

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 1. 商管人最需要培養的資料探勘技能是
- (A) 了解演算法
 - (B) 理解商業上的管理意涵
 - (C) 懂得程式設計
 - (D) 學會系統分析。
- (D) 2. 商管人應該要至少能以下列角度去看資料探勘，那就是從資料輸入、資料分析與資訊輸出等三個單純的動作去展開，資料輸入不包含哪個程序
- (A) 資料欄位選擇
 - (B) 前置作業
 - (C) 轉換資料格式
 - (D) 操作軟體工具。
- (A) 3. 資料探勘的資料輸入過程中依照需要轉換資料格式有幾種作法，在前置作業完成後要做資料簡化與轉換工作。因為資料量過大且有很多用不到的資料，將不必要的欄位刪除或不納入分析，這是程序上哪一個環節
- (A) 精簡維度
 - (B) 調整編碼方式
 - (C) 調整顆粒度
 - (D) 資料型態轉換。
- (C) 4. 資料探勘的資料輸入過程中依照需要轉換資料格式有幾種作法，在前置作業完成後要做資料簡化與轉換工作。因為資料量過大且有很多用不到的資料，將原始資料改變為較能理解的格式，這是程序上哪一個環節
- (A) 精簡維度
 - (B) 資料型態轉換
 - (C) 調整編碼方式
 - (D) 調整顆粒度。
- (B) 5. 資料探勘的資料輸入過程中依照需要轉換資料格式有幾種作法，在前置作業完成後要做資料簡化與轉換工作。因為資料量過大且有很多用不到的資料，「決定資料是要分析到每日還是每月？」，這是程序上哪一個環節
- (A) 調整編碼方式
 - (B) 調整顆粒度
 - (C) 精簡維度
 - (D) 資料型態轉換。
- (D) 6. 資料探勘的資料輸入過程中依照需要轉換資料格式有幾種作法，在前置作業完成後要做資料簡化與轉換工作。因為資料量過大且有很多用不到的資料，連續值與離散值轉換，這是程序上哪一個環節
- (A) 調整顆粒度
 - (B) 精簡維度
 - (C) 調整編碼方式
 - (D) 資料型態轉換。
- (D) 7. 資料探勘的資訊輸出過程中，解釋探勘後的資訊成為知識無法透過下列哪一種方式完成
- (A) 訪談
 - (B) 次級資料分析
 - (C) 觀察
 - (D) 解譯程式碼。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 8. 關於資料探勘描述，以下何者為非
- (A) 挖掘出有用的東西
 - (B) 挖掘出奇點
 - (C) 具備意涵
 - (D) 對應商管目的。
- (C) 9. 資料探勘上，所謂的「挖哪邊」的觀念，以下何者為非
- (A) 就是痛點
 - (B) 痛點就是對商管而言有意義的點
 - (C) 痛點可以視為替代品
 - (D) 痛點可以視為必需品。
- (C) 10. 進行資料探勘的起手式描述中，所謂可能有商機的痛點是指
- (A) 不值得我們關注的地方
 - (B) 與過往經驗相同之處
 - (C) 面對可能有商機之處應該要繼續透過質化與量化的手法繼續深入了解內在意義
 - (D) 有結果就是有商機。
- (A) 11. 在資料探勘過程中，如果遇到資料找到不多，通常可以從下列哪一點去著手改進
- (A) 我們資料找的還不夠仔細，我們去找找是否可以找到顆粒度較細的資料
 - (B) 請該單位準備好我們所需要的資料給我們
 - (C) 就用少量數據去分析
 - (D) 直接跑演算法。
- (A) 12. 關於資料探勘的描述何者正確
- (A) 是一種事後分析，需要有資料才能進行
 - (B) 不需要被驗證
 - (C) 只有社會科學可以但自然科學不可以使用
 - (D) 是一種事前分析，可以透過推衍得知結果。
- (A) 13. 「如何確認資料可用」的描述中，包含了下列哪一個選項
- (A) 確認是使用什麼演算法
 - (B) 精準定義結論
 - (C) 應用管理工具
 - (D) 確認分析軟體。
- (B) 14. 何謂有品質的資料，下列敘述者正確
