

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
B	1. 某零售企業建立一個銷售預測模型，希望評估該模型在不同月份的新資料上，是否仍能維持穩定的預測表現。資料科學團隊計畫利用統計方法檢驗模型對未觀察資料的適應能力與泛化效果。下列哪一種方法最適合用於此目的？ (A)F 檢定 (F-test); (B)交叉驗證 (Cross-Validation); (C)配對樣本 t 檢定 (Paired-sample t-test); (D)卡方檢定 (Chi-square Test)
C	2. 在建立迴歸或分類模型時，若希望避免模型過度擬合 (Overfitting)，可透過加入正則化項以限制模型的複雜度。其中，L1 正則化 (Lasso) 的主要效果為何？ (A)增加模型參數的數量，以提升表現靈活度； (B)強化梯度穩定性，避免參數更新過度震盪； (C)產生稀疏模型 (Sparse Model)，使部分參數權重收斂為零； (D)提高學習率 (Learning Rate)，加速模型收斂速度
C	3. 在訓練非線性模型時，若目標函數為非凸函數 (Non-convex Function)，演算法在參數更新過程中可能出現多個極值點，導致最佳化結果不穩定。請問此時最可能發生下列哪一種情況？ (A)梯度消失； (B)資料過少； (C)局部最優解； (D)過擬合
B	4. 在執行 DBSCAN (Density-Based Spatial Clustering of Applications with Noise) 群集分析時，若某資料點鄰域內的樣本數不足以形成核心點 (Core Point)，且該點未被任何核心點的鄰域所包含，也未與其他群集形成密度可達關係 (Density Reachability)，此資料點最終將被歸類為哪一種類型？ (A)鄰近點 (Neighbor Point); (B)雜訊點 (Noise Point); (C)邊界點 (Border Point); (D)潛在點 (Potential Point)
A	5. 某智慧製造公司開發一套影像辨識系統，用於自動檢測生產線上的瑕疵產品。系統採用卷積神經網路 (Convolutional Neural Network, CNN) 作為主要模型架構，其中第一層卷積層 (Convolutional Layer) 主要負責的功能

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	能為下列何者？ (A)自動提取輸入影像中的局部特徵； (B)降低影像維度以加速運算效率； (C)增加神經元與參數數量以提升模型容量； (D)整合所有特徵並輸出最終分類結果
C	6. 某智慧城市團隊開發一套交通監控系統，用於即時辨識路口監視器影像中的車輛與行人。團隊比較後發現，卷積神經網路 (Convolutional Neural Network, CNN) 在訓練與推論效率上，明顯優於傳統的全連接神經網路 (Fully Connected Neural Network, FCNN)。請問下列何者為主要原因？ (A)CNN 能自動學習影像的旋轉與比例不變性； (B)CNN 可直接跳過人工特徵提取步驟進行分類； (C)CNN 透過區域感知 (Local Receptive Field) 與參數共享 (Parameter Sharing) 機制，降低模型參數量與運算複雜度； (D)CNN 捨棄激勵函數 (Activation Function)，以加快運算速度
A	7. 下列哪一種應用最適合採用長短期記憶網路 (Long Short-Term Memory, LSTM) 模型？ (A)預測未來七天的電力需求變化趨勢； (B)辨識監視影像中不同類別的物件； (C)將大量顧客資料依相似特徵自動分群； (D)將高維度的感測器資料壓縮成低維表示
D	8. 資訊增益 (Information Gain) 常用於衡量特徵對分類結果的不確定性貢獻程度，並據以進行特徵選擇。此方法主要應用於下列哪一類模型架構中？ (A)使用 L1 正則化進行特徵篩選的線性模型； (B)利用激活函數 (Activation Function) 進行特徵擷取的深度神經網路； (C)透過核函數 (Kernel Function) 將特徵映射至高維空間的分類模型； (D)透過遞迴分裂方式建立分類規則的決策樹模型
A	9. 在建構以距離為基礎的機器學習模型（如 KNN、SVM）時，下列哪一項資料前處理方式最為關鍵？ (A)進行特徵縮放 (Feature Scaling)，使各特徵變數具有相似的數值範圍； (B)將連續型特徵變數轉換為類別型變數； (C)以平均值或中位數進行缺失值補齊；

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	(D)進行隨機抽樣以平衡資料筆數
C	10. 下列哪一種應用情境最適合導入 AutoML，以提升模型開發效率？ (A)公司已有完整的 MLOps 平台與資深資料科學團隊，模型更新採固定流程； (B)製造部門的生產良率模型已長期穩定運作，只需定期調整參數； (C)行銷部門希望在短時間內比較多種顧客流失預測模型，缺乏專職工程師與時間進行手動建模； (D)財務部門正在開發高度客製化的信用風險評估模型，需要精細控制特徵工程與演算法細節
D	11. 相較於 Grid Search，Random Search 在超參數調整上具備哪一項主要優勢？ (A)可自動產生模型架構； (B)可使用更大的訓練集； (C)避免模型過擬合； (D)能更有效率搜尋高維參數空間
C	12. 某智慧製造公司開發一套設備故障預測系統，利用感測器資料訓練深度神經網路 (Deep Neural Network, DNN) 模型，以提前偵測異常運作跡象。在訓練過程中，團隊發現模型收斂速度不穩定：有時過快導致過擬合，有時又遲遲無法達到最佳準確率。開發團隊可以藉由調整下列哪一項超參數 (Hyperparameter) 以改善此問題？ (A)每個神經元的輸出結果； (B)損失函數 (Loss Function) 在訓練過程中的梯度變化值 (Gradient)； (C)學習率 (Learning Rate)，控制模型權重更新的速度； (D)模型在訓練後產生的權重值
B	13. 標籤偏差(Label Bias)通常是因為什麼原因造成？ (A)訓練資料量過大； (B)標記資料本身帶有主觀偏見； (C)模型結構設計不當； (D)特徵數量設定過多
C	14. 下列哪一種 AI 應用情境中，模型的可解釋性 (Explainability) 最為關鍵？ (A)電商平台利用深度學習模型預測用戶的下一次購買時間，以優化推播行銷策略； (B)新創公司使用機器學習演算法自動調整廣告出價策略，以提升點擊轉換

