

銘傳Moodle大數據分析 與學生學習成效

■ 文／王金龍・銘傳大學學術副校長

資訊科技在高等教育的應用，一開始是應用於科學研究上，之後也用在校務行政MIS系統上以提升營運的效率，近期以來資訊科技更大量運用在學生的學習面上，其中開源碼的Moodle教學管理系統（Learning Management System）提供強大的學習環境，也記錄了許多師生的教與學過程，更為大數據的搜集提供穩定的平台。銘傳Moodle教學平台自95學年度啟用至今，每天有3至4萬人次的登入，九年來已累積近4千6百萬的登入人次。以下就大數據帶給教育領域的影響，以及銘傳如何開採Moodle寶礦的技術方法，建立預測診斷模型及早期預警的研究案例進行分享。

何謂大數據

隨著網路的普及和雲端運算的熱門發展，「大數據」（Big data，或稱巨量資料、海量資料、大資料）的應用潮流，已經遍布在各行各業，廣泛地應用在我們的生活中。這個從資料探勘（Data Mining）延伸出來的概念，現在重視的不在於分析技術的進步，而是在大量資料本身所蘊含的資訊，以及提供人們對於各種現象新的解讀方式及預測能力。大數據的取得與分析為人們帶來新的思維，顛覆了許多領域的理解與運作模式，從電子商務、社交生活、公共行政、健康醫

療，到教育學習都是（孫憶明，2014）。

全球最大的網路書店亞馬遜（Amazon.com）是大數據運用在商業領域著名的成功案例。當顧客在亞馬遜網站買東西之後，銷售系統就會記錄下顧客買了什麼、甚至還在哪些商品上面做過停留，當下次這位顧客再次進入亞馬遜書店時，系統會主動識別出顧客身分，並根據顧客的喜好做出準確的預測，進行商品的推薦。亞馬遜善用巨量資料，提出高度準確且個人化的購書建議，打造獨特的購書體驗，大數據分析提供客製化訊息的威力，讓許多實體書店無招架之力。

大數據在教育領域的研究走向

美國教育部教育科學研究所（Institute of Education Sciences, U.S. Department of Education）所長John Easton提到，運用數據來提升學校效能與學生學習的時代已經來臨（Mandinach & Gummer, 2013；陳鏗任，2014）。巨量資料的分析在教育領域近年會備受矚目快速推展，主要得益於兩大助力：

● 校務研究對資料分析之需求

校務研究（Institutional Research，簡稱IR）緣起於20世紀初，是美國高等學校內部的一項重要工作，歐美等先進國家的大學皆有設置校務研究辦公室，由研究人員負責校務評鑑、研究與規

劃的工作。研究人員運用科學的質量化統計分析，將相關資料匯整與分析，轉化成有用的資訊，有效建構科學化的校務管理，針對學校管理與運行時面臨的具體問題進行分析與諮詢研究。

臺灣高等教育在經營上面臨越來越複雜的問題與挑戰，例如：高教普及化與少子女化帶來的招生衝擊、教育資源的競爭與競合、大學治理與拓展財務之複雜等，這些複雜的問題已經不能靠過去高校管理者「以經驗為基礎」的傳統管理哲學，希冀藉由校務研究的推動，建構「以科學理性為基礎」的現代大學管理理念，以提升教育品質及校務經營的管理。

● 線上數位學習快速推展帶來巨量資料

近年來隨著線上評量、網路教學、電子書、科技輔助教學設施的進步發展，不只課內課外的教學模式有了許多變革，從學生和老師的教學應用和行為中，也留下許多學習相關的數位化資料，尤其是「大規模開放式線上課程」（Massive Open Online Courses，簡稱MOOCs）學習風潮的興起，全世界著名大學紛紛將課程上網，讓大批學生以免費的方式學習，跨國界跨時空的學習影響力，帶來了一般課堂授課前所未見的巨量學生，這股「巨量資料」趨勢也開始影響教育領域。

透過這些數據的分析，系統可以知道學習者在那個單元比較不懂，會建議學習者再次學習該單元，並且透過單元測驗確認學習者是否完全理解，大大發揮了數位教育的特性與價值，實現適性教學的教育理念。

