iPAS AI 應用規劃師（初級）

科目一：人工智慧基礎概論

易讀版研讀講義 ＋ 考前重點 ＋ 100 題模擬測驗

依使用者提供之 iPAS 學習指引整理

使用方式：第一輪先讀「一頁速記」與各章重點；第二輪做 100 題；第三輪只複習錯題與最後速查表。

重要說明：原始文件定位為學習指引，並非正式教材或題庫，也不保證實際考題。本文檔將原文件內容改寫成較容易閱讀的考前筆記，並將練習題型轉為模擬測驗。

一、先看這張：四大考試主題

① 人工智慧概念	② 資料處理與分析	③ 機器學習概念	④ 鑑別式 AI／生成式 AI
AI 定義、分類、治理、應用	資料來源、清洗、轉換、統計、EDA/CDA	監督/非監督/強化、模型、訓練、評估	DAI/GAI 原理、GAN/VAE/Diffusion、整合應用

二、一頁考前速記

考點	一句話記憶
AI	模擬人類智慧；不等於深度學習。
分析／預測／生成	分析＝找洞察；預測＝推估未來；生成＝產生文字、語音、圖像、影片等。
資料類型	結構化＝固定行列；半結構化＝XML/JSON/CSV；非結構化＝圖片、影像、音訊、文章等。
資料清洗	遺缺、重複、錯誤、不一致；離群值不一定是錯誤，須依業務需求判斷。
統計	Mean＝平均；Median＝中間；Mode＝最多；SD＝分散；IQR＝Q3－Q1。
偏態	平均數＞中位數 → 正偏態；平均數＜中位數 → 負偏態。
分析	EDA＝探索模式/關係/異常；CDA＝驗證先前假設。
監督式	有標記資料；主要做分類與迴歸。
非監督式	無標記資料；常見聚類、降維；K-Means、PCA。
強化學習	Agent 與 Environment 互動，依 Reward 試錯，追求最大累積獎勵。
模型	Linear＝連續數值；Logistic＝分類；KNN＝鄰近；Tree＝樹；Forest＝多樹；SVM＝分類邊界。
深度學習	CNN＝圖像；RNN＝序列；LSTM＝長期依賴；GAN＝Generator＋Discriminator。
訓練	Loss 衡量誤差；Gradient Descent/Optimizer 最小化 Loss；Cross Validation 評估穩健性。

過擬合	訓練很好、測試差；可用正則化、Early Stopping、Data Augmentation 等降低風險。
鑑別式 vs 生成式	鑑別式＝分類/預測；生成式＝產生新資料。

三、人工智慧概念

AI 定義

- 人工智慧（AI）旨在模擬人類智慧，使機器能執行學習、推理、解決問題、感知環境等任務。
- AI 是廣泛領域，包含機器學習、深度學習、專家系統等，不可直接把 AI 等同於深度學習。

AI 三類

- 分析型 AI：洞悉資料模式、分析大量資料並提供洞察。
- 預測型 AI：利用歷史資料預測未來趨勢與行為，如市場預測、風險評估。
- 生成型 AI：依 Prompt 產生文字、語音、圖像、影片等素材。

企業應用

- 醫療：疾病診斷、藥物研發、個人化醫療。
- 金融：風險評估、欺詐檢測、自動交易。
- 製造：自動化生產、品質控制、預測性維護。
- 交通：自動駕駛、交通流量預測。
- 娛樂：遊戲、VR、個人化內容推薦。

AI 治理

- 核心概念：透明、公平、隱私、安全、責任與倫理。
- 透明性：可透過報告、技術文件或網站揭露 AI 系統相關資訊。
- 國際合作：統一發展標準、降低濫用風險、促進技術轉移與交流。

四、資料處理與分析

資料類型

- 結構化：固定結構、行列形式，如 MySQL、PostgreSQL、規範化電子表格。
- 半結構化：有標籤、格式較彈性，如 XML、JSON、CSV。
- 非結構化：無固定結構，如圖片、影像、音訊、電子郵件、文章。

