

學習歷程的資料探勘-以布林邏輯查詢課程單元為例

Data Mining from Learning Portfolio- A case of Boolean Logics Inquiry Course Unit

*
陳昱任

*
南台科技大學資訊傳播研究生
*
yuren0807@yahoo.com.tw

**
楊錦潭

**
南台科技大學資訊傳播研究所
**
yangdav@mail.stut.edu.tw

摘要

本研究旨在探討數位學習歷程之最佳型樣，課程內容以科技大學基礎的計算機概論課程中的「布林邏輯查詢」教材單元。基於適性化教材設計，可以提升學習成效，目前有兩種方式可以達成提升學習成效，第一種方式為由資深教師設計合適的教材瀏覽順序，第二種方式為利用資料探勘找出最佳型樣，本研究為後者。本研究的執行步驟有三：(一)開發符合 SCORM 2004 順序瀏覽規範 (Sequencing and Navigation ; SN) 的「布林邏輯查詢」教材單元；(二)置於 SCORM 2004 執行環境 (Real Time Environment ; RTE) 為實驗平台，並記載學習歷程；(三)以資料探勘進行資料分析。而本研究樣本為科技大學大學部學生 20 人。本研究結果為了找出「布林邏輯查詢」課程單元的最佳型樣，本研究的貢獻為提供給教師在教材內容順序及教學活動之參考。最後，本研究針對未來研究的建議亦一併提出討論。

關鍵詞：數位學習、SCORM 2004 SN、SCORM 2004 RTE、資料探勘、學習歷程、路徑型樣

Abstract

The purpose of this study is to explore the best learning path patterns by data mining with learning portfolio. The course conducted in this study is “Boolean Logics Inquiry” in “Introduction to computer science” at University of Technology level in Taiwan. Based on web-based course design for adaptive learning, it might promote learning effects of learners. There are two approaches to accomplish the goal: the browsing sequences of course materials are decided by senior teachers or by data mining from learning portfolios. This study assumes the latter approach. The process of this study consists of three steps: developing SCORM 2004 SN compliant course unit named as Boolean Logics Inquiry, Recording each student's learning portfolios on SCORM 2004

RTE, and data mining on the learning portfolios. The subjects of this study consist of 20 students at University of Technology level in Taiwan. The result of this study is to find out the best learning path patterns for “Boolean Logics Inquiry” course unit. The contribution of this study provides a reference for teachers who intend to design adaptive learning material on the web. Finally, the implications of this study for future study are also included.

Keywords: E-learning, SCORM 2004 SN, SCORM 2004 RTE, data mining, learning portfolio、path patterns

1.緒論

數位學習(e-learning)的優勢為不僅可以打破空間、時間的限制，且可以隨時的更新教材，並提供使用者獲取最新的教材，而使用者也可以選擇性的讀取教材內容，以提升了學習的成效。(Anderson, & Elloumi, 2003)。在「學習管理系統」(Learning Management System ; LMS)中，學習者瀏覽教材過程中，系統即可記載學習者的「學習歷程」(Learning Portfolio)。因而，可以運用網路上大量的「學習歷程」，進一步針對瀏覽網頁的順序進行「資料探勘」(Data Mining)，從而讓學習者進行適性化的學習。即教材的呈現順序是依照自己本身的學習情況而定。

ADL 國際組織所提出來的「分享教材物件參考模型」(Sharable Content Object Reference Model ; SCORM)來規範數位教材的標準。有了符合 SCORM 教材之後，一份教材就可達成可重複性(Reusability)、可取得性(Accessibility)、可耐用性(Adaptability)、恆久性(Durability)、可互通性(Interoperability)、可負擔性(Affordability)等特徵(林錦泓，2002)。因而，符合 SCORM 規範所製作的學習物件，即可在 SCORM RTE(Run-Time Environment)的相容平台上，讓學習者瀏覽數位教材。

在教材設計方面，可以分為兩種形式：第一種是由老師自己依照自己的教學經驗，自己編寫上課教材內容和順序，在課堂上依照編排的教材順序傳授知識給學生。第二種是透過網路建置的學習網站，自由地讓學生點選的教材內容。前者學生的學習是屬於被動式的照單全收，而缺乏彈性化的教學模式；反之，後者則可以兼顧學生的個別差異，讓學生可以依較佳學習的順序以提高學習成效。畢竟，透過學習歷程的探勘之後而找出的序列型樣(Sequence Pattern)探勘方法，發掘出較具意義的學習瀏覽路徑型樣(path traversal pattern)，並配合課程型樣的挖掘。換句話說，將學習成效良好的學生之學習序列型樣推薦給教師課程設計之參考。

