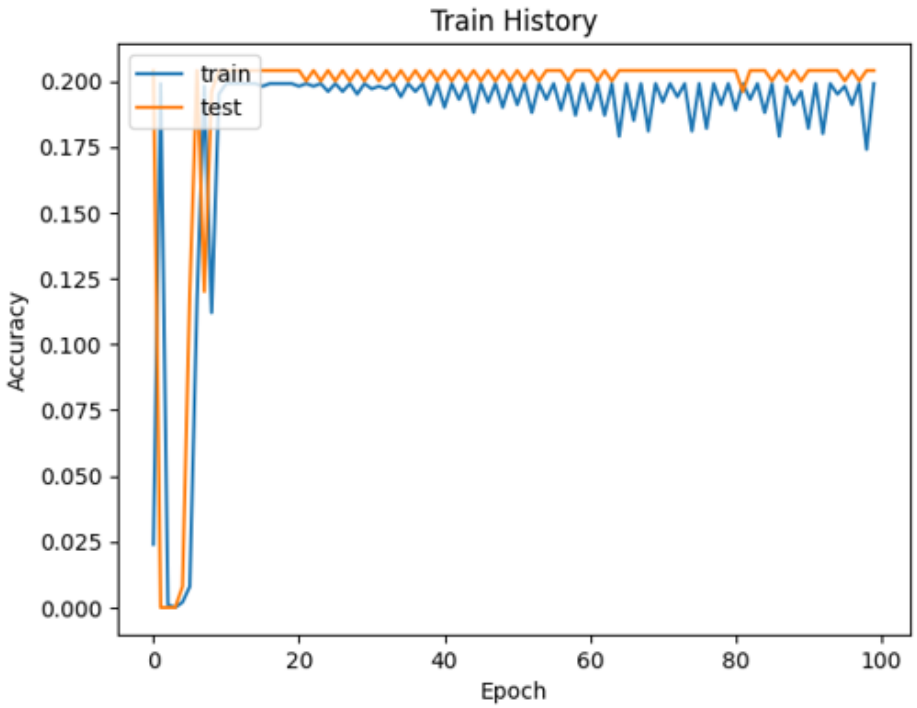
(C) 迴歸方程式的截距與斜率係數

	(D) 支援向量機中徑向基底核函數的參數
A	36. 模型選擇與評定時，經常運用重抽樣方法進行模型訓練與測試，下列敘述何者錯誤？ (A) 模型評定（Model assessment）的工作包括同一模型不同參數的調校（Within model），以及跨越不同模型的比較（Between models） (B) 模型優化（Model optimization）的工作則是在找到可以最小化訓練集損失的最佳參數 (C) 拔靴抽樣相較於其他方法有較低的變異 (D) 一般而言 k 摺交叉驗證（k-fold cross validation）相較於他法有較高的變異，但當訓練集大時則此問題較不嚴重
B	37. 當資料科學家建模時，下列何者最可能為過度配適（Overfitting）的狀況？ (A) 測試誤差高，訓練誤差高 (B) 測試誤差高，訓練誤差低 (C) 測試誤差低，訓練誤差低 (D) 測試誤差低，訓練誤差高
B	38. 如附圖所示，抓取 10 萬則 Google Play 評論，欲以 Rating 作為情感模型訓練，以附圖資料集為例，實際訓練時，採用斷詞後的詞彙轉換為詞嵌入向量後，將 rating 如附圖 1 所示，在以循環神經網路（Recurrent Neural Network, RNN）訓練之後，得到結果如附圖 2，請問下列敘述哪一項正確？

附圖 1

comments	ratings
所以 ..我為了要知道一個不確定的未來 而付錢 付了才能看結果 那我載個毛	1
使用到今天快要兩年，非常的穩定而且客服回應速度快，推一個	5
還算可以吧，一直也沒遇到什麼問題	3
垃圾! 連線品質真的很糟糕!!點個桌老半天進不去!!竟然還能刷卡儲值!?是在騙錢嗎?!! ...	1
第二章第三關下方有遊戲廣告阻礙遊戲進行 每一關都會不定時跳出廣告，第二章第三關更誇張，遊戲中...	2

附圖 2：結果



(A) 訓練得到很高的準確率，可以對文字情感做較好的預測

(B) 觀察到 RNN 模型的測試準確率幾乎高於訓練準確率，推測配適狀況可能不明

(C) 此模型的 F1 score 應為 0.88 以上

(D) 訓練更多的 Epoch 會得到更高的準確率

B 39. 下列那一種重抽樣方法是採行多次置回抽樣的方法，取出

	通常與原樣本大小相同的子集？ (A) 保留法（Holdout） (B) 拔靴抽樣（Bootstrapping） (C) 袋外樣本（Out-of-bag Samples） (D) 交叉驗證（Fold Cross Validation）
B	40. 在進行巨量資料分析時，有些時候，我們因欠缺領域的專業知識，而且可能所有屬性都要考慮，這時，會有巨量資料的維度詛咒（Curse of dimensionality）的困擾。下列何者接近函數（Proximity function）相對較適合用在高維度巨量資料的情境？ (A) 漢明距離（Hamming distance） (B) 餘弦相似度（Cosine similarity） (C) 歐幾里德距離（Euclidean distance） (D) 曼哈頓市街距離（Manhattan distance）
D	41. 進行巨量資料分析時，最佳模型（Model）參數值或參數組合的選擇，是件重要但不容易的工作，資料科學家通常會運用_(1)_進行估計，並以預測誤差估計值的全域最小值來決定模型複雜度或最佳參數組合。獲得單類模型的最佳參數後，還須跨越不同類型的模型（Models）進行_(2)_比較，透過其中的_(3)_來決定各模型差距的顯著狀況。請將下列方法論，依順序正確地填內上述的(1)(2)(3) 空格中。甲：正確率或 Kappa 係數和 p 值。乙：統計檢定。丙：交叉驗證。 (A) 甲 -> 乙 -> 丙 (B) 乙 -> 丙 -> 甲 (C) 丙 -> 甲 -> 乙 (D) 丙 -> 乙 -> 甲
B	42. 如附圖所示，某公司希望透過機器學習來預測客戶是否會購買某項產品，以便提高銷售量。附圖 1 為該公司的資料庫，請問下列哪些是正確的處理步驟？