資料蒐集

- 常見來源：問卷調查、自有產品/設備、外部公開資料 API、網路爬蟲、外部付費資料。
- 資料是 AI/ML 模型的重要基礎，品質、數量與多樣性會影響模型效果。

資料清洗

- 遺缺值：可用平均值、中位數、眾數、插補或預測模型處理；也可能刪除，但要評估樣本代表性。
- 重複值：檢查主鍵/唯一識別碼等，保留正確記錄並刪除重複項。
- 錯誤值：如年齡為 -5；應檢測並修正。

- 離群值：可能是真實異常，也可能是雜訊/錯誤；不能看到離群值就一律刪除。

資料轉換

- 格式轉換：CSV → JSON。
- 資料型態轉換：字串 → 數值。
- 正規化/標準化：縮放數值，使不同特徵具有可比性。
- 離散化：連續數值 → 區間/類別，如年齡分成年輕、中年、老年。
- 資料縮減：特徵選擇、特徵提取、降維；PCA 是重要方法。

統計重點

- Mean：平均；容易受極端值影響。
- Median：排序後中間值；對極端值較不敏感。
- Mode：出現頻率最高的值；可有多個，也可能不存在。
- Standard Deviation：衡量分散程度；越大代表資料越分散。
- IQR：Q3－Q1，代表中間 50% 資料範圍。
- 正偏態：平均數通常大於中位數；負偏態：平均數通常小於中位數。

EDA / CDA / 診斷 / 預測

- EDA：不預設假設，探索模式、關聯與異常，常用視覺化。
- CDA：驗證先前提出的假設。
- 診斷性分析：追查『為什麼發生』，如 Drill-down、關聯分析、因果分析。
- 預測性分析：利用歷史資料預測未來，如迴歸、分類、時間序列。

假設檢定

- H0：虛無假設，通常表示沒有顯著效果或差異。
- Ha/H1：對立假設，表示存在顯著效果或差異。
- 顯著水準 α 常見 0.1、0.05、0.01。
- 若 p 值 < α，依原文件的考試脈絡可拒絕 H0。
- Type I Error：H0 為真卻拒絕 H0；Type II Error：Ha 為真卻未拒絕 H0。

五、機器學習概念

基本流程

- ① 資料處理 → ② 模型訓練 → ③ 測試/評估 → ④ 調校與優化。
- 資料處理包含清洗、標準化、特徵選擇與降維。

三種學習方式

- 監督式：有標記資料；主要任務是分類與迴歸。
- 非監督式：沒有標記；找資料內在結構，常見聚類與降維。
- 強化學習：Agent 與 Environment 互動，依 Reward 試錯，最大化累積獎勵。

常見模型

- Linear Regression：連續數值預測，如房價、銷售額。
- Logistic Regression：分類模型，以 Sigmoid 將輸出轉為 0～1 機率。
- KNN：依 K 個最近鄰資料進行分類或迴歸。
- SVM：尋找最大化類別間隔的分類邊界；可用 Kernel 處理非線性。
- Decision Tree：樹狀規則，優點是可解釋性佳。
- Random Forest：多棵決策樹集成，以投票或平均產生結果。

訓練與評估

- Loss Function：衡量預測值與實際值的差異；Loss 越高代表誤差越大。
- MSE：常用於迴歸；Cross-Entropy：常用於分類。
- Gradient Descent：調整模型參數以最小化 Loss。
- 過擬合：訓練表現佳、測試表現差；可用 Regularization、Early Stopping、Data Augmentation。
- Accuracy：分類正確比例；F1：綜合 Precision 與 Recall，原文件指出適合不平衡資料。
- K-Fold Cross Validation：分成 K 折，輪流用一折測試，其餘訓練，最後平均結果。

深度學習

- 深度學習是機器學習子領域，透過多層神經網路自動學習特徵。
- CNN：擅長圖像；卷積層提取局部特徵，池化層可降維。
- RNN：適合序列資料，能利用前一時刻狀態。
- LSTM：RNN 改良架構，透過門機制改善長期依賴與梯度消失問題。
- GAN：Generator 生成資料、Discriminator 判斷真假；可能出現 Mode Collapse。