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	率； (C)醫院導入 AI 模型分析病患影像並給出腫瘤惡性可能性，作為臨床醫師診斷依據； (D)銀行導入 AI 模型預測客戶流失率，並自動推薦留客優惠方案
B	15. 在線性迴歸模型中，若 $R^2$ 值為 0.85，其意義為何？ (A)模型準確率為 85%； (B)85%的變異可被模型解釋； (C)預測誤差為 15%； (D)模型有 85%的信心水準
A	16. 在二元分類問題中，若精確率 (Precision) 為 0.8，召回率 (Recall) 為 0.6，則 F1 分數 (F1 Score) 為何？ (A)0.686； (B)0.700； (C)0.720； (D)0.750
B	17. 下列哪一種優化演算法內建動量 (Momentum) 的設計機制？ (A)SGD+Momentum； (B)Adam； (C)RMSProp； (D)Adagrad
A	18. 下列何者最能同時反映 XGBoost (eXtreme Gradient Boosting) 相較於傳統梯度提升決策樹 (Gradient Boosting Decision Tree, GBDT) 的主要技術改進？ (A)引入正則化項 (Regularization) 以抑制過擬合，並支援缺失值自動處理與並行化訓練； (B)改以隨機森林 (Random Forest) 架構取代樹模型以提升準確率； (C)以類神經網路 (Neural Network) 取代弱分類器 (Weak Learners)； (D)採用批次正規化 (Batch Normalization) 技術提升模型穩定性
C	19. 某醫療機構開發疾病早期偵測模型，正樣本（確診病例）僅佔 3%。在模型訓練與評估過程中，下列哪一種作法最不適合用於提升對少數類病例的預測能力？ (A)使用 SMOTE 過採樣； (B)調整類別權重； (C)使用準確率 (Accuracy) 作為評估指標；

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	(D)欠採樣多數類(Undersampling the majority class)
C	20. 某電子商務公司為開發商品評論情感分析模型，希望模型能捕捉評論中不同特徵之間的關聯影響，例如「商品價格」與「顧客滿意度」的互動效果。下列哪一種特徵工程設計方式最適合用於建立互動特徵（Interaction Features）？ (A)將單一特徵取平方； (B)對所有特徵進行對數轉換； (C)將兩個或多個特徵進行乘積或交互組合； (D)對特徵進行標準化
C	21. 某語音辨識系統開發團隊採用 Transformer 架構，為了讓模型能同時理解語音片段中的發音特徵、語速變化與語意脈絡等多層次資訊，團隊在設計中導入了多頭注意力（Multi-head Attention）機制。請問下列何者為此機制的主要優點？ (A)減少模型參數量以降低訓練成本； (B)加速整體注意力計算過程； (C)從不同表示子空間（Representation Subspaces）同時捕捉多樣化關聯資訊； (D)避免梯度消失（Gradient Vanishing）問題
B	22. 某電商平台希望預測顧客是否會購買特定商品。系統蒐集顧客的瀏覽紀錄、停留時間、商品類別偏好與過去購買行為，並以此推估「在觀察到這些行為特徵的情況下，該顧客會購買的機率」。若模型採用貝氏定理（Bayes' Theorem）進行推論，下列敘述何者最符合其核心運作機制？ (A)根據歷史樣本自動分群，找出行為相似的顧客群； (B)以條件機率方式計算顧客屬於「會購買」或「不會購買」的分類機率； (C)以最小平方誤差（Mean Squared Error）為損失函數，預測顧客的購買金額； (D)依據回饋信號（Feedback Signal）透過強化學習（Reinforcement Learning）動態調整推薦策略
A	23. 一家再生能源公司希望預測未來三個月太陽能發電量的波動範圍。由於氣候條件具有高度隨機性，且輸入變數（如日照時數、雲量、溫度）之間存在不確定關係，工程團隊決定以隨機抽樣方式模擬多種可能情境，以估算整體發電量的機率分佈與風險區間。請問此時所採用的技術最符合下列哪一種方法？ (A)蒙地卡羅方法（Monte Carlo Method）；