- (A) 漂亮
 - (B) 可用
 - (C) 大量
 - (D) 少量。
- (C) 15. 在資料完整的觀念中，資料有效性是指
- (A) 時間正確
 - (B) 選對列
 - (C) 沒有缺漏或空值
 - (D) 選對欄位。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (A) 16. 關於探勘可用的資料表的觀念，下列者正確
- (A) 多以二維表示
 - (B) 超過二維以上不要刪除
 - (C) 超過二維以上可以透過直接刪除來降維
 - (D) 直接跑樞紐分析結果表就好。
- (C) 17. 所謂的集群就是要從過往的資料中
- (A) 取垂直面整合
 - (B) 取個體之間的最大差異點來整合
 - (C) 取共通性來整合
 - (D) 取最大化差異。
- (B) 18. 關於集群的作法的敘述不包含
- (A) 看個別欄位的內容是否接近
 - (B) 喜不喜歡這個物件
 - (C) 選擇適當的屬性（欄位）並比較相似處
 - (D) 肘方法來判別集群數。
- (D) 19. 電腦要判別人長相是否接近不會透過下列哪一個選項
- (A) 眉毛圖像的形狀的資料庫
 - (B) 嘴唇圖像的資料庫
 - (C) 眼睛圖像的資料庫
 - (D) 聲音的資料庫。
- (A) 20. 關於歐幾里德距離計算的描述，下列何者正確
- (A) 丟入隨機參數，以選定的幾個（比如：三個）單筆資料為主，看其他資料與該種子資料的距離進行運算
 - (B) 固定選取三個單筆資料為主，看其他資料與該種子資料的距離進行運算
 - (C) 餵入判別的欄位資料是依序由小到大選取的
 - (D) 餵入判別的欄位資料是依序由大到小選取的。
- (D) 21. 透過人腦要判別一個人的長相是否是美女/帥哥的方式不包含下列哪一選項
- (A) 觀察眼睛
 - (B) 觀察頭髮
 - (C) 觀察耳朵
 - (D) 觀察書寫文字。
- (A) 22. 關於集群分析描述，下列何者為非
- (A) 所有資料都要納入
 - (B) 不合規範的要拿掉
 - (C) 不能觀察的要拿掉
 - (D) 因顆粒度無法納入的可視狀況調整後納入分析。
- (B) 23. 下列選項中何者不為集群分析的觀念
- (A) 同羽毛的鳥就是會聚一起
 - (B) 整理過往資料建立模型判別進入門內的是男生或女生
 - (C) 非監督式探勘
 - (D) 組內差異最小組間差異最大。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (A) 24. 在資料探勘的分析方法中，集群分析其英文名稱為下列哪一個選項
- (A) Clustering
 - (B) Classification
 - (C) Association Rule
 - (D) Classifier。
- (C) 25. 電子商務興起與關聯規則的描述，下列何者正確
- (A) 關聯規則分析無用武之地
 - (B) 電子商務與實體商務可以用一樣的關聯規則分析結果去操作
 - (C) 透過電子商務要設法透過關聯規則分析了解商品的關聯，適當透過線上跑馬燈來激起顧客的購買欲望
 - (D) 最重要的指標是判別度。
- (D) 26. 實體商店與電子商務之差異的描述，以下何者為非
- (A) 實體商店可以觀察貨架
 - (B) 電子商務下消費者不容易一直逛貨架
 - (C) 電子商務的彈出式廣告效果有限
 - (D) 實體商店可以賣出更多商品。
- (C) 27. 何謂關聯規則分析，以下的描述何者為非
- (A) 關聯就是關係
 - (B) 獨立存在的物件之間產生關係
 - (C) 關聯必須是「連續」而非「離散」的資料
 - (D) 信心度是用來產生關聯規則的依據。
- (D) 28. 在發現關聯規則之後，以下敘述何者為非
- (A) 搭配質化分析結果規劃行銷企劃
 - (B) 不要貿然直接看數據說故事