強化學習

- 核心：Agent、Environment、Reward。
- 流程：State → Action → Environment 回饋 Next State + Reward → 更新 Policy。
- Q-Learning 學習 Q 值；Deep Q-Learning 用深度神經網路近似 Q 值。

六、鑑別式 AI 與生成式 AI

基本差異

- 鑑別式 AI：學習 P(y|x)，主要做分類、迴歸、辨識與預測。
- 生成式 AI：學習 P(x,y) 或 P(x)，模擬資料分佈並產生新資料。
- 最簡記憶：鑑別式＝判斷；生成式＝創造。

生成式模型

- GAN：生成器＋鑑別器，透過對抗學習生成資料。
- VAE：Encoder 將資料映射至 Latent Space，Decoder 再重建/生成。
- Diffusion：逐步加雜訊，再透過反向過程逐步去除雜訊生成資料。

生成式 AI 訓練

- 資料準備 → 模型選擇與搭建 → 模型訓練 → 微調。
- 訓練階段學習大量資料的模式與特徵；微調階段用特定領域/任務資料提升特定表現。
- 效果受到資料品質、模型結構、超參數與計算資源影響。

整合應用

- 生成式 AI 可生成多樣化資料；鑑別式 AI 負責分類、預測與判斷。
- 醫療：GAN 生成病理影像 → CNN 進行分類。
- 自動駕駛：生成式 AI 模擬複雜環境 → 鑑別式 AI 學習辨識與決策。
- 網路安全：生成式 AI 模擬攻擊模式 → 鑑別式 AI 分析流量並識別異常。
- 主要挑戰：訓練穩定性、資料偏差與公平性、系統整合架構。

七、考前必背：模型與用途對照表

模型/技術	主要用途	典型例子
Linear Regression	連續數值預測	房價、銷售額
Logistic Regression	分類／機率	疾病風險、違約、購買轉換
KNN	最近鄰分類／迴歸	依相似樣本判斷
SVM	分類邊界／超平面	文本、影像分類
Decision Tree	樹狀分類/預測	醫療診斷、風險評估
Random Forest	多棵決策樹集成	分類、預測、風險
K-Means	聚類	客戶分群
PCA	降維	高維資料壓縮/視覺化
CNN	圖像特徵	影像辨識、物件偵測
RNN	序列	文字、語音、時間序列
LSTM	長期序列依賴	時間序列、文字
GAN	生成資料	影像生成、資料增強
VAE	生成/重建/異常檢測	資料修復、合成資料
Diffusion	逐步去雜訊生成	高品質圖像、文字到圖像

