

2024 年【科學探究競賽-這樣教我就懂】

大專/社會組 科學文章表單

文章題目： 探索機器學習的奇妙世界：從數據到智慧的奇蹟

摘要：介紹機器學習的運作原理、訓練模型的重要步驟、機器學習的兩種主要類型，以及機器學習在生活中的應用。通過將數據轉化為有意義的模式和趨勢，機器學習使我們能夠做出更準確的預測和更明智的決策。隨著科技的不斷發展，機器學習將繼續為我們的生活帶來更多的便利和驚喜。

文章內容：(限 500 字~1,500 字)

在當今的科技世界中，機器學習正不斷地融入我們的日常生活。從手機中的個性化推薦系統到 ChatGPT 的流行，機器學習已經開始改變著我們的生活方式和工作方式。然而，對於許多人來說，機器學習仍然是一個充滿神秘感的領域。現在，讓我們一同踏上這段探索之旅，揭開機器學習的奧秘吧！



圖一、由 Copilot 生成的 AI 意象圖

機器學習是如何運作的？

想像一下，當我們看到一個物體時，我們會**觀察其特徵**，比如形狀、顏色、大小等，然後**根據這些特徵進行分類和推斷**。機器學習也是透過類似這樣的原理來運作的，通過分析數據中的特徵來實現的。它從**大量的數據中學習模式和趨勢**，然後使用這些模式和趨勢來**做出預測**或做出決策。例如，當訓練一個圖像識別模型時，模型會從數據中學習不同物體的特徵，比如形狀、紋理、顏色等，然後根據這些特徵來識別新的圖像。通過不斷地調整和優化模型，機器學習可以不斷提高其預測準確性和性能。

訓練模型時有哪些重要的步驟？

1. 選擇適當的訓練資料

在機器學習中，首先要選擇訓練的資料是非常重要的。這些資料是機器學習的基礎，就像是學習的教科書一樣。如果我們提供的資料無效或含有偏差，如同教科書內容錯誤，機器學習的結果可能會失真或不準確。因此，我們需要確保選擇的資料具有代表性並且涵蓋了我們想要學習的範疇。這樣機器才能從中學到有用的知識，並且做出正確的預測。

2. 特徵萃取並選擇適當的特徵

特徵就像是描述資料的關鍵特點，而萃取和選擇特徵的過程就像是從一堆資料中挑出最重要的部分一樣。首先，我們將原始資料轉換成機器易於理解的數字形式，這就是特徵萃取。接著，我們要挑選哪些特徵對我們的目標最有幫助，就像從堆積如山的書中挑選出最重要的知識一樣。這個過程可能需要試驗或使用一些特殊的方法，幫助我們找到最好的特徵組合。簡單來說，特徵萃取和選擇就是幫助機器找出資料中最重要的部分，讓它可以更好地學習和做出準確的預測。

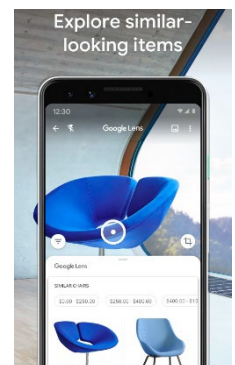
3. 構建模型

這個階段就像建造房屋需要合適的工具和材料一樣，機器學習需要選擇合適的演算法來處理資料並建立模型。這些演算法就像是建築師的藍圖，引領著機器如何從資料中學習和做出預測。不同的問題和資料可能需要不同的演算法，就像不同的建築項目需要不同的設計一樣。因此，我們需要仔細選擇適合特定任務的演算法，以確保模型的準確性和效率。

機器學習的種類有哪些？

機器學習主要可以分成監督式學習與非監督式學習兩種。在監督式學習中，我們向機器提供帶有標籤的訓練資料，也就是說，我們告訴機器每個輸入資料對應的正確輸出。機器通過分析這些帶有標籤的資料，學習到輸入與輸出之間的關係，從而建立一個預測模型。在訓練過程中，我們持續提供新的資料和正確的標籤，使機器能夠不斷地學習和改進；而在非監督式學習中，我們僅向機器提供未標籤的資料，讓機器自行發現資料中的結構和模式。機器通過分析資料的相似性和差異性，尋找資料之間的隱藏關係，從而對資料進行分群或是找出其中的規律性。非監督式學習通常應用在資料探索和分析中，幫助我們發現新的見解和趨勢。

機器學習正在改變我們的生活方式和工作方式，從個性化的推薦系統到手機上的智慧鏡頭，機器學習的技術正不斷地拓展我們的想像空間。隨著人類對數據和技術的理解不斷深入，我們可以期待機器學習為我們帶來更多的創新和驚喜。儘管機器學習還存在著挑戰和限制，但它無疑是人類追求更美好未來的重要工具之一。



圖二、智慧鏡頭

參考資料

1. 機器學習的機器是怎麼從資料中「學」到東西的？超簡單機器學習名詞入門篇！
<https://kopu.chat/ai-terms/>
2. 什麼是機器學習？
<https://www.opentext.com/zh-tw/what-is/machine-learning>
3. 科普 | 5 分鐘包你看懂「機器學習」
<https://kknews.cc/zh-tw/news/p48v53e.html>

註：

1. 未使用本競賽官網提供「科學文章表單」格式投稿，**將不予審查**。
2. 字數沒按照本競賽官網規定之限 500 字~1,500 字，**將不予審查**。
PS.摘要、參考資料與圖表說明文字不計入。
3. 建議格式如下：
 - 中文字型：微軟正黑體；英文、阿拉伯數字字型：Times New Roman
 - 字體：12pt 為原則，若有需要，圖、表及附錄內的文字、數字得略小於 12pt，不得低於 10pt
 - 字體行距，以固定行高 20 點為原則
 - 表標題的排列方式為向表上方置中、對齊該表。圖標題的排列方式為向圖下方置中、對齊該圖