

AI 讓下一代資料儲存效率倍增

■作者：Noam Mizrahi

Marvell CTO 辦公室科技副總經理暨首席科學家

今日，資料的產生速度超乎人們想像。往昔人類是主要的資料產生來源。現在則不僅有成像裝置、感應器、無人機、連網汽車、IoT(物聯網)裝置，還有許多工業設備以多元的方法和格式產生著資料。不過，區分資料和資訊這兩個詞彙至關重要，千萬不能混淆。

目前所收集到的資料中，只有一小部分稱得上是真正的資產，擁有足夠的價值。我們以成像裝置為例。在長達幾小時的影像短片中，只能捕捉到一分鐘有意義的活動，其餘部分可有可無、異常瑣碎。打個比方，「資料」就好比金礦，而「資訊」則是人們夢寐以求的金子。將此等資料轉化為寶貴資訊的能力，即所謂的「挖掘能力」，就是「分析」。

在分析公司 Statista 製作的圖表 1 中，可以看到在過去十年間，資料儲存量的增長速度簡直驚人。其中還預測出，到 2020 年，儲存需求將超過 42,000 Exabyte。然而，多數(估計至少有 80%)已儲存的資料毫無結構化可言，這就給使用資料進行

分析造成了難度。據估計只有 5% 已儲存的資料得到過分析。若我們有辦法用中繼資料在實際分析的情境下，有效描述這些非結構化資料，那麼可供分析的資料量就要大得多。此舉可以讓組織從擁有的資料中挖掘出比以往明顯更多的價值。

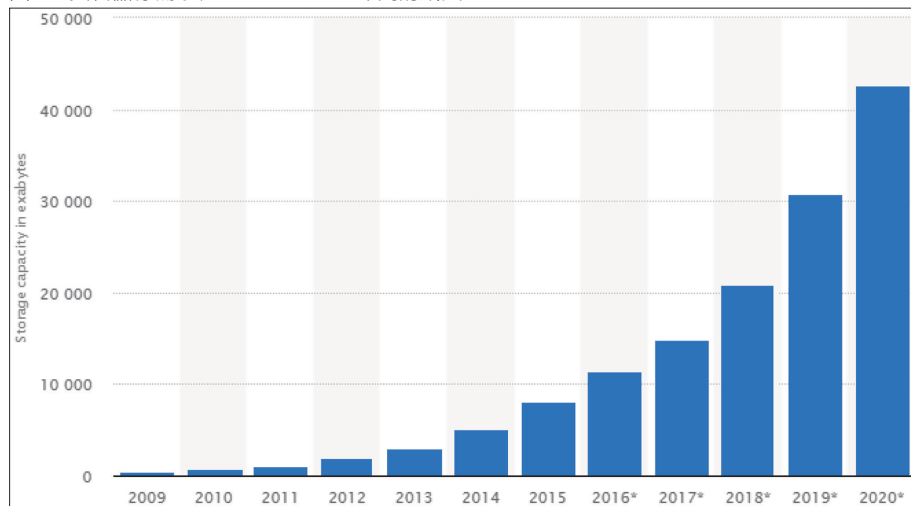
人工智慧(AI)科技必將對現代社會各個層面帶來深遠影響。電子商務推薦、自然語言翻譯、金融科技、安全、物體辨識/偵測等領域，無一例外都會受到影響；甚至在醫學方面，也能藉助 AI 快速查出威脅生命的癌細胞(或其他病變)。儘管應用領域十分多元，但我們現在有套科技，可以有效掃描眾多非結構化資料(影片、文字、聲音、圖像等)並予以處理，從而分離出真正有價值的資訊；這裡面有種萬？不離其宗的意味。

更精確地說，利用 AI 我們不僅能執行分析過程，還可以對非結構化的原始資料進行前置處理，得到經標記的中繼資料，以簡而精的方式描述這些非結構化資料。可以使用上層分析軟體分析這個簡

化過的資料庫，並蒐集有用資訊。一直以來，組織都渴盼用 AI 從他們擁有的資料中挖出更多東西，卻很遺憾的求而不得。

不過現在好了，我們想要產生中繼資料，讓分析軟體更有效地運作？有了 AI 工具，我們完全可以從龐大的非結構化資料庫中，建立起想要的那座中繼資料庫。這樣一來，就只剩下將大量資料匯入 AI 計算實體啦。打住，這麼走對嗎？