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	(B)K-means 聚類 (K-means Clustering)； (C)支持向量迴歸 (Support Vector Regression, SVR)； (D)特徵選取 (Feature Selection)
C	24. 某房地產公司利用多元迴歸模型 (Multiple Regression Model) 預測房價，並繪製殘差圖 (Residual Plot) 檢查模型品質。結果顯示部分資料點的殘差極大，且在高價區樣本中出現系統性彎曲分佈現象。根據此觀察，下列何者為最可能的正確解釋？ (A)模型過度擬合 (Overfitting)，導致在訓練資料上表現過好、泛化能力不足； (B)模型特徵數量不足，導致欠擬合 (Underfitting)； (C)模型存在異常值 (Outlier) 或非線性關係，違反迴歸假設； (D)殘差圖呈現隨機分佈，表示模型已完全符合所有假設
A	25. 某金融機構正在建立傳統信用評分卡模型，採用邏輯迴歸 (Logistic Regression) 作為建模方法，並依循監理機關建議的標準化流程進行模型開發。下列哪一項不是傳統信用評分卡模型開發流程中的常見步驟？ (A)使用生成式模型進行特徵學習； (B)進行特徵選擇與多重共線性 (Multicollinearity) 分析； (C)進行分箱 (Binning) 與資訊值 (Information Value, IV) 檢定； (D)使用樣本穩定性指標 (Population Stability Index, PSI) 檢驗模型穩定性
D	26. 在防止監督式學習模型過擬合 (Overfitting) 時，下列哪一種策略不屬於降低模型複雜度或限制學習能力的作法？ (A)採用 L1 或 L2 正則化； (B)在訓練過程中使用 Dropout 技術； (C)採取早期停止 (Early Stopping) 機制； (D)擴增輸入特徵變數以提升模型表達能力
D	27. 某智慧製造團隊在開發瑕疵影像檢測模型時，發現使用線性激活函數 (Activation Function) 後，模型的訓練準確率長期停滯，懷疑模型無法學習到足夠複雜的特徵表達。若要改善此問題，下列哪一項調整方案最為合適？ (A)增加卷積層 (Convolutional Layer) 數量，使網路更深以強化特徵提取； (B)將輸入影像先進行灰階化處理，降低運算量； (C)使用 Sigmoid 激活函數，以將輸出壓縮至 [0, 1] 範圍；

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	(D)改用 ReLU (Rectified Linear Unit) 激活函數，以引入非線性並提升模型表達能力
C	28. 一家零售電商公司希望建立顧客流失預測模型，用以判斷哪些會員可能在三個月內不再消費。團隊以去年會員資料進行訓練，並僅採用「曾經購買三次以上」的活躍顧客紀錄作為樣本。模型上線後，對整體會員進行預測時，發現模型對於新註冊會員與低消費會員的預測準確率明顯偏低。下列何者為造成此現象最可能的原因？ (A)特徵設計未排除與會員忠誠度高度相關的變數，導致特徵偏差 (Feature Bias)； (B)標記 (Label) 由人工標註，導致標籤偏差 (Label Bias)； (C)訓練樣本僅涵蓋高活躍顧客，造成取樣偏差 (Sampling Bias)； (D)模型未進行超參數調整，導致過擬合 (Overfitting)
D	29. 在工業設備故障預測專案中，模型訓練與超參數調整均依賴於一段歷史數據作為驗證集。然而，隨著設備運行環境與工況條件的變化，原有驗證集已無法充分反映現況，導致模型在實際部署後的預測準確率逐漸下降。下列哪一種策略最能有效提升模型在長期運行環境中的穩健性與泛化能力？ (A)固定驗證集內容，並透過模型正則化技巧（如 L2 正則化）強化模型泛化； (B)將全部歷史資料納入訓練，不使用驗證集，依靠早期停止 (Early Stopping) 控制訓練； (C)簡化模型架構，減少模型參數數量以降低過擬合風險； (D)採用時間序列交叉驗證 (Time Series Cross Validation) 或滑動視窗驗證 (Rolling Window Validation) 方法，動態更新驗證資料以適應時間演進
C	30. 某情感分析模型在英文資料集上取得 macro F1-score = 0.91。當該模型部署於西班牙文資料集時，F1-score 驟降至 0.58。下列哪一項解釋最合理，且與 F1-score 變化相關？ (A)macro F1-score 本身波動性高，建議改用 micro-average F1-score 評估； (B)模型在西班牙文語料上過度擬合，導致評估結果偏高； (C)語言轉移造成召回率 (Recall) 下降，模型無法正確辨識關鍵情緒詞彙； (D)以均方誤差 (MSE) 取代 F1-score 評估可獲得更準確的結果
B	31. 某能源公司利用歷史氣象與用電資料，開發長期電力需求預測模型，採用