銘傳Moodle大數據分析

銘傳Moodle教學平台自95學年度開台至今，已累積有近4千6百萬的登入人次、6萬多門的課程量，以及數以億計的學生活動日誌。基於這些

珍貴的巨量資料，若能像Amazon一樣，透過巨量資料的導入，分析每位學習者的個性化與多元化發展需求，研究學生的學習軌跡與教師的教學歷程，便能打造一個智慧的教與學平台，提供學生一個適性學習的環境，並對學習者的特徵與行為做出預測，做到早期預警的提醒與輔導。

資料探勘應用在數位學習系統中是個反覆的循環程序，整個過程牽涉了大量的準備工作與規劃過程，此一循環的過程不僅是為了將龐大的資料淬煉成知識，並可過濾探勘的知識去支援決策的判定。銘傳Moodle大數據分析的過程是由以下一般資料探勘的四個步驟所組成（Romero, Ventura & Garcia, 2008）：

● 收集資料（Collect data）

教學管理系統通常將學生資料、使用歷程、互動資訊儲存在資料庫，我們在Moodle系統中收集了學生參與學習活動歷程的log檔資料。

● 前置處理資料（Preprocess the data）

資料須先清理和轉換成合適的形式，去除資料錯誤與不一致的問題，進行格式轉換，將log資料經統計成為關聯性資料表格，之後將資料進行離散化、分類等轉換的動作，此可使後續的階段更順利進行，是個不可輕忽的重要基本工作。

● 進行資料探勘（Apply data mining）

使用資料探勘演算法構建和執行模型，去發掘並總結出我們感興趣的知識。在這步驟，我們採用了自由軟體Weka作為資料探勘的工具。

● 說明、評估和運用結果（Interpret, evaluate and deploy the results）

得到的結果或模型可以讓教師說明和使用，運用在後續的教學行動上。透過這些已發掘的資訊，教師可以針對學生的學習進行預測，以及對Moodle上的課程活動進行調整，以提升教學和學習成效。

Moodle資料探勘技術與教學應用

資料探勘的方法和其相關的演算法有很多，例如：分類（Classification）、分群（Clustering）、迴歸分析（Regression）、關聯性規則（Association rule mining）、社會網絡分析（Social network analysis）、文本挖掘（Text mining）等，以下僅就銘傳在執行Moodle大數據分析的過程中，主要運用的技術與方法加以說明：

● 統計（Statistics）

Moodle沒有提供統計模組讓教師可以針對每個學生學習表現的詳細狀況進行查詢，其報表功能只能查詢單一活動或資源的區間狀況，我們的技術團隊即針對每個使用者參與教學活動的狀況，自行開發了一個總和性的統計報表（圖一），讓教師對於學生的課程參與度一目瞭然。

但在測驗方面，Moodle提供很好的統計模組（圖二），包括試題的難易度、標準差、鑑別係數等數據，教師便能透過上述的數據分析，進行試題和試卷的維護與調整；透過學生的答題狀況紀錄，教師能夠掌握學生知識點的學習障礙，了解學生的學習狀況，為教師在教學決策時提供了有用的參考依據。在一般紙本評量中，教師通常僅能針對最後的考試分數對學生的學習成果進行評價，對於試題分析、試卷分析、學生題目完成情況的

10302-05351APP 程式設計

使用者	課程檢視	發表文章	回覆文章	檢視討論區	檢視文章	線上資源	上傳檔案
94 14 郭	570	4	0	33	4	76	0
01 602 張	46	0	0	1	0	47	0
00 101 陳	82	0	2	5	5	44	1
00 374 黃	6	0	0	0	0	1	1
00 896 曹	119	0	2	6	6	101	1
00 060 莊	85	0	1	3	2	74	1
00 030 張	35	0	2	4	6	4	1
00 171 王	107	0	2	3	7	91	1
00 296 黃	55	0	2	2	4	31	1
00 1126 張	51	0	2	2	4	27	1
00 1196 林	85	0	2	7	6	45	1
00 054 劉	61	0	2	3	7	25	1
00 044 陳	100	0	2	6	6	54	0
00 42 陳	83	0	1	3	5	38	1
00 32 黃	73	0	2	4	5	32	1
00 66 安	80	0	1	2	3	10	1
00 80 甘	63	0	2	2	5	34	2
00 53 楊	86	0	2	2	4	67	1

