

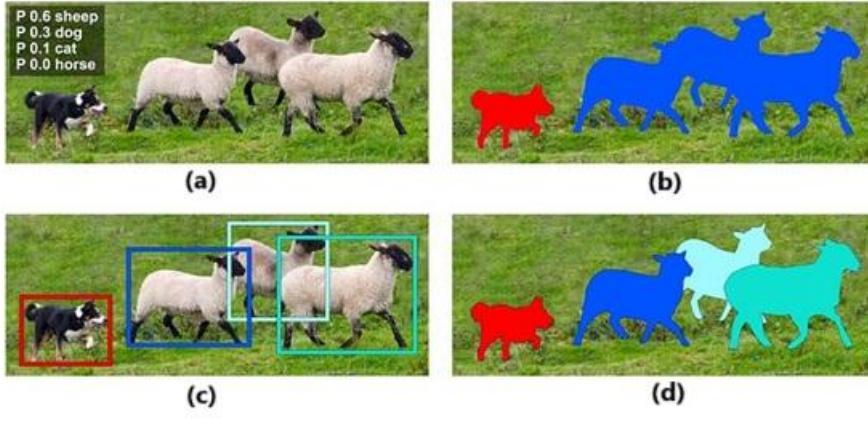
iPAS AI 應用規劃師中級能力鑑定-考試樣題

(樣題範例非正式考題 · 僅供參考使用)

114.09 版

◆ 科目一：人工智慧技術應用與規劃

題號	答案	題目
1	A	<p>下列何者並未使用人工智慧 (AI) 或機器學習 (ML) 技術？</p> <p>(A) 使用一組預定義規則來確定最佳移動做法的象棋遊戲</p> <p>(B) 使用深度神經網路來提高其準確性的語音識別系統</p> <p>(C) 使用感測器和預定義的自動駕駛汽車-定義導航規則</p> <p>(D) 使用自然語言處理算法來理解和反應用戶查詢的聊天機器人</p>
2	C	<p>在文本資料處理過程中，通常會需要「將接續的文本轉換為詞彙單位」，以便後續的處理。請問上述所指的是文本資料處理中的哪一個方法？</p> <p>(A) 詞形還原 (Lemmatization)</p> <p>(B) 停用詞移除 (Stopword Removal)</p> <p>(C) 斷詞 (Tokenization)</p> <p>(D) 詞頻 - 逆向文件頻率 (Term Frequency-Inverse Document Frequency, TF-IDF)</p>
3	A	<p>下列何者為自然語言處理 (NLP) 在機器學習應用中的主要用途？</p> <p>(A) 情緒分析</p> <p>(B) 圖像識別</p> <p>(C) 預測性維護</p> <p>(D) 供應鏈優化</p>
4	D	<p>關於深度學習模型，下列敘述何者不正確？</p> <p>(A) 卷積神經網路 (Convolutional Neural Networks) 適合影像辨識</p> <p>(B) ReLU (Rectified Linear Unit) 比 tanh 和 Sigmoid 好，因為 ReLU 可以減緩梯度爆炸與消失的現象</p> <p>(C) 遞迴神經網路 (Recurrent Neural Networks) 適合處理序列相關資料</p> <p>(D) Elman 神經網路 (Elman Neural Networks) 適合處理影像辨識</p>
5	A	<p>下列何者為機器學習模型在業界部署的主要趨勢？</p> <p>(A) 越來越多地採用自動化機器學習 (AutoML) 技術</p> <p>(B) 轉向使用更簡單的機器學習算法</p> <p>(C) 基於雲的機器學習平台的使用率下降</p> <p>(D) 依賴手動超參數調整進行模型優化</p>