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	<p>深度神經網路架構進行訓練。在訓練過程中，模型在訓練集上的損失值持續下降，但在驗證集上，損失在第 80 輪後開始波動，呈現週期性上升與下降。團隊懷疑模型受到季節性資料波動與隨機噪音影響，導致驗證損失難以穩定收斂。若要在此情境下合理運用早期停止法（Early Stopping）以確保模型具最佳泛化能力，下列哪一項策略最為適當？</p> <p>(A)直接根據訓練集損失最低點停止訓練，以確保模型充分擬合所有樣本； (B)監控驗證集損失並設定適度的耐心值（Patience），在連續多輪未改善後再停止訓練； (C)改以測試集損失作為早停依據，以提升模型最終評估一致性； (D)將所有資料重新合併後訓練至收斂，避免因資料分割導致評估波動</p>
D	<p>32. 某電信公司開發客戶流失預測模型，使用大量顧客行為特徵，例如通話時長、上網頻率、帳單金額、客服聯絡次數等。在訓練過程中，團隊發現部分特徵彼此高度相關，但同時也懷疑有些特徵對流失預測的貢獻度有限。若希望模型在避免過擬合（Overfitting）的同時，能自動篩選出較具代表性的特徵，採用下列哪一種方法最為合適？</p> <p>(A)使用早期停止法（Early Stopping）控制訓練回合數，避免過擬合（Overfitting）； (B)同時移除多重共線性特徵並採用 L2 正則化（Ridge），以確保模型穩定收斂； (C)僅使用 L2 正則化（Ridge），抑制所有權重幅度但保留全部特徵； (D)採用 L1 正則化（Lasso），透過懲罰項使部分特徵係數縮為 0</p>
B	<p>33. 某資料科學團隊正在開發一個客戶相似度比對系統，用於計算所有客戶之間的相似度分數。若系統需逐一比對每一位客戶與其他所有客戶的資料組合，此時演算法的時間複雜度最可能為哪一種？其代表意義為何？</p> <p>(A)<math>O(n)</math> — 執行時間與資料量成線性關係； (B)<math>O(n^2)</math> — 執行時間與資料量平方成正比； (C)<math>O(1)</math> — 執行時間固定不變； (D)<math>O(\log n)</math> — 執行時間與資料量呈對數成長關係</p>
D	<p>34. 某醫療人工智慧團隊正在開發心臟病風險預測模型，資料量僅有 150 筆，其中陽性個案不到 8%。由於樣本數稀少且類別分布極不平衡，團隊希望在有限資料下，仍能準確評估模型在不同資料上的表現穩定性，同時避免訓練資料被過度切分而影響模型效能。若團隊希望在有限樣本下，同時兼顧資料的利用率與各類別在驗證折中的比例一致性，最適合採用下列哪一種交叉驗證方法？</p>

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	(A)5-Fold 交叉驗證 (5-Fold Cross Validation); (B)留一法交叉驗證 (Leave-One-Out Cross Validation); (C)隨機交叉驗證 (Random Cross Validation); (D)分層留一法交叉驗證 (Stratified Leave-One-Out Cross Validation)
A	35. 某公司針對製程感測器資料進行主成分分析 (PCA)，經標準化與協方差矩陣分解後，得到三個主成分的特徵值如下： $\lambda_1=6.0$ , $\lambda_2=3.0$ , $\lambda_3=1.0$ 。若團隊決定僅保留能解釋至少 80% 總變異量的主成分，以進行後續模型建構，下列哪一項敘述最合理且數據解讀正確？ (A)前兩個主成分合計解釋 90% 的總變異量，因此可安全降維至二維，且仍保留大部分資訊； (B)第一主成分解釋 60% 的變異量，表示資料結構呈現明顯線性關係，僅保留一維即可避免過擬合； (C)雖然前兩個主成分可解釋超過 80% 變異量，但第二主成分貢獻仍高達 30%，不宜捨棄第三主成分； (D)三個特徵值相差不大，顯示各主成分變異均衡，降維可能導致資訊損失
D	36. 某銀行計畫與多家合作機構共同訓練一個 AI 信用風險預測模型，為避免客戶交易資料在傳輸與運算過程中外洩，技術團隊評估使用同態加密 (Homomorphic Encryption) 技術。下列何者最能正確描述同態加密在此應用中的關鍵特性？ (A)系統以隨機雜訊 (Noise) 干擾輸出，確保統計結果不洩漏個資； (B)各參與銀行透過安全通道交換私鑰，確保模型參數一致； (C)將原始資料壓縮並同時加密，以減少加密後資料量與運算時間； (D)資料在加密狀態下仍可進行數值運算，模型訓練可於未解密資料上完成
B	37. 某跨銀行風控平台希望整合多家銀行的用戶行為資料，用於訓練信用風險預測模型。由於競爭與法規限制，各銀行僅願意提供加密後資料，且資料在任何時間不得被平台解密。同時，平台需建立安全通訊協議以確保資料在傳輸過程未被竊改或重放。 下列哪一組技術最能完整對應上述需求？ (A)對稱加密 (Symmetric Encryption) + 單向雜湊 (Hash Function) + 非對稱加密 (Asymmetric Encryption) + 差分隱私 (Differential Privacy)； (B)同態加密 (Homomorphic Encryption) + 非對稱加密 (Asymmetric Encryption) + 單向雜湊 (One-way Hash Function) + 對稱加密

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	(Symmetric Encryption)； (C) 差分隱私 (Differential Privacy) + 對稱加密 (Symmetric Encryption) + 同態加密 (Homomorphic Encryption) + 數位簽章 (Digital Signature)； (D) 同態加密 (Homomorphic Encryption) + 安全多方計算 (Secure Multi-party Computation, MPC) + 雜湊函數 (Hash Function) + 對稱加密 (Symmetric Encryption)
B	38. 附圖程式碼所計算的是哪一類型的評估指標？ <pre>def metric(y_true, y_pred):     return sum((y_true - y_pred) ** 2) / len(y_true)</pre> (A)MAE；(B)MSE；(C)RMSE；(D) $R^2$
C	39. 附圖程式碼實現的是哪一種正則化技術？ <pre>def forward(x, p, training=True):     if training:         mask = np.random.binomial(1, p, size=x.shape)         return x * mask / p     else:         return x</pre> (A)L1 正則化；(B)L2 正則化；(C)Dropout；(D)Batch Normalization
C	40. 依據附圖程式碼進行資料處理，下列何者正確？ <pre>import numpy as np  v1 = np.array([1, 2, 3]) v2 = np.array([4, 5, 6]) A = np.array([[1, 2], [3, 4]])</pre> (A)np.linalg.inv(A) 計算矩陣 A 的行列式； (B)v1 * v2 結果為 array([5, 7, 9])； (C)np.dot(v1, v2) 結果為 np.int64(32)； (D)np.linalg.eig(A) 計算矩陣 A 的反矩陣
D	41. 考慮擲出骰子並採用 Monte Carlo 方法估算條件機率，參考附圖程式碼。