八、100 題考前模擬測驗

作答建議：先完成全部題目，再對答案。此版本刻意不附答案，方便自行測驗。

1. 關於人工智慧（AI）的描述，下列何者最正確？

A. AI 僅能處理結構化資料

B. AI 涵蓋多種技術與應用領域

C. AI 等同於深度學習

D. AI 只能應用於學術研究

2. 下列何者最符合人工智慧的基本目標？

A. 讓電腦取代所有人類工作

B. 讓機器模擬人類智慧，完成學習、推理與問題解決等任務

C. 讓所有資料轉換成數字

D. 讓電腦完全不需要人類控制

3. 下列何者不是人工智慧常見的應用領域？

A. 自然語言處理

B. 電腦視覺

C. 機器學習

D. 純粹的電源供應管理

4. 企業利用 AI 分析客戶歷史購買資料，預測客戶下個月可能購買的產品，最接近：

A. 預測型 AI

B. 生成型 AI

C. 資料庫管理

D. 網路安全

5. 利用 AI 自動產生文章、圖片或音樂，最符合：

A. 鑑別式 AI

B. 生成式 AI

C. 描述性分析

D. 聚類分析

6. 下列何者最能描述「鑑別式 AI」？

A. 產生全新的圖片

B. 產生新的文章

C. 對資料進行分類或判斷

D. 自動產生音樂

7. 下列何者最能描述「生成式 AI」？

A. 只負責分類

B. 只負責資料排序

C. 學習資料模式並產生新的內容

D. 只負責資料庫查詢

8. 下列何者正確描述 AI、機器學習與深度學習的關係？

A. AI 是機器學習的一部分

B. 深度學習包含整個 AI

C. 機器學習與深度學習都是 AI 領域中的重要技術

D. 三者完全沒有關係

9. 銀行使用 AI 建立聊天機器人，最主要會涉及：

A. 自然語言處理與機器學習

B. 電源管理

C. 硬碟分割

D. 網路佈線

10. 智慧製造中利用 AI 預測機台可能發生故障，主要屬於：

A. 預測性維護

B. 資料備份

C. 網路交換

D. 文字生成

11. AI 治理最主要關注下列哪一組議題？

A. 公平、透明、隱私、安全與責任

B. CPU、RAM、硬碟容量

C. 螢幕尺寸與解析度

D. 網路速度與儲存容量

12. 下列何者最符合 AI 治理的精神？

A. 只要 AI 準確就不必考慮公平性

B. AI 風險應在設計、部署及運作過程中持續管理

C. AI 系統完全不需要人類監督

D. AI 發生問題後再處理即可

13. AI 系統的「透明性」主要是希望：

A. 增加 CPU 效能

B. 讓相關人員了解系統的運作、限制與決策相關資訊

C. 增加模型參數數量

D. 讓模型變得更複雜

14. AI 系統若因訓練資料偏差而對不同族群產生不公平結果，主要涉及：

A. 公平性

B. 網路頻寬

C. 資料壓縮

D. 儲存效能

15. AI 使用大量個人資料進行模型訓練時，最需要注意：

A. 隱私與資料保護

B. 顯示器更新率

C. CPU 時脈

D. 印表機速度

16. AI 治理中的國際合作，其重要性包括：

A. 促進共同標準

B. 降低 AI 技術遭濫用的風險

C. 促進技術交流與合作

D. 以上皆是

17. 下列何者最符合「人在迴圈內（Human-in-the-loop）」？

A. AI 完全自主決策，人類不參與

B. 人類在重要決策或行動前進行審核

C. AI 執行後人類完全不管

D. 人類只負責維護硬體

18. 如果 AI 系統可以自主運作，但人類持續監督並在必要時介入，較接近：

A. Human-in-the-loop

B. Human-over-the-loop

C. Human-out-of-the-loop

D. Machine-only-loop

19. 下列何者最可能造成 AI 模型產生偏見？

A. 訓練資料本身存在偏差

B. 顯示器太大

C. 網路線太長

D. 鍵盤種類不同

20. AI 的「可解釋性」主要希望：

A. 了解模型為何產生某項結果

B. 增加硬體容量

C. 減少資料筆數

D. 增加網路速度

21. 下列何者最符合 AI 安全性的概念？

A. 防止 AI 系統遭到惡意利用並降低錯誤造成的風險

B. 讓 AI 永遠產生相同答案

C. 增加螢幕解析度