 - (C) 量化分析與質化分析需要一起看
 - (D) 可以直接應用。
- (C) 29. 關於 Apriori 演算法的描述，下列何者正確
- (A) 只需找出最頻繁項目集
 - (B) 找出好用的項目集
 - (C) 一定要給定支持度與信心度
 - (D) 可以快速且節省資源的生成規則。
- (B) 30. 關於關聯規則的支持度的描述，以下何者為是
- (A) 決策變數在聯集中所出現的比例
 - (B) 就是商品組合在納入模型計算的樣本中出現的總次數
 - (C) 可以看成統計裡面的條件機率
 - (D) 決策變數在交集中所出現的比例。
- (D) 31. 關於關聯規則的信心度的描述，以下何者為非
- (A) 信心就是主管希望知道在賣出 X 的時候，搭著賣 Y 的可能性
 - (B) 可以看成統計裡面的條件機率
 - (C) 看看能否有創造搭售的可能性
 - (D) 信心度一定要非常高才有意義。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 32. 關於決策樹的基本原理敘述，以下何者錯誤
- (A) 監督式學習
 - (B) 非監督式學習
 - (C) 有目標
 - (D) 用以分類。
- (B) 33. 關於決策樹的描述，以下何者不正確
- (A) 樹狀的結構
 - (B) 向上長的樹
 - (C) 有樹根
 - (D) 可能會過度擬合。
- (D) 34. 關於決策樹分枝的判別標準的中英文名稱，下列哪一個選項是正確的
- (A) 熵 (Gini Coefficient)
 - (B) 吉尼係數 (Entropy)
 - (C) 商 (Entropy)
 - (D) 吉尼係數 (Gini Coefficient)。
- (A) 35. 關於決策樹分枝的判別標準值描述中，下列何者正確
- (A) 熵越低越好
 - (B) 吉尼係數越高越好
 - (C) 熵越高越好
 - (D) 隨機性越高越好。
- (D) 36. 關於 CART 決策樹方法的特色描述，以下何者為非
- (A) 可以分類連續變數與類別變數
 - (B) 已用過的變數可以重複使用
 - (C) 透過計算吉尼係數來判別分類
 - (D) 不能分類連續變數。
- (A) 37. 關於使用決策樹建立分類模型的描述中，何謂過度擬合(overfitting)的概念
- (A) 切太細了
 - (B) 切太粗
 - (C) 分支越多越好
 - (D) 分支越少越好。
- (A) 38. 在決策樹建立模型的過程中，關於事後剪枝的描述，以下何者正確
- (A) 在樹已經生成之後，才對過度擬合的部分進行修剪
 - (B) 生成樹的時候就對樹的生長進行限制
 - (C) 預先給定標準
 - (D) 多為主觀剪枝。
- (A) 39. 關於決策樹的概念，以下選項何者不正確
- (A) 吉尼係數為正數表示分類已經一致
 - (B) 是一種貪婪(greedy)演算法
 - (C) 可能會過度擬合
 - (D) 成本複雜度剪枝法 (Cost-Complexity Pruning) 就是一種事後剪枝的手法。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (D) 40. 關於隨機森林的方法，下列選項何者正確
- (A) 是衍生於類神經網路
 - (B) 是衍生於深度學習
 - (C) 是衍生於關聯規則自動化方法
 - (D) 是衍生於決策樹。
- (B) 41. 關於隨機森林方法的可能用途
- (A) 判別顧客年紀
 - (B) 推算顧客重要性
 - (C) 確定員工薪資
 - (D) 找出關聯規則。
- (C) 42. 關於隨機森林演算法的觀念，下列何者正確
- (A) 採取 bootstrapping 法抽取後不放回
 - (B) 同一人不會被抽到多次
 - (C) 採取 bootstrapping 法抽取後還會放回
 - (D) 隨機森林的 bootstrapping 創造資料集的方法就是 begging。