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目																											
	<pre>import numpy as np  np.random.seed(123) n = 100000 dice_rolls = np.random.randint(1, 7, size=n)  A = (dice_rolls % 2 == 0) B = (dice_rolls &gt; 3) A_and_B = A &amp; B</pre> <p>事件 A：擲出偶數 事件 B：擲出大於 3 請問下列何者為條件機率 <math>P(A   B)</math> 的正確值？</p> <p>(A)A_and_B.sum() / (A.sum() * B.sum()); (B)A_and_B.sum() / (A.sum() + B.sum()); (C)A_and_B.sum() / A.sum(); (D)A_and_B.sum() / B.sum()</p>																											
	<p>VGG16 是由牛津大學 Visual Geometry Group(VGG)在 2014 年提出的經典卷積神經網路 (Convolutional Neural Network, CNN) 架構。該模型以簡潔且規則的層堆疊設計聞名，廣泛應用於影像分類、特徵提取及遷移學習等任務。附圖程式碼載入了預訓練的 VGG16 模型，並輸出其完整層級結構及參數統計摘要 (如附表)。請根據此資訊回答 42~45 題。</p> <pre>from torchsummary import summary from torchvision import models model = models.vgg16(weights='IMAGENET1K_V1') summary(model, (3,150,150))</pre> <hr/> <table><thead><tr><th>Layer (type)</th><th>Output Shape</th><th>Param #</th></tr></thead><tbody><tr><td>Conv2d-1</td><td>[ -1, 64, 150, 150]</td><td>1,792</td></tr><tr><td>ReLU-2</td><td>[ -1, 64, 150, 150]</td><td>0</td></tr><tr><td>Conv2d-3</td><td>[ -1, 64, 150, 150]</td><td>36,928</td></tr><tr><td>ReLU-4</td><td>[ -1, 64, 150, 150]</td><td>0</td></tr><tr><td>MaxPool2d-5</td><td>[ -1, 64, 75, 75]</td><td>0</td></tr><tr><td>Conv2d-6</td><td>[ -1, 128, 75, 75]</td><td>73,856</td></tr><tr><td>ReLU-7</td><td>[ -1, 128, 75, 75]</td><td>0</td></tr><tr><td>Conv2d-8</td><td>[ -1, 128, 75, 75]</td><td>147,584</td></tr></tbody></table>	Layer (type)	Output Shape	Param #	Conv2d-1	[ -1, 64, 150, 150]	1,792	ReLU-2	[ -1, 64, 150, 150]	0	Conv2d-3	[ -1, 64, 150, 150]	36,928	ReLU-4	[ -1, 64, 150, 150]	0	MaxPool2d-5	[ -1, 64, 75, 75]	0	Conv2d-6	[ -1, 128, 75, 75]	73,856	ReLU-7	[ -1, 128, 75, 75]	0	Conv2d-8	[ -1, 128, 75, 75]	147,584
Layer (type)	Output Shape	Param #																										
Conv2d-1	[ -1, 64, 150, 150]	1,792																										
ReLU-2	[ -1, 64, 150, 150]	0																										
Conv2d-3	[ -1, 64, 150, 150]	36,928																										
ReLU-4	[ -1, 64, 150, 150]	0																										
MaxPool2d-5	[ -1, 64, 75, 75]	0																										
Conv2d-6	[ -1, 128, 75, 75]	73,856																										
ReLU-7	[ -1, 128, 75, 75]	0																										
Conv2d-8	[ -1, 128, 75, 75]	147,584																										

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目		
ReLU-9	[ -1, 128, 75, 75]	0	
MaxPool2d-10	[ -1, 128, 37, 37]	0	
Conv2d-11	[ -1, 256, 37, 37]	295,168	
ReLU-12	[ -1, 256, 37, 37]	0	
Conv2d-13	[ -1, 256, 37, 37]	590,080	
ReLU-14	[ -1, 256, 37, 37]	0	
Conv2d-15	[ -1, 256, 37, 37]	590,080	
ReLU-16	[ -1, 256, 37, 37]	0	
MaxPool2d-17	[ -1, 256, 18, 18]	0	
Conv2d-18	[ -1, 512, 18, 18]	1,180,160	
ReLU-19	[ -1, 512, 18, 18]	0	
Conv2d-20	[ -1, 512, 18, 18]	2,359,808	
ReLU-21	[ -1, 512, 18, 18]	0	
Conv2d-22	[ -1, 512, 18, 18]	2,359,808	
ReLU-23	[ -1, 512, 18, 18]	0	
MaxPool2d-24	[ -1, 512, 9, 9]	0	
Conv2d-25	[ -1, 512, 9, 9]	2,359,808	
ReLU-26	[ -1, 512, 9, 9]	0	
Conv2d-27	[ -1, 512, 9, 9]	2,359,808	
ReLU-28	[ -1, 512, 9, 9]	0	
Conv2d-29	[ -1, 512, 9, 9]	2,359,808	
ReLU-30	[ -1, 512, 9, 9]	0	
MaxPool2d-31	[ -1, 512, 4, 4]	0	
AdaptiveAvgPool2d-32	[ -1, 512, 7, 7]	0	
Linear-33	[ -1, 4096]	102,764,544	
ReLU-34	[ -1, 4096]	0	
Dropout-35	[ -1, 4096]	0	
Linear-36	[ -1, 4096]	16,781,312	
ReLU-37	[ -1, 4096]	0	
Dropout-38	[ -1, 4096]	0	
Linear-39	[ -1, 1000]	4,097,000	
<hr/>			
Total params: 138,357,544			
Trainable params: 138,357,544			
Non-trainable params: 0			
<hr/>			
Input size (MB): 0.26			