D. 讓所有資料公開

22. 企業導入 AI 時，下列何者較合理？

A. 只考慮模型準確率

B. 同時考慮效益、成本、風險、隱私與安全

C. 完全不需要測試

D. 只要模型越大越好

23. 下列何者最適合描述「弱人工智慧（Narrow AI）」？

A. 具有完整人類智慧

B. 能處理所有人類工作

C. 專注於特定任務或領域

D. 完全不需要訓練資料

24. AI 在醫療影像中辨識腫瘤，主要屬於：

A. 電腦視覺／影像分析

B. 網路路由

C. 資料備份

D. 作業系統管理

25. 下列何者是 AI 導入企業時較合理的流程？

A. 直接購買最大模型

B. 確認需求 → 準備資料 → 選擇模型 → 測試 → 部署與監控

C. 先部署再確認需求

D. 完全不需要評估資料品質

26. 下列何者屬於結構化資料？

A. MySQL 資料表

B. MP4 影片

C. JPEG 圖片

D. MP3 音樂

27. 下列何者通常屬於半結構化資料？

A. JSON

B. JPEG

C. MP4

D. 純手寫文件

28. 下列何者最典型屬於非結構化資料？

A. Excel 表格

B. 關聯式資料庫

C. 圖片

D. SQL Table

29. 大數據常用的 3V 包含：

A. Volume、Velocity、Variety

B. Value、Video、Virtual

C. Version、Visual、Voice

D. Volume、Virtual、Video

30. Big Data 的 Volume 主要代表：

A. 資料量

B. 資料速度

C. 資料多樣性

D. 資料價值

31. Big Data 的 Velocity 主要代表：

A. 資料量

B. 資料產生與處理速度

C. 資料種類

D. 資料準確度

32. Big Data 的 Variety 主要代表：

A. 資料量

B. 資料產生速度

C. 資料類型與格式多樣性

D. 資料安全性

33. 資料清洗最主要的目的為：

A. 增加資料錯誤

B. 提升資料品質

C. 增加資料重複

D. 刪除所有資料

34. 下列何者屬於資料清洗常見問題？

A. 遺缺值

B. 重複資料

C. 錯誤資料

D. 以上皆是

35. 資料中出現「年齡 = -20」最可能屬於：

A. 合理資料

B. 異常或錯誤資料

C. 結構化資料

D. 生成式資料

36. 資料出現缺失值時，下列何者可能是處理方法？

A. 使用平均數或中位數填補

B. 必須全部保留空白

C. 一定全部刪除

D. 將所有資料轉成圖片

37. 下列關於離群值的敘述何者正確？

A. 離群值一定是錯誤

B. 離群值一定要刪除

C. 離群值可能是真實且有意義的資料

D. 離群值與資料分析無關

38. 將 CSV 資料轉成 JSON，屬於：

A. 資料格式轉換

B. 資料分類

C. 強化學習

D. 模型訓練

39. 將「男、女」轉換成 0、1，較接近：

A. 資料編碼／轉換

B. 資料備份

C. 模型部署

D. 聚類

40. Normalization 的主要目的之一是：

A. 調整資料尺度

B. 增加資料錯誤

C. 刪除所有資料

D. 產生圖片

41. 平均數的主要缺點之一是：

A. 無法計算

B. 容易受到極端值影響

C. 永遠等於中位數

D. 只能處理文字

42. 中位數的主要特性是：

A. 對極端值較不敏感

B. 一定是最大值

C. 一定是最小值

D. 一定等於平均數

43. 眾數（Mode）是：

A. 最大值

B. 最小值

C. 出現頻率最高的值

D. 所有數值的平均

44. 標準差主要用來描述：

A. 資料分散程度

B. 資料筆數

C. 資料格式

D. 資料名稱

45. 若某生產線產品尺寸的標準差大幅增加，通常代表：

A. 生產波動變大

B. 所有產品完全相同

C. 資料一定沒有問題

D. 平均數一定增加

46. IQR 的計算方式為：

A. Q1 + Q3

B. Q3 − Q1

C. Q3 ÷ Q1

D. Q1 × Q3

47. IQR 主要代表：

A. 最大值與最小值差距

B. 中間 50% 資料的範圍