基於以上的所述，本研究採用探勘方法在實驗課程中，進行實徵研究。作為提供教師針對數位教材內容的編序等教學活動設計之參考。

2.文獻探討

配合本研究進行，本節以三個議題進行文獻探討，包括：「SCORM 簡易順序規格」、「資料探勘」與「學習歷程實徵研究」。

2.1 SCORM 簡易順序規格 (Simple Sequencing Specification, SSS)

SCORM 針對數位教材的終極目的就是學習之共享性、可重複性，當教材遵循其所規範之標準，讓這些學習資源能在不同的學習管理系統(LMS)上互通，其規範包含學習資源製定的方法與其所應遵循的事項。執行環境提供了共同的方法以存取遵循標準的共享內容元件(SCO)，讓這些共享元件可以在不同的學習管理系統上互相交流、分享；它也定義了元件存取時的應用程式介面(API, Application Program Interface)及資料模型(Data Model)。(李昇暉、林居鴻、馮志成、游合正、陳怡妃，2003)

SCORM 2004 中的「簡易順序規格」(Simple Sequencing Specification, SSS)包含兩個資料模型，「追蹤狀態模型」與「順序定義模型」：(莊宛毓，2005)

一、追蹤狀態模型 (Tracking Status

Model, TSM)：TSM 定義了當學習者與學習管理系統產生互動時，所應紀錄的相關資料，例如：成績、教材學習次序...等等，其中尚包含三種資料定義：1.課程活動學習資訊 (Activity Progress Information) 2.教學目標學習資訊 (Objective Progress Information) 3.課程經驗學習資訊 (Attempt Progress Information)

二、順序定義模型 (Sequencing Definition Model, SDM)：SDM 定義了在學習活動中，不同之學習順序的次序建議與限制。在 SDM 所制訂的 Metadata 中，有兩種資料模型：1.

順序控制模式 (Sequencing Control Modes, SCM)：SCM 主要定義一般常用的學習路徑型樣，例如：隨機學習 (Randomization)、只向前學習 (Forward Only) 等等。2.順序規則 (Sequencing Rules)：Sequencing Rules 則主要定義規則中的條件與動作，並設定相關標準字彙庫，可以讓教師或是教材編輯者自由填寫教材的建議學習規則，而此規則為 if <condition set> Then <actions>之格式。

2.2 資料探勘

資料探勘 (data mining) 就是從「大量」資料中獲取「知識」或者擷取先前未知的有價值之資訊或資料型樣，以協助制定決策或進行預測，就可以納入資料探勘的事業版圖。

資料探勘使用了資料庫之知識發掘 (Knowledge Discovery in Database; KDD)、資料視覺化(Data Visualization)、線上及時分析 (OLAP)、統計分析 (statistics analysis)、類神經網路 (Artificial Neural Network)、機器學習 (Machine Learning)、模糊系統 (Fuzzy Systems)...等技術，在資料量龐大的資料庫中進行分析探索，找尋隱藏在資料中的訊息 (Hand, Mannila, & Smyth, 2001)，例如：趨勢 (trend)、特徵模型 (feature pattern) 或相關性規則...等，並且以人類能理解的形式呈現所尋獲的資訊或知識。

一般而言，為了達成資料探勘的目的，資料探勘在不同的領域上有許多不同的應用功能，包含下列六項功能：分類 (Classification)、推估 (Estimation)、預測 (Prediction)、關聯法則 (Association)、叢集群組 (Clustering) 及序列型樣 (Sequence Pattern) ...等 (Berry,

Michael, & Gordon, 1997)。

本研究針對受測者在網站中瀏覽的行為狀況進行探勘，找出瀏覽路徑型樣(path traversal pattern)。瀏覽路徑型樣探勘主要是針對使用者瀏覽網頁的順序進行探勘，因此適合運用序列型樣的探勘方法，因為序列型樣探勘主要在尋找資料項目之間的發生順序，亦即利用項目之間的先後順序，探勘出資料庫中所隱含的資訊。(曾憲雄、蔡秀滿、蘇東興、曾秋蓉、王慶堯，2005)。

有關序列型樣的研究由 Agrawal & Srikant(1995)首度提出，發展至今，序列型樣探勘的演算法大致可分為兩大類研究方向，第一類是 Apriori-Like 的作法，它沿用關聯規則分析的 Apriori 演算法，加入時間維度的考量，使具有前後順序關係，常見的演算法有 Full-Scan(FS)、Selective-Scan(SS)(Chen, Park, & Yu, 1996)；第二類是 Suffix-Tree Like 的作法，將所有使用者瀏覽序列建立成樹(tree)或圖(graph)的形式，再進行序列探勘，這樣的作法是希望藉此減少對資料庫的存取，以加快執行效率，常見的有 WUM(Spiliopoulou, & Faulstich, 1998)、WAP-Tree(Han, Pei, & Yin, 2000)。Apriori 演算法其主要的基本概念，如下所示：