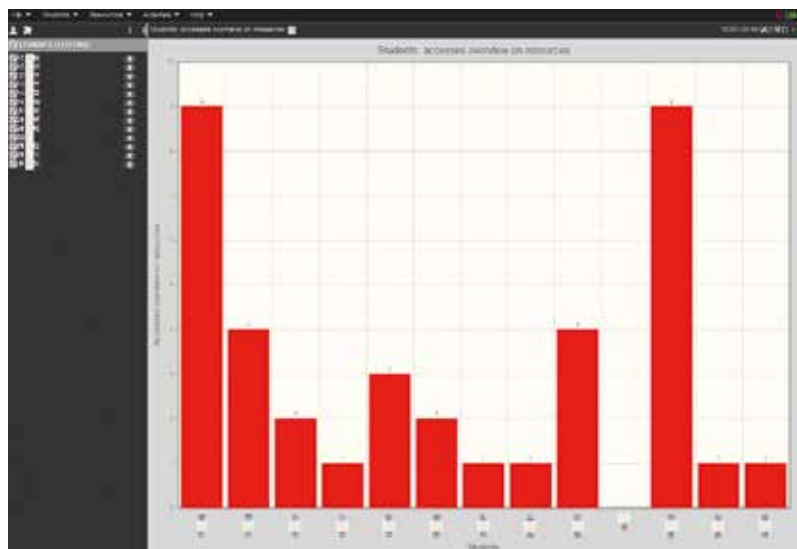
圖一 銘傳大學開發的統計日誌功能

第一次作答的標準誤									
7.35%									
測驗結構分析									
以右列格式下載 Comma separated values text file 下載									
題號	試題名稱	作答次數	難易度指數	標準差	隨機猜測分數	重要的知識	有效的知識	鑑別指數	鑑別係數
1	MVC	56	90.48%	24.38%	16.67%	15.00%	15.14%	36.23%	43.46%
2	concept	56	86.61%	29.37%		10.00%	12.51%	30.70%	34.68%
3	interface & implementation	56	100.00%	0.00%		5.00%	0.00%		
4	method call	56	93.45%	22.40%		10.00%	10.35%	30.33%	44.58%
5	method declare	56	96.43%	12.99%		10.00%	7.59%	34.17%	48.95%
6	string	56	94.64%	22.72%	25.00%	10.00%	8.13%	9.56%	13.71%
7	macos	56	96.43%	18.73%	50.00%	5.00%	2.11%	-3.35%	-5.94%
8	IBOutlet	56	83.04%	29.04%	0.00%	10.00%	14.45%	54.85%	60.39%
9	IBAction	56	75.00%	31.62%	0.00%	10.00%	14.75%	48.70%	54.61%
10	project template with multiple view	56	87.65%	22.19%		10.00%	10.47%	32.25%	35.35%

圖二 Moodle測驗分析功能

分析，沒有教學管理系統的協助是難以如此深入進行的。

此外，以往探討如何進行學生的學習成果評量時，傳統的教學與測驗評量模式有其侷限之處，只能進行總結性評量，然而在Moodle的環境下，可以即時的獲得學生的測驗結果，並蒐集學生學習歷程的總體表現，提高了形成性評量的可行性。



圖三 GISMO統計圖表

● 視覺化（Visualization）

Moodle對於學生的使用歷程紀錄僅有提供文字資訊，如：log日誌報表、項目分析等，沒有提供視覺化的工具，但是面對密密麻麻的一堆數據與文字報表，實在讓人難以一眼就看出重點為何？有句話說：「一圖抵萬言」，學生資料若能以圖表呈現，便能讓教師和管理者在研究分析數據時更加事半功倍。於是，我們的技術團隊在Moodle系統上引入GISMO模組，它是一個圖形化的互動監控工具，提供教師視覺化的學生活動統計資料（圖三），如：出缺席、閱讀教材、繳交作業等各方面的狀況。