C. 所有資料平均值

D. 最大值

48. 若資料呈現正偏態，通常：

A. 平均數 > 中位數

B. 平均數 < 中位數

C. 平均數 = 0

D. 中位數一定為最大值

49. 探索性資料分析（EDA）的主要目的為：

A. 探索資料中的模式、關係與異常

B. 只驗證既有假設

C. 直接建立硬體

D. 只進行資料備份

50. 驗證性資料分析（CDA）主要著重於：

A. 探索未知資料

B. 驗證既有假設

C. 刪除所有異常值

D. 資料格式轉換

51. 機器學習的基本概念是：

A. 讓機器從資料中學習規律

B. 讓機器不需要資料

C. 讓機器只能執行固定指令

D. 讓所有資料轉成圖片

52. 監督式學習最大的特徵是：

A. 使用有標記資料

B. 完全沒有資料

C. 只有獎勵

D. 只能進行分群

53. 非監督式學習的主要特徵是：

A. 有明確標籤

B. 沒有標籤，從資料中尋找結構或模式

C. 必須有人逐筆給答案

D. 只能處理圖片

54. 強化學習的核心概念是：

A. Label

B. Reward

C. Table

D. SQL

55. 利用大量已標記的醫療影像訓練模型判斷良性或惡性，屬於：

A. 監督式學習

B. 非監督式學習

C. 強化學習

D. 資料清洗

56. 利用客戶購買行為，自動將客戶分成不同群組，屬於：

A. 監督式學習

B. 非監督式學習

C. 強化學習

D. 文字生成

57. 訓練 AI 玩遊戲，依照得分高低調整策略，最適合：

A. 監督式學習

B. 非監督式學習

C. 強化學習

D. K-Means

58. Classification 的主要目的是：

A. 預測類別

B. 預測連續數值

C. 產生圖片

D. 資料格式轉換

59. Regression 通常用於：

A. 預測連續數值

B. 將資料分成群組

C. 產生文章

D. 判斷圖片格式

60. 預測房屋價格最適合使用：

A. Linear Regression

B. K-Means

C. GAN

D. CNN

61. 判斷 Email 是否為垃圾郵件，最典型屬於：

A. 分類問題

B. 分群問題

C. 生成問題

D. 降維問題

62. Logistic Regression 主要用於：

A. 分類

B. 圖片生成

C. 資料分群

D. 影像壓縮

63. KNN 的基本概念是：

A. 根據鄰近資料進行判斷

B. 產生新的資料

C. 建立神經網路

D. 使用生成器與鑑別器

64. Decision Tree 的主要特徵：

A. 樹狀決策結構

B. 只能處理圖片

C. 只能進行生成

D. 必須使用 GAN

65. Random Forest 的核心概念：

A. 多棵決策樹的集成

B. 單一神經元

C. 單一迴歸線

D. 一個生成器

66. Random Forest 相較於單一 Decision Tree 的主要優勢之一：

A. 可以提高穩定性並降低過擬合風險

B. 一定不需要資料

C. 一定比所有模型準確

D. 只能處理文字

67. K-Means 中的 K 通常代表：

A. 特徵數

B. 群組數

C. 資料筆數

D. 迭代時間

68. K-Means 最適合：

A. 將資料分成 K 個群組

B. 產生文章

C. 判斷垃圾郵件

D. 預測房價

69. K-Means 的一項限制是：

A. 容易受到離群值影響

B. 完全不能計算距離

C. 一定需要人工標記

D. 只能產生圖片

70. K-Means 通常較不適合：

A. 具有明顯群集的數值資料

B. 非球形、密度差異很大且含大量離群值的複雜資料

C. 客戶分群

D. 簡單群集分析

71. 模型在訓練資料上表現非常好，但在新資料上表現很差，稱為：

A. Underfitting

B. Overfitting

C. Clustering

D. Normalization

72. 下列何者可以降低過擬合？

A. 增加正則化

B. 無限制增加模型複雜度

C. 移除所有驗證資料

D. 只使用訓練資料

73. Cross Validation 的主要用途：