- (A) 43. 關於隨機森林演算法的敘述，以下何者錯誤
- (A) out-of-bag 是指都沒有被列入考慮的變數（欄位）
 - (B) out-of-bag 是指都沒有被列入考慮的資料
 - (C) 可以用 out-of-bag 資料作為測試用資料來修正我們隨機選擇變數（欄位）的數量
 - (D) 選擇變數數量的依據是輸入 out-of-bag 的資料的錯誤率。
- (A) 44. 關於 kNN 演算法中 k 的條件之描述，以下何者錯誤
- (A) 用來集群
 - (B) k 是指最近鄰居
 - (C) 隨機找 k 個
 - (D) k 不可以是 1。
- (D) 45. 關於 kNN 演算法的觀念，以下何者錯誤
- (A) k 不可以太小
 - (B) k 需要是奇數
 - (C) 選擇的鄰居針對特定議題的意見居多數者勝出
 - (D) 與 k-Means 相同。
- (C) 46. 關於模型預測的驗證之描述，下列何者正確
- (A) k-fold 不適用於隨機森林
 - (B) k-fold 不可以用來驗證訓練出的決策樹的好壞
 - (C) 在隨機森林中沒被抽到的樣本數可以當作驗證資料用
 - (D) 在隨機森林中沒被抽到的樣本數不可以當作驗證資料用。
- (B) 47. 關於隨機森林方法的觀念，下列何者正確
- (A) 就是少數決
 - (B) 隨機森林的樣本抽樣設計是採取抽取之後還可以放回的方法
 - (C) 我的意見只會出現在一棵樹裡面
 - (D) 隨機森林的樣本抽樣設計是採取抽取之後不可以放回。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 48. 以下的哪一選項的描述何者不為預測
- (A) 了解下一個進來教室的人的性別
 - (B) 今天買了一個東西 A 下次會買 B
 - (C) 戰爭是否勝利
 - (D) 離職的可能性。
- (D) 49. 關於類神經網路的神經元運作，以下何者為非
- (A) 有範圍
 - (B) 答案通常就是二元值(True/False)
 - (C) 仿照大腦神經元運算原理去施作
 - (D) 可以天馬行空去思考。
- (C) 50. 過往在商管的統計上，要預測大多是透過迴歸分析，關於迴歸分析以下何者描述是正確的
- (A) 多個因變數
 - (B) 一個自變數
 - (C) 因變數就是預測的結果
 - (D) 自變數一定要是類別變數。
- (D) 51. 關於類神經網路運算能否合於預期的判斷方式之描述，下列何者錯誤
- (A) 要有訓練組
 - (B) 要有測試組
 - (C) 不需要生成函數
 - (D) 逐步修正錯誤值。
- (D) 52. 為何類神經演算法會很準的描述中，以下何者不正確
- (A) 因為會一直重複進行運作
 - (B) 重複將誤差值配上權重丟入隱藏層一直訓練，直到訓練成功為止
 - (C) 電腦速度跑這些重複的運算非常很快
 - (D) 在訓練數量極少下可以產生最佳結果。
- (C) 53. 關於倒傳遞神經網路方法的描述，以下何者為非
- (A) 要給定學習率
 - (B) 權重為一個大於-1 但小於 1 的百分比數值
 - (C) 會經過反函數判斷或調整其數據
 - (D) 將誤差回饋到上一層進行修訂。
- (D) 54. 在倒傳遞神經網路方法中，不會透過誤差回饋修訂下列哪一個選項
- (A) 閾值
 - (B) 權重
 - (C) 隱藏層權重
 - (D) 激勵函數。
- (B) 55. 關於類神經的特性之描述，以下何者為非
- (A) 類神經可以透過誤差值回饋修正參數以求達到最佳結果
 - (B) 類神經分析的結果一定有一個合理的解釋
 - (C) 若能應用大量資料去訓練類神經應能提升準確程度
 - (D) 激勵函數就是要去除線性結果的可能性。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (A) 56. 集群分析又可以稱為