GISMO除了具有將資料視覺化的功能優點外，還能針對全班的狀況提供教師一個綜觀的視覺統計圖表，讓教師可以更容易的觀察和掌握學生學習概況的全貌。

● 分群（Clustering）

分群分析是資料探勘的方法之一，將比較相似的樣本叢集在一起，形成集群（cluster）。以「距離」作為分類的依據，「相對距離」愈近的，「相似程度」愈高，歸類成同一群組。此分析方法為非監督式學習（unsupervised

learning），事先不需要任何的假設，樣本是依靠自身的相近性叢集在一起，其意義是靠事後的集群特徵解釋才能得知，用以找出事物相似性的內部結構。

在教學應用上，教師可以利用分群分析的結果，辨識出學生的學習特性和行為，例如：分析結果將學生分成了三群，群組一的學生共同特徵是沒繳交作業、很低的訊息閱讀數、完成少數的測驗等，這些特徵都代表了群組一的學生學習態度較不積極。以此類推，群組二的學生學習態度普通，群組三則是較積極的學生群，教師便能利用這個分析結果進行合作學習的活動，在分組時教師視教學設計的需要，將同類型的學生編組在一起，或是將三種類型的學生平均分配在小組裡，讓同儕互相討論激盪，激發學習動機。

● 分類（Classification）

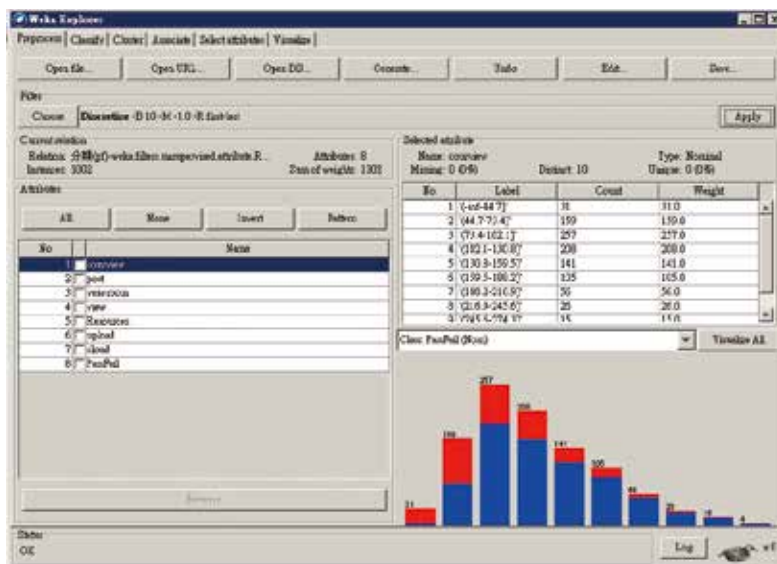
分類分析屬監督式學習（supervised learning），即透過機器學習，預測結果是屬於哪一種類別的方法，例如：利用學生通過小考的次數，預測學生最後的學期成績是及格或不及格。在分類問題中，除了提供預測的分類結果之

外，同時也會提供發生這個分類結果的可能機率。

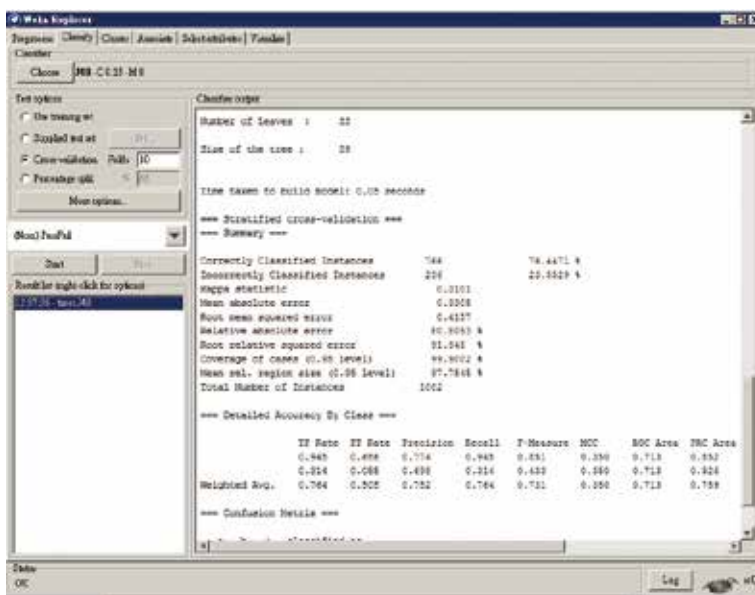
在分類分析的過程中，我們會用一些已經分類的資料來研究它們的特徵，然後再根據這些特徵對其他未經分類或是新的資料做預測。這些我們用來尋找特徵的已分類資料可能是來自我們現有的歷史性資料，或是將一個完整資料庫做部分取樣，再經由實際的運作來測試（資料探勘Data Mining的介紹，2005），例如：我們透過Moodle教學管理系統，追蹤學生上線頻率、瀏覽教材、發表文章、回覆文章等參與學習活動的歷程，以過往的課程學生作為樣本模組，對現行的課程學生進行學習行為和結果的預測；或是，將該學期前幾週學生的學習歷程表現作為樣本模組，對後幾週的學習成效做出預測，幫助教師辨識出有潛在學習困難的學生，便可適時的採取行動通知教師和學生，做到更即時的預警制度，幫助學生達到最大的學習成效，有效地減少輟學和退學率。