附圖 1

客戶 ID	性別	年齡	職業	購買金額	是否購買
001	男	30	工程師	500	否
002	女	40	記者	300	是
003	男	50	教師	800	否
004	女	35	銀行家	1200	是

處理步驟：

A. 從資料集中移除客戶 ID，因為它不是用來建立模型的特徵。

B. 將「性別」轉換為字串型資料，例如男性為 M，女性為 F。

C. 使用 K-fold Cross Validation 來將資料集分成訓練集和測試集。

D. 在測試集上使用混淆矩陣（Confusion Matrix）來評估模型效能，計算準確率、召回率、F1 score 等指標。

(A) ABD
(B) ACD
(C) ABC
(D) BCD

B

43. 關於迴歸分析與變異數分析的異同，下列敘述何者錯誤？
(A) 皆在探討反應變數（Y）與解釋變數（X）間的統計關係
(B) 反應變數（Y）可為數量變數或屬質變數
(C) 變異數分析是用來比較不同刺激對反應變數平均數的影響的統計方法
(D) 迴歸分析討論反應變數與解釋變數間關係的數學式；變異數分析利用解釋變數去解釋反應變數的變異

C

44. 關於兩變量關聯（Association）、相關（Correlation）與因果（Causation）的敘述，下列何者錯誤？
(A) 關聯不代表相關
(B) 相關代表關聯
(C) 因果代表關聯

(D) 關聯代表因果

A

45. 如附圖所示，某家零售企業希望透過深度神經網路（Deep Neural Networks, DNN）模型預測產品銷售量，以便進行產品管理和庫存優化。附圖 1 是企業的銷售資料，請問下列哪些是正確的處理步驟？

附圖 1：企業銷售資料

日期	產品編號	廠商編號	廠商區域	廣告費用	宣傳標語	價格	銷售量
2022/1/1	001	101	區域 A	500	健康第一	100.00	200
2022/1/1	002	102	區域 B	600	新年新氣象	80.00	150
2022/1/2	001	101	區域 A	700	快樂運動	110.00	250
2022/1/2	003	103	區域 C	400	輕鬆健身	90.00	100
2022/1/3	002	102	區域 B	550	健身王	85.00	180
2022/1/3	003	103	區域 C	350	快樂運動	95.00	120
2022/1/4	001	101	區域 A	600	健康第一	105.00	240
2022/1/4	002	102	區域 B	500	新年新氣象	80.00	170
2022/1/5	003	103	區域 C	450	輕鬆健身	90.00	130

處理步驟：

A. 將廣告費用、價格、銷售量等數值資料進行標準化，以避免不同特徵間的數值差異對模型效能造成影響。

B. 將宣傳標語轉換為詞向量表示，以便模型能夠處理文本資料。

C. 將產品編號、廠商編號、廠商區域轉換為 one-hot encoding，以便模型能夠處理分類資料。

D. 將日期欄位轉換為時間戳記格式，以便模型能夠處理時間序列資料。

(A) ABCD
(B) ABC
(C) BD
(D) C

A

46. 下列何種卷積神經網路（Convolution Neural Networks, CNN）首度提出添加棄卻層（Dropout）的方式，以降低模型過度配適（Overfitting）的現象？

(A) AlexNet
(B) Inception
(C) LeNet
(D) VGG19

D	47. 關於長短期記憶（Long Short Term Memory, LSTM）神經網路的閘門（gate）與參數矩陣的敘述，下列何項正確？ (A) 四個閘門（gates），有四個參數矩陣需要從資料中估計 (B) 三個閘門（gates），有兩個參數矩陣需要從資料中估計 (C) 四個閘門（gates），有三個參數矩陣需要從資料中估計 (D) 三個閘門（gates），有四個參數矩陣需要從資料中估計
B	48. 下列何種卷積神經網路（Convolution Neural Networks, CNN）是將卷積層加寬而非加深？ (A) R-CNN (B) Inception (C) ResNet (D) VGG19
C	49. 下列何種卷積神經網路（Convolution Neural Networks, CNN）提出跳層（skip connections）概念，解決梯度消失（gradient varnishing）的問題，讓深層網路更容易訓練，開啟了超深網路的時代？ (A) R-CNN (B) Inception (C) ResNet (D) VGG19
B	50. 關於循環神經網路（Recurrent Neural Networks, RNN）在自然語言之應用的敘述，下列何者錯誤？ (A) 多對一的 RNN 用於情感分析（Sentiment Analysis） (B) 一對多的 RNN 用於對話回應（Question Answering） (C) 一對一的 RNN 可建立語言模型（Language

	Modeling） (D) 多對多的 RNN 用於語言翻譯（Language Translation）