A. 評估模型穩健性

B. 增加圖片解析度

C. 將 JSON 轉成 CSV

D. 產生新的資料庫

74. 梯度下降（Gradient Descent）主要用於：

A. 模型參數最佳化

B. 資料庫備份

C. 圖片壓縮

D. 資料格式轉換

75. 模型中的 Loss Function 主要用途：

A. 衡量模型預測結果的誤差

B. 增加硬碟容量

C. 建立網路連線

D. 儲存圖片

76. 神經網路通常包含：

A. 輸入層、隱藏層、輸出層

B. CPU、RAM、硬碟

C. Generator、Router、Switch

D. SQL、HTML、CSS

77. 神經網路透過反向傳播（Backpropagation）主要進行：

A. 調整模型權重

B. 刪除所有資料

C. 產生網路封包

D. 格式化硬碟

78. CNN 最適合處理：

A. 圖像資料

B. 關聯式資料庫

C. 網路設定檔

D. 純文字檔案名稱

79. CNN 的卷積層主要功能之一：

A. 擷取資料中的局部特徵

B. 儲存資料庫

C. 進行網路路由

D. 產生 SQL

80. RNN 特別適合處理：

A. 序列資料

B. 硬碟分割

C. 網路交換

D. 靜態資料庫結構

81. LSTM 是：

A. RNN 的一種改良架構

B. 一種資料庫

C. 一種網路協定

D. 一種資料格式

82. LSTM 主要改善 RNN 的哪項問題？

A. 長期依賴與梯度消失等問題

B. 硬碟容量不足

C. 網路封包遺失

D. 資料格式錯誤

83. GAN 是由哪兩個主要元件構成？

A. Generator 與 Discriminator

B. Encoder 與 Router

C. CNN 與 KNN

D. CPU 與 GPU

84. GAN 中 Generator 的主要任務：

A. 生成資料

B. 判斷真假

C. 執行資料庫查詢

D. 進行分群

85. GAN 中 Discriminator 的主要任務：

A. 產生圖片

B. 判斷資料真偽

C. 預測房價

D. 將資料分群

86. GAN 訓練過程常見的問題之一：

A. Mode Collapse

B. SQL Injection

C. IP 位址不足

D. 硬碟碎片

87. 生成式 AI 的核心特色是：

A. 生成新的內容

B. 只做分類

C. 只做排序

D. 只做資料清洗

88. 下列何者最可能屬於生成式 AI 的應用？

A. 生成產品文案

B. 判斷垃圾郵件

C. 客戶分群

D. 預測房價

89. 下列何者較典型屬於鑑別式 AI？

A. 分類醫療影像

B. 生成病理影像

C. 產生文章

D. 產生音樂

90. 鑑別式 AI 通常著重於：

A. 學習資料之間的判別關係

B. 產生全新資料

C. 資料備份

D. 資料格式轉換

91. 生成式 AI 與鑑別式 AI 整合時，下列何者最合理？

A. 生成式 AI 產生資料，鑑別式 AI 進行判斷

B. 兩者都只能做分類

C. 兩者都不能處理資料

D. 生成式 AI 只能做資料清洗

92. 在醫療影像中，GAN 產生模擬病理影像，再使用 CNN 判斷疾病，屬於：

A. 生成式與鑑別式 AI 的整合

B. 純資料庫應用

C. 純強化學習

D. 純資料清洗

93. 在網路安全領域，生成式 AI 可以：

A. 模擬潛在攻擊情境

B. 只負責更換 IP

C. 只負責建立帳號

D. 只負責備份硬碟

94. 在網路安全中，鑑別式 AI 可以：

A. 分析網路流量並識別異常

B. 只生成圖片

C. 只產生文章

D. 只進行資料壓縮

95. 生成式 AI 提供資料增強（Data Augmentation）的一個重要用途是：

A. 改善資料稀缺或不平衡問題

B. 增加資料錯誤

C. 完全取消模型訓練

D. 讓模型不需要測試

96. 生成式 AI 與鑑別式 AI 整合時，可能面臨：

A. 資料偏差與公平性問題

B. 模型訓練穩定性問題

C. 生成資料品質問題

D. 以上皆是

97. 如果生成式 AI 所產生的資料放大原始訓練資料的偏差，可能導致：