研究案例分享

我們以「資訊科技：辦公室應用課程」作為分析樣本，利用銘傳自行研發的統計模組，蒐集學生在Moodle平台中參與課程教學活動的數據，將課程檢視、發表文章、檢視討論區、檢



圖四 Weka樣本資料輸入畫面



圖五 利用J48演算法執行分類的預測結果

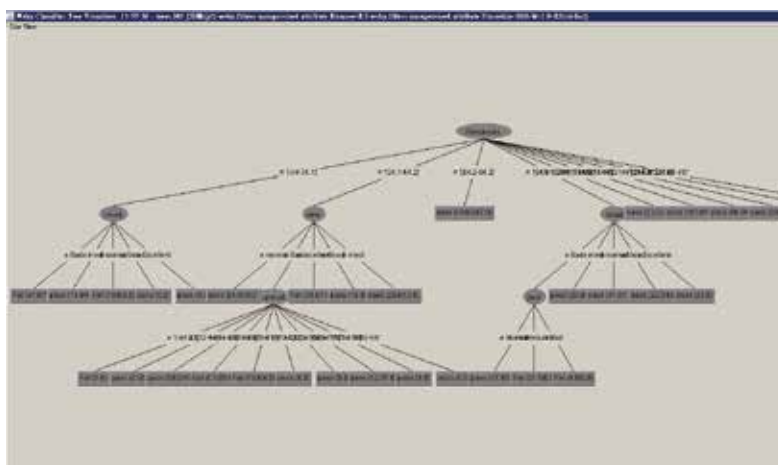
視文章、線上資源、上傳檔案等六個重要教學活動的次數，再加上進入銘傳大學雲端教室操作練習題的次數，對1,002位學生參與TOC Word技能檢定的成績通過或不通過（pass/fail），利用J48決策樹演算法進行建模預測。（圖四、圖五、圖六）

藉由數據的執行結果，預測的準確率可達七成五左右，的確可以作為學習預警的目標，期望作為學生核心重要課程的學習監控。數據分析預測的準確度要高，除該課程在Moodle中必須是門互動活躍的課程之外，針對學生在課程中活動的項目也會影響預測的準確度，所以若要在學校全面大力推動透過大數據來達到課程前期預警的目標，必須引動教師們對數位教學平台有更多的投入和使用。

透過Moodle做出準確的前期學習預警 打造高教體系Amazon

以往各校都相當重視落實期中預警，對學習成效不佳之學生，透過補救課程輔導、課業諮商、學習導師輔導等機制，提升及強化學生學習成效。但是，有時候拖延到期中再介入給予學習輔助已經緩不濟急，我們期盼能藉由Moodle的大數據，打造學習殿堂的亞馬遜，提前在學期初的幾週課程後，便能做出準確的前期學習預警，提供學生個

別化的學習建議，讓教師能更精確掌握學生的學習狀況，達到最佳的教學與學習成效。



圖六 決策樹建模結果畫面



▲銘傳大學積極打造Moodle教學平台，希望成為臺灣高教體系的Amazon。
(學校提供)

◎參考文獻

- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30-37.
- Romero, C., Ventura, S, & Garcia, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384.
- 陳鏗任(2014)。大專院校應用學習分析之概況。*教育資料與圖書館學*，51(4), 597-636。
- 孫憶明(2014年2月10日)。大數據(Big Data)改變未來教育樣貌的三種可能。*The News Lens關鍵評論*。取自<http://www.thenewslens.com/post/24794/>
- 資料探勘Data Mining的介紹(2005)。*資料探勘知識分享*。取自<http://www.uniminer.com/center01.htm>