- (A) 聚類分析
 - (B) 因素分析
 - (C) 因果分析
 - (D) 主成份分析。
- (B) 57. 關於 RFM 模型的英文的基本名稱中 R、F、M 的正確選項為
- (A) F 是 Fluently
 - (B) R 是 Recency
 - (C) M 是 Mental
 - (D) M 是 Many。
- (A) 58. 關於 K-Means 演算法的觀念，下列正確選項為
- (A) K 是指 K 個分群數量的意思
 - (B) 計算隨機種子與其他資料的因果關係
 - (C) K 值永遠為 5
 - (D) K 值可以設定為-2。
- (B) 59. 讓消費習慣有差異的客戶可以分別集結到不同的客戶組群中，同時也將有相類似消費習慣的客戶集結到相對應的客戶組群中，這就是
- (A) 獨熱編碼
 - (B) 集群分析
 - (C) 關聯規則
 - (D) 主題模型分析。
- (C) 60. 「公司的訂單收入中，有 80%的銷售額來自於 20%的客戶」這是下列哪一種分析
- (A) 決策樹
 - (B) 奇異值分解
 - (C) 柏拉圖法則(Pareto Principle)
 - (D) 分類法。
- (D) 61. 在 RFM 模式中說明客戶最近一次消費的時間點距離目前數據分析工作的時間點的時間間距 (Time interval)的是下列哪一個選項
- (A) Frequency
 - (B) Monetary
 - (C) Money
 - (D) Recency。
- (B) 62. 關於 RFM 模型中 F 的觀念是下列哪一個選項
- (A) 消費金額
 - (B) 消費頻率
 - (C) 最近一次消費
 - (D) 最久遠的一次消費。
- (D) 63. 關於「買啤酒(Beer)也會買尿布(Diaper)」的故事當時是發生在
- (A) Apple 公司
 - (B) 亞馬遜公司
 - (C) SAP 公司
 - (D) 沃爾瑪(Walmart)大型零售賣場中。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 64. 「啤酒(Beer)與尿布(Diaper)」的故事是採用下列哪一類演算法
- (A) 決策樹分類法
 - (B) 關聯規則
 - (C) 集群分析
 - (D) 貝氏分類法。
- (B) 65. 購物籃(Shopping basket)分析所採用方法是
- (A) 隨機森林
 - (B) 關聯規則
 - (C) 集群分析
 - (D) kNN 分類法。
- (A) 66. 分析購物籃的採購內容就等同是在分析客戶的
- (A) 消費行為(Consumer behavior)
 - (B) 倉庫管理行為
 - (C) 財務分析行為
 - (D) 預期心理行為。
- (B) 67. 關於 Market basket analysis 的中文意思為
- (A) 出貨分析
 - (B) 購物籃分析
 - (C) 投籃命中分析
 - (D) 進貨分析。
- (C) 68. 購物籃(車)中的內容可以提供許多珍貴的資訊，這對賣場中商品的下列哪一項活動非常重要
- (A) 出貨
 - (B) 會計
 - (C) 補貨(Replenishment)
 - (D) 財務。
- (D) 69. 關於 Apriori 方法的功能與描述，下列何者正確
- (A) 是一種決策樹分類法
 - (B) 是一種貝氏分類法
 - (C) 是一種集群分析方法
 - (D) 是一種關聯規則方法。
- (A) 70. 因為當交易資料量很大的時候常常會遇到計算上瓶頸，因此需要將原來的 list 資料型態轉為二維表格的資料框(DataFrame)，而這個資料框的內容填滿真(True)與假(False)的邏輯值，這些真、假值也就是什麼值
- (A) 布林值
 - (B) 焦點值
 - (C) 奇異值
 - (D) 閾值。
- (A) 71. 關於決策樹分類演算法的描述何者正確