題號	答案	題目
6	D	<p>下列何者「最適合」使用迴歸 (Regression) 模型進行預測？</p> <p>(A) 根據顧客的購買行為，預測他最可能加入購物車的商品類別 (B) 根據病患的症狀與病史，判斷其可能患有的疾病類別 (C) 根據電影的特徵與用戶評分，推薦適合的電影類型 (D) 根據歷年銷售數據，預測明年某產品的總銷售額</p>
		<p>附圖中，展示了電腦視覺辨識的幾項技術。</p> <p>請問下列選項中，何者正確配對了圖中 (a)、(b)、(c)、(d) 所代表的技術名稱？</p>  <p style="text-align: right;">(圖片來源：MIT 課程—Object Detection and Segmentation)</p>
7	B	<p>(A) (a) 實例分割 (Instance Segmentation) (b) 語義分割 (Semantic Segmentation) (c) 物件偵測 (Object Detection) (d) 圖像分類 (Image Classification);</p> <p>(B) (a) 圖像分類 (Image Classification) (b) 語義分割 (Semantic Segmentation) (c) 物件偵測 (Object Detection) (d) 實例分割 (Instance Segmentation);</p> <p>(C) (a) 圖像分類 (Image Classification) (b) 物件偵測 (Object Detection) (c) 語義分割 (Semantic Segmentation) (d) 實例分割 (Instance Segmentation);</p> <p>(D) (a) 圖像分類 (Image Classification) (b) 物件偵測 (Object Detection) (c) 實例分割 (Instance Segmentation) (d) 語義分割 (Semantic Segmentation)</p>
8	A	企業在生成式 AI 導入中，可選擇下列哪一種模型壓縮 (Model Compression) 技術以減少記憶體使用？

題號	答案	題目
		(A) 參數剪枝； (B) 增加訓練數據量； (C) 增加模型層數； (D) 使用更高維數據
9	B	供應鏈攻擊(Supply Chain Attacks)如何影響企業內部 AI 系統安全？ (A) 供應鏈攻擊的常見狀況，包括員工因釣魚郵件而被盜取帳號密碼； (B) AI 系統依賴的第三方開源模型或數據可能被植入惡意內容； (C) AI 系統只會受到 DDoS 攻擊影響，不會有供應鏈風險； (D) 供應鏈攻擊僅影響硬體層級、不影響 AI 模型
10	B	下列何者為自然語言處理(NLP)中的詞嵌入技術，能將文字轉換為向量以利機器學習處理？ (A) TF-IDF (B) Word2Vec (C) Stop Words (D) Bag-of-Words
11	B	在多模態學習中，早期融合 (Early Fusion) 方法的主要特徵為何？ (A) 將不同模態資料的輸出結果進行決策合併 (B) 在模型輸入階段或特徵提取階段整合不同模態資料 (C) 僅處理單一來源的資料輸入 (D) 利用注意力機制在深層隱藏層進行語意比對與融合
12	D	下列哪項技術最有助於強化醫療多模態 AI 系統在處理影像與文本數據時的整合能力？ (A) 利用預先定義的規則產生診斷結果 (B) 僅使用 CNN 架構同時處理影像與文字資訊 (C) 利用單一模態資料建立通用醫療模型 (D) 採用 Transformer 架構整合醫療影像與臨床文本資訊
13	C	某線上音樂平台希望根據用戶的聽歌與查詢行為，將用戶劃分為不同的類型。若事前沒有定義用戶類型，下列哪一種模型最適合用於此任務？ (A) 邏輯迴歸 (Logistic Regression); (B) 決策樹 (Decision Tree); (C) 基於密度之含噪空間聚類法 (DBSCAN); (D) 線性迴歸 (Linear Regression)
14	B	在訓練模型時，若數據中出現特徵尺度差異極大（例如：年齡為 0–100、收入為 0–1,000,000），容易導致模型偏向特定特徵。為提升模