圖 1：資料儲存需求在 2009 至 2020 年間的成長



真的？

考慮到現如今「雲端」和「邊緣」是產生和儲存資料的兩大所在，您瞬間就能明白四處移動大量資料所費不貲，著實應當加以避免。在雲端，將這些資料路由到資料中心，會影響構成網路的基礎結構，消耗許多電力並提高延遲程度（因此會增加整體處理時間）。而邊緣這一側，計算和電力資源有限。其中小型裝置的網路功能亦受限，這就使得上傳大量資料至雲端做處理特別不切實際。在這兩例中，關鍵在於極力縮減要移動的資料量，並以中繼資料取而代之，方能將操作效率最大化。

比起四處移動資料，在源頭分配中繼資料的做法來得更有效率；所謂源頭也就是儲存資料的裝置。SSD（固態硬碟）已經包含了用作計算實體的基本元素。這類儲存裝置通常僅用於硬碟運作，但也可重新規劃目的，去執行功能相關的工作，並負責標記資料，或是輔助整合硬體 / 軟體 / 韌體組塊，來執行上述功能。其中一種操作模式，是利用硬碟的閒置視窗執行背景對應工作。另一種方法，則是在資料寫入硬碟的同時進行處理。在儲存時部署此類加速方法，若能正確應用，好處信手拈來：像是節省電力和花費、最小化資料移動、大幅減少延遲，以及降低整體網路流量。而且此方法還有其天然的延展性，讓企業和雲端服務供應商可利用 AI 的力量，拓展各自的能力範圍。

於去年八月在聖克拉拉舉行的快閃記憶體高峰會中，Marvell 演示了突破性的 AI SSD 控制器，力證資料標記無須存取主機 CPU（中央處理器）處理資源，也能有效執行的概念；而且此舉還能避免先前提過的花費和延遲問題。與會者見識了現成的 Marvell 資料中心和客戶端 SSD 控制器 IC（積體電路），在結合開放原始碼的 NVIDIA 深度學習加速器（NVDLA）科技後，如何沿用經訓練的 AI 模型、編入整合 AI 推斷 IP，並掃描儲存於硬碟本機內並充滿非結構化資料的大型資料庫（例如影片資料庫）。由此可產生標記，建立起一座中繼資料庫，方便在搜尋時描述原始資料。

若以偵測和辨識物體或場景為目標，AI 推斷引

擎可掃描儲存於硬碟中的影片檔，並建立中繼資料，其中標記它們在影片中的出現時間。拜此一新 AI 增強型儲存科技所賜，讓中繼資料庫可儲存於 SSD 本機，且可供分析軟體按要求進行檢查。

舉個例子，執法部門想從超長的影片檔中找出一處疑點。他們可以載入一個經訓練的模型，該模型知道如何準確辨識那處疑點，並可在背景中跨所有載入該模型的硬碟，平行推斷一切可用影片內容。該處疑點的任何足跡都會被標示和標記，待後續的進一步分析創造便利。

同樣地，試想這個架構若用於背景聊天機器人分析，效果該有多麼好。一般為了改善服務品質，需要掃描聊天機器人通話的大型資料庫。衡量使用者對收到的回應是滿意或不悅、通話時間太長或太短，統統不再只是夢想。建立一個懂得如何追蹤這些衡量標準的 AI 模型，就能將它們編入 AI 儲存推斷引擎，以及離線掃描的通話。以點播視訊（VOD）服務中的個人化廣告置入來說，在人物或物品搜尋及眾多額外大量的 IO（輸入 / 輸出）密集型使用案例中，利用與資料的接近程度，可取得重大的效能效益。

Marvell 率先推出的 AI SSD 控制器科技，展示出新資料儲存架構不必透過昂貴的客製化 IC，即可用於處理範圍益發龐大、運算難度又較高的「大數據」相關應用程式。透過讓市場現有的 SSD 硬碟得以存取可大幅增強智慧的附加邏輯，可直接處理對下一代分析工作負載至關重要的中繼資料和標記。無須連線至專門的處理資源。

使用有別於依賴傳統中央處理模式的另類招數，可讓整個過程更有效率。它只需占用最小的可用網路頻寬，而且不會遭遇瓶頸。AI 加速器直接合併至符合成本效益的 SSD 控制器 IC 中，使得更快地完成分析工作成為可能。不僅減少處理量、節省可用電力預算，同時還完全避開了從頭打造複雜 ASIC（特定應用積體電路）的需求。因為用的是可程式化架構，有充足的佈建可資升級已採用的 AI 模型，如此也能在遇有隨時出現的新使用案例時一舉解決。 **CTA**