勞動部勞動力發展署桃竹苗分署 114 年度自辦在職訓練甄試考卷

Python 機器學習建模與評估班 試卷編號-□01■02□03

(114 年 1 月 11 日■上午□下午)

專業科目選擇題 50 題（單選題，1 題 2 分）考試不得使用計算機。

題號	試　　題

01. 什麼是「資料清理」？（A）刪除資料集中所有資料（B）將資料儲存到檔案中（C）處理資料中的缺失值和錯誤（D）以上皆非
02. 什麼是「資料轉換」？（A）將資料從一種格式轉換為另一種格式（B）將資料儲存到資料庫中（C）使用資料進行預測（D）以上皆非
03. 什麼是「特徵選擇」？（A）選擇對模型重要的特徵（B）建立新的特徵（C）將特徵縮放到特定範圍（D）以上皆非
04. 什麼是「降維」？（A）減少資料集中特徵的數量（B）增加資料集中特徵的數量（C）將資料分成 K 個群集（D）以上皆非
05. PCA（主成分分析）常用於哪種資料處理任務？（A）特徵縮放（B）特徵選擇（C）降維（D）以上皆非
06. 什麼是「訓練集」？（A）用於訓練模型的資料（B）用於評估模型的資料（C）用於測試模型的資料（D）以上皆非
07. 什麼是「測試集」？（A）用於訓練模型的資料（B）用於評估模型的資料（C）用於測試模型的資料（D）以上皆非
08. 什麼是「損失函數」？（A）用於評估模型性能的指標（B）用於衡量模型預測誤差的函數（C）用於優化模型的演算法（D）以上皆非
09. 什麼是「學習率」？（A）模型訓練的速度（B）模型的複雜度（C）模型的泛化能力（D）以上皆非
10. 什麼是「epoch」？（A）模型訓練的迭代次數（B）模型的層數（C）模型的參數數量（D）以上皆非
11. 什麼是「偏差-方差權衡」？（A）模型的準確率和召回率之間的關係（B）模型的偏差和方差之間的關係（C）模型的複雜度和泛化能力之間的關係（D）以上皆非
12. 什麼是「正則化」？（A）用於提高模型準確率的技術（B）用於防止模型過擬合的技術（C）用於處理缺失值的技術（D）以上皆非
13. 什麼是「超參數優化」？（A）調整模型的超參數以提高性能的過程（B）訓練模型的過程（C）評估模型的過程（D）以上皆非
14. 什麼是「模型部署」？（A）將模型應用於實際問題的過程（B）訓練模型的過程（C）評估模型的過程（D）以上皆非
15. Python 中，用於表示字串的資料類型是？（A）int（B）float（C）str（D）bool
16. 哪個 Python 函數可以用於取得列表的長度？（A）size()（B）length()（C）len()（D）count()
17. Python 中，哪種資料結構可以儲存鍵值對？（A）列表（List）（B）字典（Dictionary）（C）元組（Tuple）（D）集合（Set）
18. 哪個 Python 函數可以用於將字串轉換為整數？（A）int()（B）str()（C）float()（D）bool()
19. Python 中，如何定義一個空列表？（A）[]（B）{}（C）()（D）None
20. 什麼是「機器學習」？（A）一種程式設計語言（B）一種人工智慧技術（C）一種資料庫管理系統（D）以上皆非
21. 什麼是「監督式學習」？（A）從帶有標籤的資料中學習（B）從不帶標籤的資料中學習（C）使用強化學習方法（D）以上皆非
22. 什麼是「迴歸」問題？（A）預測連續值的任務（B）預測類別的任務（C）將資料分成 K 個群集的任務（D）以上皆非
23. 什麼是「分類」問題？（A）預測連續值的任務（B）預測類別的任務（C）將資料分成 K 個群集的任務（D）以上皆非
24. 什麼是「特徵」？（A）資料集中每個樣本的預測目標（B）資料集中每個樣本的屬性或特徵（C）資料集中所有樣本的集合（D）以上皆非
25. 決策樹（Decision Tree）屬於哪種類型的機器學習演算法？（A）監督式學習（B）無監督式學習（C）強化學習（D）以上皆非
26. K-Means 分群（K-Means Clustering）的主要目標是什麼？（A）找到一個超平面來最大化資料間的間隔（B）將資料分成 K 個群集（C）建立一個決策樹來進行預測（D）以上皆非
27. 哪種模型適合用於處理影像分類問題？（A）線性迴歸（B）卷積神經網路（CNN）（C）遞迴神經網路（RNN）（D）以上皆非
28. 哪種模型適合用於處理自然語言處理（NLP）問題？（A）邏輯迴歸（B）支持向量機（C）遞迴神經網路（RNN）（D）以上皆非
29. 什麼是「深度學習」？（A）一種機器學習技術，使用多層神經網路（B）一種資料庫管理系統（C）一種程式設計語言（D）以上皆非
30. scikit-learn 套件的主要功能是什麼？（A）資料視覺化（B）資料清理（C）機器學習模型訓練和評估（D）以上皆非
31. pandas 套件的主要功能是什麼？（A）資料視覺化（B）資料清理和分析（C）機器學習模型訓練和評估（D）以上皆非
32. NumPy 套件的主要功能是什麼？（A）資料視覺化（B）數值計算（C）機器學習模型訓練和評估（D）以上皆非
33. matplotlib 套件的主要功能是什麼？（A）資料視覺化（B）資料清理和分析（C）機器學習模型訓練和評估（D）以上皆非
34. 哪個 Python 套件可以用於深度學習模型的開發？（A）scikit-learn（B）TensorFlow（C）PyTorch（D）以上皆是
35. 如何在 Python 中使用 while 迴圈？（A）while condition:（B）while condition: else:（C）while: else:（D）while condition then:
36. 如何在 Python 中使用 try-except 語句處理異常？（A）try: except:（B）try: except Exception as e:（C）try: finally:（D）以上皆可
37. 如何在 Python 中讀取檔案？（A）open('file.txt', 'r')（B）read('file.txt')（C）load('file.txt')（D）get('file.txt')
38. 如何在 Python 中寫入檔案？（A）open('file.txt', 'w')（B）write('file.txt')（C）save('file.txt')（D）put('file.txt')
39. 如何在 Python 中使用 def 關鍵字定義一個帶有參數的函數？（A）def function_name(parameter1, parameter2):（B）def function_name():（C）function_name(parameter1, parameter2):（D）以上皆非
40. 什麼是「資料缺失」？（A）資料集中缺少某些值（B）資料集中包含錯誤的值（C）資料集太大無法處理（D）以上皆非
41. 什麼是「資料不平衡」？（A）資料集中不同類別的樣本數量不均勻（B）資料集中包含缺失值（C)資料集中包含錯誤的值（D）以上皆非
42. 什麼是「特徵工程」？（A）收集和清理資料的過程（B）選擇和轉換特徵以提高模型性能的過程（C)訓練和評估模型的過程（D）部署和監控模型的過程
43. 什麼是「One-Hot 編碼」？（A）一種特徵縮放技術（B）一種特徵選擇技術（C）一種用於處理類別型特徵的技術（D）以上皆非
44. 什麼是「資料標準化」？（A）將特徵值縮放到 0 到 1 之間（B）將特徵值轉換為具有零均值和單位方差（C）將特徵值轉換為具有特定分佈（D）以上皆非