- (A) 監督式方法
 - (B) 非監督式方法
 - (C) 兼容監督與非監督方法
 - (D) 加強非監督式方法。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (D) 72. 關於決策樹分類演算法的描述何者錯誤
- (A) 有目標變數
 - (B) 有自變數
 - (C) 運用自變數去預測目標變數
 - (D) 自變數就是要預測的未來資料。
- (D) 73. 關於決策樹的預測的驗證議題，下列何者錯誤
- (A) 使用訓練資料集來建立模型
 - (B) 將資料拆成訓練與測試組，前者訓練模型，後者用以驗證
 - (C) 測試組資料量不要多於訓練組資料量
 - (D) 測試組資料量一定要是訓練組資料量的九倍。
- (C) 74. 使用分類六大步驟完成建立分類與預測決策樹不包含
- (A) 資料載入與準備
 - (B) 訓練與測試資料切割
 - (C) 產生訓練資料集的預測結果
 - (D) 預測未知類別的新資料。
- (D) 75. 在建立模型的過程中，需要將原始資料切割成兩部分，第一部分是訓練資料集，第二部分是
- (A) 交叉資料集
 - (B) 缺失值資料集
 - (C) 常數值資料集
 - (D) 測試資料集。
- (A) 76. 分類模型可以由隨機森林方法所建立，而隨機森林方法的方法是衍生於
- (A) 決策樹
 - (B) 倒傳遞類神經方法
 - (C) 關聯規則
 - (D) 集群分析。
- (A) 77. 關於 kNN 預測分析以下何者正確
- (A) 設定 k 必須要是奇數
 - (B) 只能餵入偶數筆資料
 - (C) 預測結果可以有多个答案
 - (D) 設定 k 必須要設定成 0。
- (A) 78. 分類模型又可以稱為
- (A) 分類器
 - (B) 分歧器
 - (C) 橢圓曲線加密器
 - (D) 迴歸解密器。
- (B) 79. 所謂「近朱者赤，近墨者黑」的觀念類似下列哪一種商業數據分析的技術
- (A) 決策樹
 - (B) k 最近鄰演算法
 - (C) 簡單貝氏法
 - (D) 奇異值分解法。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 80. 建立模型過程中，基本上會將資料集切割成兩部分，其一為測試資料集，另一為
- (A) 開放資料集
 - (B) 訓練資料集
 - (C) 結案資料集
 - (D) 半折資料集。
- (A) 81. 整個建立模型過程中的啟動點在於資料的取得與區分，如果將測試資料集與訓練資料集合併在一起就等同於
- (A) 原始資料集
 - (B) 開放資料集
 - (C) 驗證資料集
 - (D) 鑑證資料集。
- (C) 82. 如果使用 kNN 建立分類模型，則其原理與 CART 決策樹分類模型
- (A) 相同
 - (B) 偽陽性資料量過半則相同
 - (C) 不同
 - (D) 偽陰性資料量過半則相同。
- (B) 83. 在訓練 kNN 的分類模型程中，最近鄰資料點基本上可以仰賴什麼機制完成
- (A) 資料點之間的虛值大小
 - (B) 資料點之間的距離
 - (C) 資料點之間的空值數量
 - (D) 資料點之間的噪音聲量。
- (D) 84. 以 kNN 來方法建立分類模型屬於非常簡單就可以完成，在機器學習眾多方法中歸屬於
- (A) 非監督式機器學習
 - (B) 強化式學習
 - (C) 深度學習
 - (D) 監督式機器學習。
- (B) 85. 關聯規則實作上利用 Apriori 尋找客戶的消費習性的方式非常簡單，只需要事先準備兩個欄位就可以執行 Apriori 找出有趣的商品之間的關聯規則，除了交易的商品編號與名稱欄位之外還需要什么欄位資料
- (A) 供應商編號
 - (B) 交易編號
 - (C) 製造商編號
 - (D) 緩衝鍵編號。
- (D) 86. 在決策樹的模型建立步驟中經過測試資料測試過後的學習模型稱為
- (A) 供給彈性模型