題號	答案	題目
		<p>型效能與穩定性，以下哪一種預處理方式最能有效解決此問題？</p> <p>(A) 移除尺度較小的欄位以避免對模型影響過低</p> <p>(B) 對所有特徵進行 Z-score 標準化 (Standardization)</p> <p>(C) 將特徵縮放至 0–1 區間進行最小–最大正規化 (Min-Max Normalization)</p> <p>(D) 對所有數值欄位加上常數使其不為零</p>
15	B	<p>某企業即將部署 AI 模型至現有營運系統，進入系統整合測試階段。測試工程師需確認所有模組在實際環境中能正確協同運作。下列哪項驗證最應優先執行？</p> <p>(A) 確認模型在開發機上訓練集表現是否達標</p> <p>(B) 驗證模型服務與資料平台、前後端介面協同是否正常，資料格式與流程一致性是否維持</p> <p>(C) 檢查開發團隊提交程式碼時是否有依照 Git commit message 規範</p> <p>(D) 審閱模型報告與 API 文件的格式是否符合交付標準</p>



◆ 科目二：大數據處理分析與應用

題號	答案	題目
1	D	<p>在巨量資料分析班中，共有一年級至四年級，每個年級有 50 個學生，且學生身高呈常態分佈。下列敘述何者不正確？</p> <p>(A) 要檢測一年級和二年級的平均身高是否有差異，可以利用 t 檢定 (B) 要檢測一年級、二年級、三年級之間的平均身高是否有差異，可以利用 t 檢定 (C) 要檢測二年級、三年級、四年級之間的平均身高是否有差異，可以利用 F 檢定 (D) 要檢測一年級的平均身高是否等於 170 公分，可以利用卡方檢定</p>
2	A	<p>關於接受者操作特徵 (ROC) 曲線，下列敘述何者正確？</p> <p>(A) ROC 曲線繪製了真陽性率與假陽性率的關係 (B) ROC 曲線用於評估模型的準確性 (C) ROC 曲線下的面積 (AUC-ROC) 始終等於 1 (D) ROC 曲線只適用於二元分類問題</p>
3	D	<p>下列何者不屬於特徵工程 (Feature Engineering) ？</p> <p>(A) 轉換 (Transformation) (B) 萃取 (Extraction) (C) 挑選 (Selection) (D) 預測 (Prediction)</p>
4	D	<p>拉拉網路商城的老闆擬透過機器學習的方式，利用過往的產品銷售資料，預測下一季的產品銷售數量，以調整現有的庫存水位。下列哪一個類型的模型，比較適合應用在老闆期望的預測目標？</p> <p>(A) 決策樹分類器 (Decision Tree Classifier) (B) K-means 分群 (K-means Clustering) (C) 主成份分析 (Principal Component Analysis, PCA) (D) 線性迴歸 (Linear Regression)</p>
5	C	<p>對於低結構化的文本或圖像資料，下列哪一種特徵工程 (Feature Engineering) 方法最為適用？</p> <p>(A) 特徵改善 (Feature Improvement) (B) 特徵建構 (Feature Construction) (C) 特徵學習 (Feature Learning) (D) 特徵選擇 (Feature Selection)</p>
6	D	<p>下列哪種方法屬於非監督式學習中的降維技術？</p> <p>(A) K-均值聚類 (K-means) (B) 隨機森林 (Random Forest) (C) 支持向量機 (SVM) (D) 主成分分析 (PCA)</p>

