

數據深海中的羅盤：

特徵提取、向量嵌入與相似性搜索的融合之旅

在這個資訊爆炸的時代，數據如同浩瀚的海洋，其中蘊藏著無限的知識和機遇。面對這海量的數據，如何有效地挖掘並提取出有價值的信息，成為了當代科技領域中一項重大的挑戰。特徵提取、向量嵌入與相似性搜索，這三項關鍵技術，如同數據海洋中的羅盤和航標，引導我們在廣闊的數據中尋找並抵達有價值的目的地。

特徵提取：精煉數據的藝術

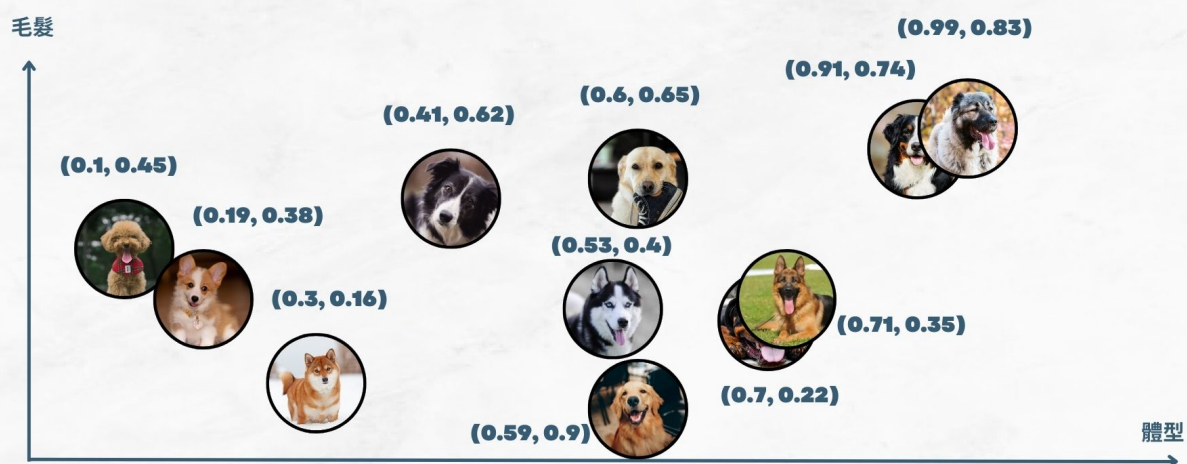
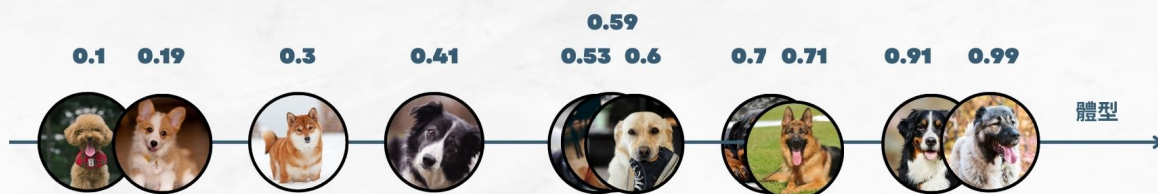
特徵提取是從原始數據中識別出最具代表性的信息的過程，它要求我們將數據轉化為更加表徵性的形式。這一過程本質上是一門藝術，它需要人們深入數據的內在結構，提取出能夠代表數據本質特徵的關鍵信息。無論是在文本分析中選取有效的關鍵詞，還是在圖像識別中識別出關鍵的形態特徵，特徵提取都扮演著至關重要的角色。