 - (B) 第一正規化模型
 - (C) 物件導向模型
 - (D) 分類模型。
- (A) 87. 分類與迴歸樹方法的英文名稱是下列哪一個選項
- (A) CART
 - (B) PARK
 - (C) Apriori
 - (D) FP-Tree。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 88. 請問下列哪一個選項是 Classification 的中文名稱
- (A) 分群
 - (B) 分類
 - (C) 聚類
 - (D) 集群。
- (A) 89. 綜合多棵決策樹分類結果的方法被稱為
- (A) 隨機森林(Random forests)
 - (B) 梯度森林
 - (C) 獨熱編碼分類法
 - (D) 關聯規則。
- (B) 90. 有句俗語「三個臭皮匠勝過一個諸葛亮」，這句話也提醒大家集思廣益也能想出贏過諸葛亮的想法，上述說明其實就是下列哪一種演算法的概念
- (A) 貝氏分類法
 - (B) Random forests
 - (C) Association Rules
 - (D) K-Means。
- (C) 91. CART 決策樹所描述的過程，不管是使用何種方法所產生的決策樹其實都是一個資料分類模型，即每一棵決策樹都可以被視為一個
- (A) 關聯規則
 - (B) 相似集群
 - (C) 分類器
 - (D) 資料表。
- (D) 92. 下列哪一個選項的中文名稱被稱為隨機森林
- (A) CART
 - (B) Apriori
 - (C) One-hot encoding
 - (D) Random forests。
- (A) 93. 隨機森林的方式就是綜合多個分類器的結果，這就好像每一個分類器都是一個
- (A) 領域專家
 - (B) 銷售員
 - (C) 班長
 - (D) 採購人員。
- (B) 94. 隨機森林的方式綜合多個分類器的結果，最後猶如一場會議投票表決計票方式，以什麼方式來當作答案，如此那將會是綜合全體意見的答案，其預測正確率將可以有所提升
- (A) 取最低正確率者為決定
 - (B) 取多數決定
 - (C) 取正確率為 50% 者為決定
 - (D) 取少數決定。
- (C) 95. Random forests 的中文名稱稱為
- (A) 貝氏隨機
 - (B) 梯度下降
 - (C) 隨機森林
 - (D) K 平均法。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (A) 96. 下列哪一選項是模仿生物神經網路(特別是大腦)的方式來建立資料分類模型與預測
- (A) 類神經網路
 - (B) 隨機森林
 - (C) k 最近鄰演算法
 - (D) 關聯規則。
- (B) 97. Perceptron 是一種簡單形式的人工神經網路，稱為
- (A) 運算元
 - (B) 感知機
 - (C) 運算子
 - (D) 關聯表。
- (A) 98. 類神經網路(Artificial neural network, ANN)又可稱為人工神經網路，經常簡稱為
- (A) NN
 - (B) AI
 - (C) BI
 - (D) ERM。
- (C) 99. 下列哪一個選項不是類神經網路中 Perceptron 的名稱
- (A) 感知器
 - (B) 感應機
 - (C) 純量機
 - (D) 感知機。
- (A) 100. 遠在 1943 年就有學者 McCulloch 和 Pitts 首度提出用數學模型來模擬生物大腦的神經元(Neurons)、神經網路，那時電腦還沒被發明出來，直到 1946 年人類才創造了第一部電腦，稱為
- (A) ENIAC
 - (B) 算盤
 - (C) 加法器
 - (D) 差分機。
- (C) 101. 類神經網路發展過程中曾遇到了計算技術上的瓶頸，當時只要神經網路架構太多層就幾乎沒有分類與預測的良好效果，此即梯度計算上不太穩定所產生的問題，如果離輸入層較近的隱藏層，其梯度值可能因為太小而消失，進而無法正確修正相關的權重參數值，稱為
