

【教學任意門】大數據資料分析介紹與教學應用

學習新視界

【記者柯家媛淡水校園報導】由本校學習與教學中心教師發展組主辦之「教學系列技巧講座：大數據資料分析介紹與教學應用」於9月27日在淡水校園覺生綜合大樓I501，以及蘭陽校區CL423同

步進行。本次演講邀請教學特優教師—統計系副教授陳景祥，介紹大數據資料分析及分享大數據應用在課程教學的經驗。

陳景祥首先列舉自身研發經驗，如：遙控飛機、金屬蒸汽小火車、電動三輪車、自學電鉗、三台自製咖啡烘豆機、作曲……等，且表示退休後想在一塊空地架一道鐵軌，製作出縮小版蒸氣推動載人火車。以此說明「跨領域」與「多領域」的不同，跨領域是指以某個專業為主，其他為業餘興趣或嗜好；而多領域是指一個人的一生可以有多種興趣與專長。陳景祥表示：「人生是不斷學習的過程。這輩子沒學完的、下輩子再學」，並勉勵大家往「多領域」發展。

對於「大數據」，陳景祥先從4V來定義何謂Big Data。4V分別為：Volume(資料量)、Velocity(速度、時效性)、Variety(資料型態多樣性)、Veracity(資料真實性、可信度)。其中，陳景祥認為Velocity(速度、時效性)最為重要。而依照數據量大小定義Big Data，則是單一文件的資料量以TB~PB為單位做計算。

對於大數據應用的實例，例如：2012年美國大選，總統歐巴馬的競選團隊透過大數據來分析歐巴馬會在哪些州勝出、日本衛生機關利用大數據預測流感人數，陳景祥表示可信度還待商榷。台灣也曾有媒體報導超商業者累積了巨量資料而歸納出幾種消費現象，陳景祥表示絕大部分宣稱用大數據分析的新聞與宣傳，其實資料量並沒有那麼大，幾億筆資料未必就等於大數據。這些媒體並非真正了解大數據，造成名詞亂用、誤導閱聽者。大數據的應用尚未成熟，現今仍用傳統的分析技術，記數(Counting)、平均值(Average)、比例(Ratio)，來進行數據的分析。

坊間也有許多「5到10年內，AI人工智慧就會取代現在人力資源」諸如此類的迷思，陳景祥表示AI要贏過人類必須具備三個條件：第一，環境是要封閉的；第二，規則是完備的，不能隨時變動；第三，約束是有限的，不能無限自我遞迴。對於現今社會仍充斥著變數，所以陳景祥認為AI並無法完全取代人類。另外，陳景祥引用「中文字坊間思想實驗」說明機器可以完成許多聯結規則，來找出問題的答案，但未必是智慧。

總結整個演講，陳景祥說明大數據分析運算類型分為兩種：資料處理及圖形展現。大數據分析案例在課堂中主要效用是「說服」與「參考」，最常見的是圖形展現。因圖型對學生來說是最據有說服力和容易理解的。舉例來說，在 R-bloggers 網站有一篇關於美國紐約市計程車與Uber載客率比較的分析文章，內含許多淺顯易懂的圖形，可讓學生快速理解 Uber 在美國大城市的快速成長狀況。

台灣甚至東方世界，真正缺乏的是「數據分析」，而非「大數據分析」，這牽涉到尊重專業與遠見視野的文化問題，並非科技問題。（編輯／梁淑芬）

2018/10/23



陳景祥老師精彩演講讓與會老師們聽得入神。（攝影／黃柏耀）



本校學習與教學中心教師發展組主辦之「教學系列技巧講座：大數據資料分析介紹與教學應用」於9月27日在淡水校園覺生綜合大樓I501，以及蘭陽校區CL423同步進行。（攝影／黃柏耀）