A. 鑑別模型產生不公平結果

B. CPU 速度增加

C. 網路頻寬增加

D. 儲存空間自動增加

98. 大型生成式 AI 模型通常需要大量資料訓練，其主要目的之一：

A. 學習資料中的模式與特徵

B. 讓電腦不用作業系統

C. 取代資料庫

D. 取消所有人工監督

99. 下列哪一個配對最正確？

A. CNN－圖像、RNN－序列、GAN－生成

B. CNN－資料庫、RNN－防火牆、GAN－SQL

C. K-Means－文字生成、GAN－資料清洗

D. Logistic Regression－圖片生成

100. 下列哪一項最完整描述本課程人工智慧基礎概論的核心？

A. 只要會使用 ChatGPT 就代表了解 AI

B. 了解 AI 基本概念、資料處理、機器學習，以及鑑別式與生成式 AI

C. 只需要學習 Python 程式設計

D. 只需要熟悉 GPU 硬體

九、100 題答案作答卡

請先自行作答；每 10 題為一組，方便檢查。

1. ＿＿	2. ＿＿	3. ＿＿	4. ＿＿	5. ＿＿	6. ＿＿	7. ＿＿	8. ＿＿	9. ＿＿	10. ＿＿
11. ＿＿	12. ＿＿	13. ＿＿	14. ＿＿	15. ＿＿	16. ＿＿	17. ＿＿	18. ＿＿	19. ＿＿	20. ＿＿
21. ＿＿	22. ＿＿	23. ＿＿	24. ＿＿	25. ＿＿	26. ＿＿	27. ＿＿	28. ＿＿	29. ＿＿	30. ＿＿
31. ＿＿	32. ＿＿	33. ＿＿	34. ＿＿	35. ＿＿	36. ＿＿	37. ＿＿	38. ＿＿	39. ＿＿	40. ＿＿
41. ＿＿	42. ＿＿	43. ＿＿	44. ＿＿	45. ＿＿	46. ＿＿	47. ＿＿	48. ＿＿	49. ＿＿	50. ＿＿
51. ＿＿	52. ＿＿	53. ＿＿	54. ＿＿	55. ＿＿	56. ＿＿	57. ＿＿	58. ＿＿	59. ＿＿	60. ＿＿
61. ＿＿	62. ＿＿	63. ＿＿	64. ＿＿	65. ＿＿	66. ＿＿	67. ＿＿	68. ＿＿	69. ＿＿	70. ＿＿
71. ＿＿	72. ＿＿	73. ＿＿	74. ＿＿	75. ＿＿	76. ＿＿	77. ＿＿	78. ＿＿	79. ＿＿	80. ＿＿
81. ＿＿	82. ＿＿	83. ＿＿	84. ＿＿	85. ＿＿	86. ＿＿	87. ＿＿	88. ＿＿	89. ＿＿	90. ＿＿
91. ＿＿	92. ＿＿	93. ＿＿	94. ＿＿	95. ＿＿	96. ＿＿	97. ＿＿	98. ＿＿	99. ＿＿	100. ＿＿

十、最後 5 分鐘速查

- AI ≠ 深度學習；AI 是較大的領域。
- 分析型＝洞察資料；預測型＝預測未來；生成型＝創造內容。
- 結構化＝固定行列；半結構化＝JSON/XML/CSV；非結構化＝圖片/音訊/文章。
- 資料清洗：遺缺、重複、錯誤、不一致；離群值不一定刪除。
- Mean 易受極端值影響；Median 對極端值較不敏感；Mode＝最多。
- SD 越大 → 分散越大；IQR＝Q3－Q1；正偏態 Mean＞Median。
- EDA＝探索；CDA＝驗證；Diagnostic＝為什麼；Predictive＝未來。
- 監督式＝有 Label；非監督式＝無 Label；強化式＝Reward。
- Linear＝連續數值；Logistic＝分類；K-Means＝分群；PCA＝降維。
- CNN＝圖像；RNN＝序列；LSTM＝長期依賴；GAN＝生成器＋鑑別器。
- Overfitting＝訓練好、測試差；Regularization / Early Stopping / Data Augmentation 可降低風險。
- Loss 衡量誤差；Gradient Descent 最小化 Loss；Cross Validation 評估穩健性。
- 鑑別式 AI＝分類/迴歸/判斷；生成式 AI＝生成新資料。
- GAN：Generator 產生、Discriminator 判斷；VAE：Encoder＋Decoder；Diffusion：加雜訊→反向去雜訊。
- 整合應用：生成式 AI 產生資料/情境，鑑別式 AI 分類/判斷；注意偏差、公平、穩定性與資料品質。

祝你 iPAS 考試順利！