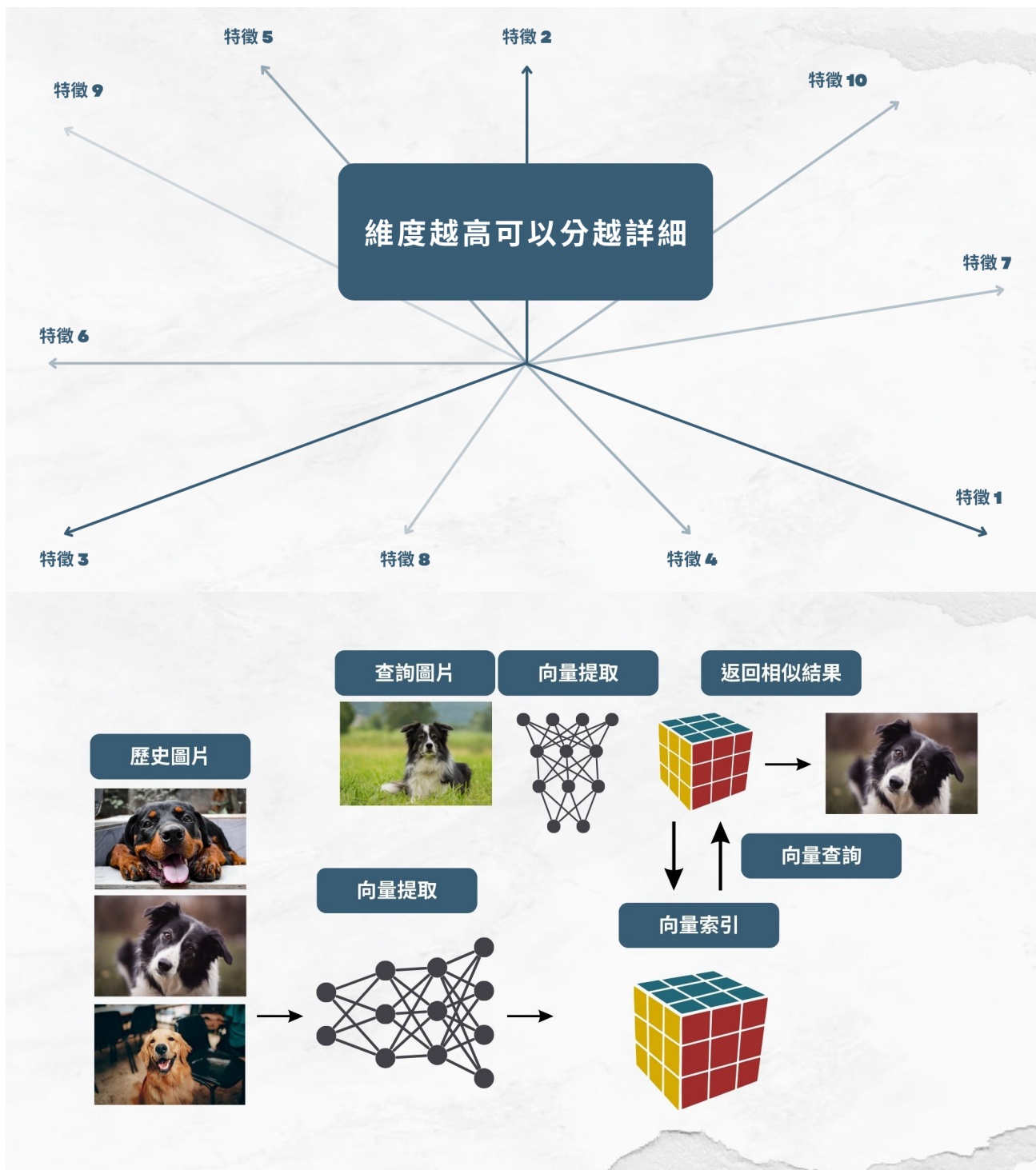
向量嵌入：維度的魔術師

隨著人工智慧和機器學習技術的發展，向量嵌入成為了一種將特徵轉化為向量形式的強大工具。這些向量，或稱為嵌入，能夠在高維空間中精確地捕捉和表達數據的細微差異和深層關聯。這一過程彷彿一位魔術師，將數據轉化為了一系列的數學表徵，這些表徵富含豐富的語義信息，為後續的數據分析和機器學習提供了堅實的基礎。

相似性搜索：高維空間中的導航系統

當我們擁有了向量化的數據表示後，相似性搜索便成為了一種強有力的工具，幫助我們在高維空間中定位和檢索訊息。透過計算向量間的距離或相似度，我們能夠快速找到與特定查詢最相似的數據點。這一過程就像是高維空間中的導航系統，指引我們在複雜的數據結構中迅速找到所需的訊息。





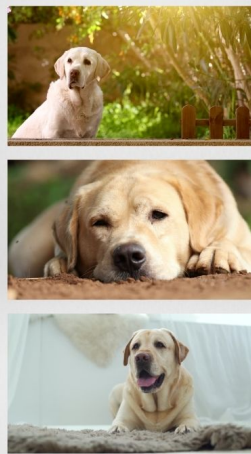
圖片向量化

以圖搜圖



影片向量化

相關影片推薦



商品向量化

瀏覽商品推薦