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	Forward/backward pass size (MB): 96.93 Params size (MB): 527.79 Estimated Total Size (MB): 624.98 -----
B	42. 在深度神經網路中，不同層的參數量 (parameter count) 差異極大。有些層雖然數量少但計算量大，有些則相反。了解參數分佈情形，有助於模型壓縮與遷移學習設計。請問在 VGG16 中，下列何者的參數量最多？ (A)卷積層(Conv2d)； (B)全連接層(Linear)； (C)ReLU 激活函數； (D)池化層(MaxPool2d, AdaptiveAvgPool2d)
A	43. 在神經網路中，了解各層的運算量分佈，有助於模型壓縮與硬體加速的策略設計。請問在 VGG16 中，下列何者運算量(FLOPs)最多？ (A)卷積層(Conv2d)； (B)全連接層(Linear)； (C)ReLU 激活函數； (D)池化層(MaxPool2d, AdaptiveAvgPool2d)
D	44. VGG16 層數深且結構規則，由多層卷積、池化及全連接層組成。了解各層的輸入/輸出維度、參數量及記憶體需求，有助於掌握 CNN 模型的組成邏輯與實作技巧。根據 VGG16 的模型架構，下列敘述何者正確？ (A)AdaptiveAvgPool2d 的輸出會被攤平後傳入第一個全連接層；由於前一層池化輸出空間為 4×4，所以第一個線性層的輸入維度是 $512 \times 4 \times 4 = 8192$ ； (B)Linear-33 (第一個全連接層) 報出的 102,764,544 參數只包含權重，偏差 (bias) 沒有算在內； (C)根據列出的「Estimated Total Size (MB) = 624.98」，表示訓練此模型只需大約 625MB 的 GPU 記憶體 (包含所有 optimizer state 與梯度)，所以一張 1 GB 的 GPU 就足夠訓練； (D) VGG16 包含 13 層卷積層 (conv) 與 3 層全連接層 (FC)，總參數數目約為 138,357,544 (約 138.36M)
B	45. 在實務應用中，我們常使用遷移學習 (transfer learning) 技巧，即載入預訓練模型 (如 VGG16)，凍結部分層的參數，只針對特定任務重新訓練最後幾層，這種做法可節省訓練時間並提升模型效能。假設你要對 VGG16 進行遷移學習 (transfer learning)，希望凍結卷積層的參數，只訓練最後全連

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

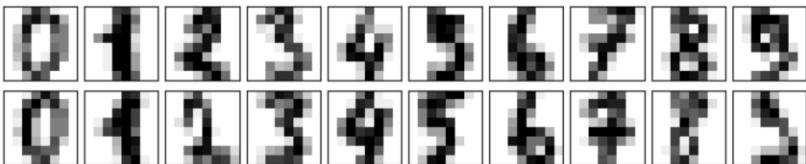
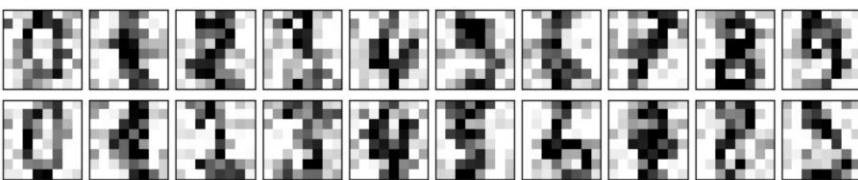
考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	<p>接層(classifier)。下列哪段程式碼寫法正確？</p> <p>(A)</p> <pre>import torch import torchvision.models as models  model = models.vgg16(pretrained=True) for param in model.parameters():     param.requires_grad = False model.classifier[6] = torch.nn.Linear(4096, 10);</pre> <p>(B)</p> <pre>import torch import torchvision.models as models  model = models.vgg16(pretrained=True) for param in model.features.parameters():     param.requires_grad = False model.classifier[6] = torch.nn.Linear(4096, 10);</pre> <p>(C)</p> <pre>import torch import torchvision.models as models  model = models.vgg16(pretrained=True) for param in model.classifier.parameters():     param.requires_grad = False model.classifier[6] = torch.nn.Linear(4096, 10);</pre> <p>(D)</p> <pre>import torch import torchvision.models as models  model = models.vgg16(pretrained=True) model.requires_grad = False model.classifier[6] = torch.nn.Linear(4096, 10)</pre>
	<p>在郵遞區號自動辨識的研究中，研究人員收集了一份手寫數字影像資料集，每一張影像為 8x8 的灰階圖片，共包含多個手寫數字樣本。這份資料集來自 UCI Machine Learning Repository，常被用於數字辨識與機器學習方法的教學與實驗。</p> <p>在過程中，研究人員發現資料中可能存在雜訊，例如筆跡模糊或影像中附加的干擾點，這會影響後續分類模型的效能。因此，他們希望透過資料降噪的方法，提升後續分類的準確度。同時，他們也想透過 KNN (K-Nearest Neighbors) 搭配交叉驗證來評估模型表現，確保模型在不同資料切割下都能有穩定的預測能力。</p> <pre>from sklearn.datasets import load_digits digits = load_digits() digits.data.shape</pre> <p>(1797, 64)</p> <p>部分樣本經繪製後的外觀如下圖所示：</p>