- (A) 梯度爆炸
 - (B) 梯度加強
 - (C) 梯度消失
 - (D) 梯度平方。
- (A) 102. 生成式人工智慧的主要目的是
- (A) 創造新內容
 - (B) 存儲數據
 - (C) 分析數據
 - (D) 傳輸資訊。
- (A) 103. 以下哪一項屬於生成式人工智慧的應用
- (A) 文本生成
 - (B) 數據清理
 - (C) 資料儲存
 - (D) 數據壓縮。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 104. 生成式人工智慧主要依賴什麼類型的數據
- (A) 結構化數據
 - (B) 非結構化數據
 - (C) 文本數據
 - (D) 音頻數據。
- (A) 105. 生成式人工智慧通常使用以下哪種技術
- (A) 深度學習
 - (B) 專家系統
 - (C) 決策樹
 - (D) 線性回歸。
- (B) 106. ChatGPT 屬於生成式人工智慧中的哪一類應用
- (A) 圖像生成
 - (B) 文本生成
 - (C) 數據分類
 - (D) 聲音合成。
- (B) 107. 生成式人工智慧需要什麼來進行學習
- (A) 標籤數據
 - (B) 大規模數據
 - (C) 壓縮數據
 - (D) 單一數據。
- (B) 108. SVM 中的“邊界”是什麼
- (A) 分類準確度
 - (B) 支持向量到超平面的最短距離
 - (C) 所有數據的平均值
 - (D) 模型訓練時間。
- (B) 109. 什麼是支援向量
- (A) 與分類結果無關的數據點
 - (B) 距離超平面最近的數據點
 - (C) 最大的數據點
 - (D) 隨機選擇的數據點。
- (B) 110. SVM 使用什麼來處理非線性數據
- (A) 線性回歸
 - (B) 核函數
 - (C) 深度學習
 - (D) 資料清理。
- (B) 111. 核函數的主要作用是什麼
- (A) 減少計算成本
 - (B) 將數據轉換到更高維度
 - (C) 降低錯誤率
 - (D) 增強模型速度。

中華企業資源規劃學會 專業認證

商用數據應用師-參考題型

- (B) 112. “硬邊界” SVM 的特點是什麼
- (A) 允許部分分類錯誤
 - (B) 不允許任何分類錯誤
 - (C) 僅適用於非線性數據
 - (D) 使用核函數計算。
- (B) 113. 支援向量的主要作用是什麼
- (A) 增強模型速度
 - (B) 支持和確定分類邊界
 - (C) 減少錯誤率
 - (D) 增加計算資源。
- (B) 114. 核函數的作用是什麼
- (A) 減少數據量
 - (B) 在高維空間中找到分隔邊界
 - (C) 增強數據存儲
 - (D) 降低模型複雜度。
- (C) 115. SVM 模型可以應用於哪種類型的數據
- (A) 僅限線性數據
 - (B) 僅限非線性數據
 - (C) 線性與非線性數據
 - (D) 僅限結構化數據。
- (B) 116. 什麼情況下需要使用軟邊界 SVM
- (A) 數據是完美分隔的
 - (B) 數據有少量噪聲
 - (C) 數據點非常少
 - (D) 數據是完全非結構化的。
- (D) 117. 核函數有哪些常見類型
- (A) 線性核
 - (B) 多項式核
 - (C) RBF 核
 - (D) 以上皆是。
- (B) 118. 下列哪一選項是透過不同的間隔最大化方式建立分類模型
- (A) FP-Tree
 - (B) 支援向量機 SVM
 - (C) 3NF
 - (D) Apriori。
- (C) 119. 下列哪一選項是透過建立決策邊界方式建立分類模型
- (A) Apriori
 - (B) DBScan
 - (C) 支援向量機 SVM
 - (D) Association Rules。