1. 任何一個高頻項目集的”子集合”必定是個高頻項目集。這是指若{B,C,E}是高頻3項目集，則{B, C}、{B, E}、{C, E}是高頻2項目集，{B}、{C}、{E}是高頻1項目集。

2. 一個項目集若存在一個子集合不是高頻項目集，則此項目集不可能成為高頻項目集。例如：3項目集{A, C, D}中，{A, C}和{A, D}為高頻2項目集，但是{C, D}不是高頻2項目集，則{A, C, D}不可能為高頻3-項目集。

經由 Apriori 演算法，可以知道在挖掘關聯式法則時，其主要共有五個探勘的階段(Phases)，如下述：

步驟 1. 排序階段(Sort Phase)：以顧客代號(Customer ID)為主要鍵(Major Key)，交易時間(Transaction Time)為次要鍵(Minor Key)作排序。

步驟 2. 最大產品項目集合階段(Large Itemset Phase)：找出所有滿足最小支持度的產品項目

集合。

步驟 3. 轉換階段(Transform Phase)：將上一階段找到的最大產品項目集合轉換成連續的正整數，以形成轉換資料庫(Transform Database)。

步驟 4. 序列階段(Sequence Phase)：找出所有的最大序列。此階段，Agrawal 與 Srikant 共提出了 AprioriAll、AprioriSome 及 DynamicSome 三個演算法來達成。

步驟 5. 最大階段(Maximal Phase)：找出所有最大的序列，形成循序樣式。

顧客購買商品紀錄如表1所示，而圖1說明Apriori關聯規則演算法，使大家了解Apriori演算法的整個運作過程。

表1：顧客購買商品的紀錄表

顧客 ID	購買日期	購買商品	顧客 ID	購買日期	購買商品
1002	11/2	ABCD	1009	10/29	AB
1006	10/31	AEBCD	1011	10/31	BDE
1007	11/3	AC	1011	11/1	ACD
1008	11/4	ACD	1013	11/2	CDE
1008	11/5	BCDE	1016	11/3	ACE

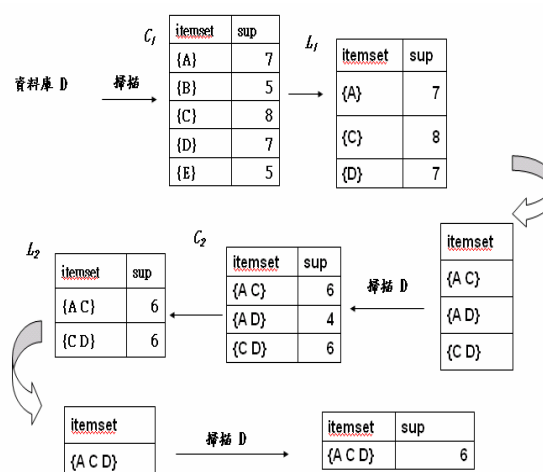


圖 1：Apriori 演算法執行過程示意圖

2.3 學習歷程的實徵研究

「學習歷程檔案」(learning portfolio) 最主要的目的或用途是用來收集學習者在學習過程中各面向的資料,以反應或顯現學習者的真實表現或改變的情況(童宜慧、張基成, 1996)。由於學生需參與有關內容、評判標準... 等的選擇與決定,並記錄自我反省的證據,因此學習歷程檔案不但可用來展現學習者在某領域學習的努力、進步與成就情形,亦可提供學習者自我反思其認知過程或學習策略運用的實際情況(林孟鴻, 2001)。以下簡述三篇學習歷程實徵研究的論文

首先,倪家祥(2001)指出教師可以透過學習系統,隨著時間變化完整地記錄學生的學習歷程(Portfolio)與時序上的考量作為觀察與預測學生成就的依據,並建構具時間性的學習者模式。第一個模組是利用 Dual model 架構,將學生在網路學習環境中具時間性的學習行為歷程及學習成果等學習變數以 Decision Tree 的形式表示而成的學習成就預測模型;第二個模組是老師平常所執行的教學策略和學生在網站上的學習行為歷程,以序列探勘(Sequential Pattern Mining)方法分析教學策略與學生學習行為改變的影響,以做為老師輔助學生依據及教學策略擬定的參考。