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	 <p>根據這份資料來回答 46~47 題。</p>
B	<p>46.</p> <pre>import numpy as np noisy = np.random.normal(digits.data, 4)</pre> <p>假設研究人員已將含有雜訊的手寫數字影像存放在變數 noisy 中。部分資料經視覺化後的外觀如下：</p>  <pre>from sklearn.decomposition import PCA          #程式碼A pca = PCA()                                     #程式碼B pca.fit(noisy)                                  #程式碼C components = pca.transform(noisy)                #程式碼D filtered = pca.inverse_transform(components)      #程式碼E</pre> <p>他們嘗試使用 PCA 進行降噪，並希望能保留影像的主要特徵，同時去除影像中的雜訊。然而，當程式執行後，觀察到影像仍然含有明顯的雜訊。研究人員懷疑是程式中某個步驟的設定不正確，導致 PCA 沒有發揮降噪的作用，需要修改程式碼才能讓降噪有效。請問哪一段程式碼需要修改，才能讓 PCA 對 noisy 影像有效去噪？</p> <p>(A) 程式碼 A； (B) 程式碼 B； (C) 程式碼 C； (D) 程式碼 D</p>
B	<p>47.</p> <p>研究人員在對 digits 資料集進行分類時，決定使用 KNN 並搭配交叉驗證來評估模型準確率。他們撰寫了四組不同的程式碼來進行 KNN 訓練與交叉驗證，但不確定哪幾組程式碼能正確執行並輸出準確率。每組程式碼在資料切割、模型訓練、交叉驗證函數的使用上略有差異，研究人員希望找出可以正確完成任務的程式碼組合，以確保模型評估的可靠性。</p> <pre># 程式碼 A : from sklearn.model_selection import StratifiedKFold, cross_val_score from sklearn.neighbors import KNeighborsClassifier  X, y = digits.data, digits.target model = KNeighborsClassifier(n_neighbors=3) cv = StratifiedKFold(n_splits=5, shuffle=True) scores = cross_val_score(model, X, y, cv=cv, scoring="accuracy") print(scores.mean())</pre>

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	<pre># 程式碼 B : from sklearn.model_selection import StratifiedKFold, cross_val_score from sklearn.neighbors import KNeighborsClassifier  X, y = digits.data, digits.target model = KNeighborsClassifier(n_neighbors=3) cv = StratifiedKFold(n_splits=5, shuffle=True) scores = cross_val_score(model, X, y, cv=cv, scoring="f1") print(scores.mean())  # 程式碼 C : from sklearn.model_selection import StratifiedKFold, cross_val_score from sklearn.neighbors import KNeighborsClassifier  X, y = digits.data, digits.target model = KNeighborsClassifier(n_neighbors=3) scores = cross_val_score(model, X, y, cv=5, scoring="accuracy") print(scores.mean())  # 程式碼 D : from sklearn.model_selection import StratifiedKFold, cross_val_score from sklearn.neighbors import KNeighborsClassifier  X, y = digits.data, digits.target model = KNeighborsClassifier(n_neighbors=3) scores = cross_val_score(model, X, y, cv=5, scoring="f1") print(scores.mean())</pre> <p>請問哪幾組程式碼能正確使用 KNN 搭配交叉驗證，對 digits 資料集進行訓練並輸出準確率？</p> <p>(A) 程式碼 A、程式碼 B、程式碼 C、程式碼 D； (B) 程式碼 A、程式碼 C； (C) 程式碼 A、程式碼 B； (D) 程式碼 C、程式碼 D</p>
	<p>使用鐵達尼號(Titanic)資料集進行多層感知機(Multilayer Perceptron, MLP)分類預測分析，其中 survived 為反應變數(1 表示存活，0 表示死亡)。附圖程式碼展示資料載入與處理，請根據此資料回答第 48~50 題。</p> <pre>import numpy as np import pandas as pd from keras import Sequential from keras.layers import Input, Dense  np.random.seed(123) df_train = pd.read_csv("titanic_train.csv") df_test = pd.read_csv("titanic_test.csv")</pre> <p>下圖顯示資料集的前 5 筆資料。</p>