題號	答案	題目
7	D	<p>異常值偵測使用 IQR (Interquartile Range) 法時，下列哪一種範圍外的數值會被視為異常？</p> <p>(A) 超過 $Q_2 \pm 2 \times IQR$ (B) 高於 $Q_1 - IQR$ 或低於 $Q_3 + IQR$ (C) 高於 $Q_1 - 2 \times IQR$ 或低於 $Q_3 + 2 \times IQR$ (D) 低於 $Q_1 - 1.5 \times IQR$ 或高於 $Q_3 + 1.5 \times IQR$</p>
8	A	<p>差分隱私 (Differential Privacy) 在數據應用中的目的為何？</p> <p>(A) 添加隨機噪聲掩蓋個體訊息 (B) 加密所有數據 (C) 移除多數類樣本 (D) 提高數據完整性</p>
9	B	<p>若使用主成分分析 (PCA) 將資料降維至兩個主成分，這表示哪一種情況？</p> <p>(A) 僅保留兩筆樣本資料進行模型訓練 (B) 將原始高維資料轉換為兩個主成分表示 (C) 每筆資料僅保留兩個原始欄位的數值 (D) 模型將限制只能用於分類任務</p>
10	C	<p>假設有一組成績資料呈現常態分配，平均數為 70，標準差為 10，若某位學生的成績為 90，該學生的 Z 分數約為多少？</p> <p>(A) 0 (B) 1 (C) 2 ; (D) 3</p>
11	C	<p>在 AI 倫理治理的背景下，「透明性 (Transparency)」通常是指什麼？</p> <p>(A) AI 系統的運算速度公開； (B) AI 系統僅使用公開數據； (C) AI 系統決策流程清晰且可解釋； (D) 所有 AI 模型都是開源的</p>
12	C	<p>若企業希望即時監控交易異常，應選擇下列哪一類數據處理架構？</p> <p>(A) 批次式資料倉儲 (B) 離線式資料湖 (C) 串流處理系統 (D) 冷資料備援架構</p>
13	C	為了分析社群網路使用者之間的互動結構，應使用下列哪種分析方

題號	答案	題目
		<p>法？</p> <p>(A) 文字探勘 (B) 主成分分析 (C) 圖論分析 (D) 分群分析</p>
14	C	<p>為了加速大數據環境下的 AI 模型訓練，以下哪一項為常見技術？</p> <p>(A) 早期停止 (Early Stopping) (B) 批次分群 (Mini-batching) (C) 混合精度訓練 (Mixed Precision Training) (D) 主成分分析 (PCA)</p>
15	B	<p>考慮使用 CIFAR-10 資料集進行資料處理，資料包括 32×32 像素的多筆彩色照片。下列程式碼的資料處理，請選出正確的選項。</p> <pre>[1]: from tensorflow.keras import datasets, utils import pandas as pd [2]: (x_train, y_train), (x_test, y_test) = datasets.cifar10.load_data() [3]: type(x_train) numpy.ndarray [4]: print(x_train.shape, y_train.shape, x_test.shape, y_test.shape) (50000, 32, 32, 3) (50000, 1) (10000, 32, 32, 3) (10000, 1) [5]: print(x_train.min()) 0 [6]: print(x_train.max()) 255</pre> <p>(A) 訓練集(x_train)資料集個數為 100000 筆 (B) 測試集(x_test)資料集個數為 10000 筆 (C) 訓練集(x_train)是 Pandas 資料框(DataFrame)物件 (D) 如果希望將訓練集(x_train)像素值轉換為[0, 1]的範圍，則可以輸入：x_train = x_train.astype('int32') / 255.0</p>

◆ 科目三：機器學習技術與應用

題號	答案	題目
1	A	<p>在 MapReduce 計算框架中，關於 Map 和 Reduce 所負責處理資料的問題，下列敘述何者正確？</p> <p>(A) Map：一組資料映射成另一組資料；Reduce：統合與歸納資料 (B) Map：地圖式的搜索資料；Reduce：統合與歸納資料 (C) Map：一組資料映射成另一組資料；Reduce：過濾不符合的資料 (D) Map：一組資料映射成另一組資料；Reduce：生成更多的資料</p>
2	B	<p>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？</p> <p>(A) R-CNN (B) Inception (C) ResNet (D) VGG19</p>
3	B	<p>當模型的訓練誤差 (Training Error) 低、但測試誤差 (Test Error) 很大時，這通常是在訓練過程中產生下列哪一種情況？</p> <p>(A) 模型的泛化能力強 (B) 模型出現過度擬合 (Overfitting) (C) 模型出現欠擬合 (Underfitting) (D) 訓練資料和測試資料之間沒有相關性</p>
4	A	<p>下列哪一種指標通常用於評估迴歸模型的效能？</p> <p>(A) R^2 (B) F1-分數 (C) 曲線下面積 (AUC) (D) Precision</p>
5	A	<p>近年來，深度學習研究與應用蓬勃發展，但數據本身可能存在什麼潛在問題？</p> <p>(A) 數據標註品質鮮少被討論，但它卻直接影響模型性能 (B) 數據品質是完美可信賴的 (C) 大部分情況下，數據不存在類別不平衡問題 (D) 數據不需要領域知識的輔助</p>
6	C	<p>在分類任務中，深度學習模型通常搭配哪一種輸出函數？</p> <p>(A) Tanh (B) ReLU (C) Sigmoid 或 Softmax (D) Dropout</p>