45. 什麼是「過擬合」？（A）模型在訓練資料上表現良好，但在測試資料上表現不佳（B）模型在訓練資料和測試資料上表現都不佳（C）模型在訓練資料和測試資料上表現都很好（D）以上皆非
46. 什麼是「欠擬合」？（A）模型在訓練資料上表現良好，但在測試資料上表現不佳（B）模型在訓練資料和測試資料上表現都不佳（C）模型在訓練資料和測試資料上表現都很好（D）以上皆非
47. 什麼是「驗證集」？（A）用於訓練模型的資料（B）用於評估模型的資料（C）用於調整模型超參數的資料（D）以上皆非
48. 哪個 Python 函數可以用於產生 0 到 1 之间的随机数？（A）random.randint()（B）random.random()（C）random.uniform()（D）random.randrange()
49. 在機器學習中，什麼是「訓練資料」？（A）用於訓練模型的資料（B）用於評估模型的資料（C）用於測試模型的資料（D）以上皆非
50. 哪種機器學習演算法常用於處理「異常偵測」問題？（A）線性迴歸（B）支持向量機（C）K-Means 分群（D）Isolation Forest

交卷時請連同答案卷一起繳回

Python 機器學習建模與評估班 答案卷編號-□01■02□03

選擇題（單選）共 50 題；每題 2 分(答錯不倒扣)

C	A	A	A	C	A	C	B	A	A
1	2	3	4	5	6	7	8	9	10

B	B	A	A	C	C	B	A	A	B
11	12	13	14	15	16	17	18	19	20

A	A	B	B	A	B	B	C	A	C
21	22	23	24	25	26	27	28	29	30

B	B	A	D	A	D	A	A	A	A
31	32	33	34	35	36	37	38	39	40

A	B	C	B	A	B	C	B	A	D
41	42	43	44	45	46	47	48	49	50