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	<pre>df_train.head()    pclass  sex    age  sibsp  parch     fare  embarked_C  embarked_Q  embarked_S  survived 0      1    1  29.0000      0      0  211.3375    False    False     True      1 1      1    0  0.9167      1      2  151.5500    False    False     True      1 2      1    1  2.0000      1      2  151.5500    False    False     True      0 3      1    0 30.0000      1      2  151.5500    False    False     True      0 4      1    0 48.0000      0      0  26.5500    False    False     True      1  df_test.head()    pclass  sex    age  sibsp  parch     fare  embarked_C  embarked_Q  embarked_S  survived 0      1    1 25.000000      1      2  151.5500    False    False     True      0 1      1    1 18.000000      1      0  227.5250    True     False    False      1 2      1    0 29.881135      0      0  25.9250    False    False     True      0 3      1    1 32.000000      0      0  76.2917    True     False    False      1 4      1    1 47.000000      1      1  52.5542    False    False     True      1  dataset_train = df_train.values.astype('float32') dataset_test = df_test.values.astype('float32')  X_train = dataset_train[:, 0:9] y_train = dataset_train[:, 9] X_test = dataset_test[:, 0:9] y_test = dataset_test[:, 9]</pre>
C	48. 參考下圖程式碼，下列何者正確？  <pre>X_train -= X_train.mean(axis=0) X_train /= X_train.std(axis=0) X_test -= X_test.mean(axis=0) X_test /= X_test.std(axis=0)</pre> <p>A : <math>X_{train} \leftarrow X_{train} - \text{mean}(axis=0)</math> 將每個訓練集特徵的平均值調整為 0 B : <math>X_{train} \leftarrow X_{train} / \text{std}(axis=0)</math> 將每個訓練集特徵的標準差調整為 0 C : <math>X_{train}</math> 處理結果會將資料壓縮到 0 和 1 之間 D : 標準化結果防止梯度爆炸或消失 E : 標準化是屬於特徵選擇(Feature Selection)方法 F : <math>X_{train}</math> 程式碼應修正為 <math>X_{train} = X_{train}.std(axis=0)</math>, <math>X_{test}</math> 程式碼應修正為 <math>X_{test} = X_{test}.std(axis=0)</math> (A)A、B、C、D；</p>

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

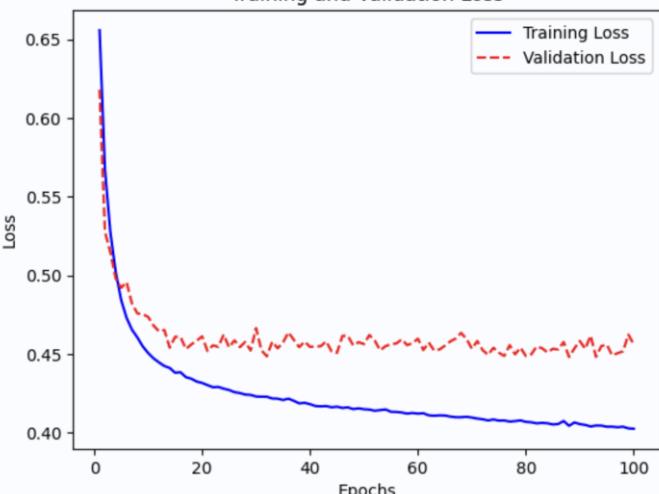
考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目												
	(B)A、E； (C)A、D； (D)A、C、F												
C	49. 參考下圖執行結果，下列何者正確？ <pre>model = Sequential() model.add(Input(shape=(X_train.shape[1],))) model.add(Dense(10, activation="relu")) model.add(Dense(10, activation="relu")) model.add(Dense(1, activation="sigmoid")) model.summary() model.compile(loss="binary_crossentropy", optimizer="adam", metrics=["accuracy"])</pre> <p>Model: "sequential"</p> <table border="1"><thead><tr><th>Layer (type)</th><th>Output Shape</th><th>Param #</th></tr></thead><tbody><tr><td>dense (Dense)</td><td>(None, 10)</td><td>空格1</td></tr><tr><td>dense_1 (Dense)</td><td>(None, 10)</td><td>空格2</td></tr><tr><td>dense_2 (Dense)</td><td>(None, 1)</td><td>11</td></tr></tbody></table> <p>(A)activation="relu" 其數學式為 <math>f(x) = \frac{1}{1+e^{-x}}</math>； (B)空格 1 值為 110，空格 2 值為 100； (C)空格 1 值為 100，空格 2 值為 110； (D)activation="sigmoid" 一般用於多類別分類預測模型</p>	Layer (type)	Output Shape	Param #	dense (Dense)	(None, 10)	空格1	dense_1 (Dense)	(None, 10)	空格2	dense_2 (Dense)	(None, 1)	11
Layer (type)	Output Shape	Param #											
dense (Dense)	(None, 10)	空格1											
dense_1 (Dense)	(None, 10)	空格2											
dense_2 (Dense)	(None, 1)	11											
C	50. 參考下圖執行結果，下列何者正確？ <pre>history = model.fit(X_train, y_train, validation_split=0.2, epochs=100, batch_size=10) loss, accuracy = model.evaluate(X_train, y_train, verbose=0)</pre> <pre>import matplotlib.pyplot as plt loss = history.history["loss"] epochs = range(1, len(loss)+1) val_loss = history.history["val_loss"] plt.plot(epochs, loss, 空格1, label="Training Loss") plt.plot(epochs, val_loss, 空格2, label="Validation Loss") plt.title("Training and Validation Loss") plt.xlabel("Epochs") plt.ylabel("Loss") plt.legend() plt.show()</pre>												

# 114 年第二次 AI 應用規劃師-中級能力鑑定【公告試題】

## 第三科：機器學習技術與應用

考試日期：114 年 11 月 08 日 試題公告日期：114 年 11 月 20 日

答案	題目
	<p style="text-align: center;"><b>Training and Validation Loss</b></p>  <p>A : 空格 1 須填入 "b-" B : 空格 2 須填入 "b--" C : 空格 1 須填入 "r-" D : 空格 2 須填入 "r--" E : 驗證損失明顯較訓練損失減少更明顯 (A)B、C； (B)A、C、D； (C)A、D； (D)C、D、E</p>

《以下空白》