題號	答案	題目
7	C	<p>下列哪一種學習任務不適合使用監督式學習方法處理？</p> <p>(A) 客戶信用風險分類 (B) 預測未來銷售額 (C) 找出資料中的潛在群集 (D) 判斷影像是否為貓或狗</p>
8	B	<p>在神經網路中，前向傳播(Forward Propagation)主要依賴下列哪一種數學操作？</p> <p>(A) 機率積分 (B) 矩陣乘法與向量內積 (C) 對數變換 (D) 條件機率推論</p>
9	B	<p>特徵縮放(Feature Scaling)中，下列何者為標準化(Standardization)的主要作用？</p> <p>(A) 將數據範圍限制在 0 到 1 且標準差-1 (B) 使數據平均值為 0 且標準差為 1 (C) 移除數據中的異常值 (D) 增加特徵間的相關性</p>
10	B	<p>關於準確率 (Accuracy) 的計算方式，下列何者正確？</p> <p>(A) $TP / (TP+FP)$ (B) $(TP+TN) / (TP+FP+TN+FN)$ (C) $TN / (TN+FP)$ (D) $TP / (TP+FN)$</p>
11	D	<p>關於損失函數 (Loss Function) 的主要功能，下列何者正確？</p> <p>(A) 記錄模型預測歷史 (B) 計算模型結構複雜度 (C) 控制模型的學習率 (D) 衡量模型預測與真實值之間的誤差</p>
12	B	<p>關於歐盟《一般資料保護規則 (GDPR)》，所謂的被遺忘權 (Right to be Forgotten) 主要賦予資料主體哪一項權利？</p> <p>(A) 要求平台永久備份其個人資料以防遺失 (B) 在符合條件下請求刪除其個人資料 (C) 將個人資料轉換為匿名格式保存 (D) 限制企業將資料輸出至境外伺服器</p>
13	A	<p>在進行模型訓練前，若針對資料中不同群體（例如分類標籤）之間樣本數量不平衡的情況進行比例調整，此方法通常屬於下列哪一種技</p>

題號	答案	題目
		<p>術？</p> <p>(A) 資料重抽樣 (B) 特徵選擇 (C) 模型正則化 (D) 超參數調整</p>
14	B	<p>在優化器中，哪一個方法會自動調整每個參數的學習率，特別適用於稀疏資料？</p> <p>(A) Momentum (B) Adagrad (C) Adam (D) SGD</p>
	A	<p>某零售公司希望利用顧客的年齡與每月消費金額，預測顧客是否為高價值顧客。提供相關資料 data.csv，包含欄位 Age、Spending、HighValue。請將下列程式碼片段依正確順序排序，以完成模型的建立與預測。</p> <p>a. <pre>from sklearn.model_selection import train_test_split X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)</pre></p> <p>b. <pre>from sklearn.linear_model import LogisticRegression model = LogisticRegression() model.fit(X_train, y_train)</pre></p> <p>c. <pre>import pandas as pd data = pd.read_csv("data.csv") X = data[['Age', 'Spending']] y = data['HighValue']</pre></p> <p>d. <pre>y_pred = model.predict(X_test) print("Predictions:", y_pred[:5])</pre></p> <p>(A) c → a → b → d (B) a → c → b → d (C) c → b → a → d (D) b → a → c → d</p>