其次,林秀如(2004)的論文則利用資料探勘(Data Mining)的關聯式法則(Association Rules)方法,來發掘學習者常犯錯的英語介係詞概念彼此之間的關聯性,預測可能發生的錯誤概念。來對學習者進行預防教學,以減少相關概念學習錯誤的發生。在實驗流程分成兩大階段,第一階段從作答記錄資料庫中挖掘無難度控制組和有難度控制組兩組關聯法則,作為後續預防教學法則啟動規則的基礎。第二階段為預防教學,以中華大學資訊管理系105位學生來進行實驗,提供的教材版本和法則啟動規則為操弄因素,將所有學生隨機分為三組,在不同的操弄因素下進行實驗;分別收集各組學生在學習前的前測評量資料,依照前測作答資料以判斷是否要啟動法則來給予預防教材或一般教材,再收集各組學生在瀏覽教材學習後的後測學習評估資料,將各組的後測學習評量資料利用統計假說進行檢定。

最後,黃秋燕(2006)在探討符合SCORM 2004 SN規範教材之學習歷程分析研究—以資料探勘為工具,其研究目的有二:(1)取得符

合SCORM 2004 SN規範之有效學習歷程資料;(2)找出高學習成效學習者的學習模式作為學習典範。在教材設計方面以高職「計算機概論」課程中的「數值系統轉換」單元為例,以高雄縣私立高苑工商職業學校一年級103個學生為實驗對象,透過Aggregation Tree以挖掘瀏覽路徑型樣,發掘出較具意義的學習歷程路徑型樣,將學習成效良好的受測者之學習序列型樣推薦給學習成效不佳的學習者參考,也可提供給尚未進行的受測者,或教師在進行教學策略規劃時參考。以學習成效良好之學習者的學習歷程資料作為學習路徑型樣探勘的資料來源,進行資料前處理,取出有效學習紀錄,再尋找出有效學習路徑型樣。

3.研究方法

3.1 研究架構

本研究架構是設計出一套「布林邏輯查詢」的網路學習系統,並且符合 SCORM 2004 SN 規則的數位教材,使學習者透過 SCORM 2004 RTE 實際地瀏覽課程教材,以取得有效學習歷程檔案,並導入了資料探勘的方法,對學習者在線上學習的情況進行瀏覽路徑序列探勘,發現最佳學習路徑型樣。

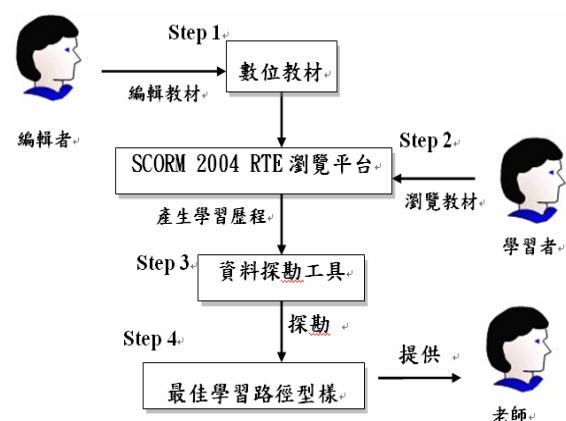


圖 2：研究架構圖

本研究架構如圖 2 所示,共有 4 步驟,首先編輯者完整編輯全部的數位教材。接著學習者透過網路瀏覽器閱讀 SCORM 教材,遠端的線上學習系統透過 SCORM 2004 RTE,把學習者的課程教材之學習歷程能記錄下來,當作是資料探勘的樣本。接著利用 SQL Server 2005 關聯規則進行分析學習歷程的資料探勘,找出一個最適當的學習路徑型樣。最後,把最適當的學習路徑型樣提供建議給教師在教材內容編排序順及教學活動設計之參考。最後,本研究亦針對未來研究提出建議。

3.2 實驗設計

本實驗以「大學基礎的計算機概論」課程中的「布林邏輯查詢」課程之教材為。教材內容分為 5 個子單元，包括有：「布林邏輯查詢摘要」、「And 的數學教材內容」、「OR 的數學教材內容」、「Not 的數學教材內容」、「XOR 的數學教材內容」，讓學習者可依不同順序在 SCORM 2004 RTE 上自由瀏覽，透過 RTE 記錄學習者之學習歷程檔案；最後，再施以課程學習成就測驗，以了解學習者之學習成效。

本研究為取得有效學習之學習歷程紀錄，單元學習活動包含該子單元之學習內容瀏覽。

教材設計以 SCORM 2004 RTE 為實驗平台，學生依其帳號及密碼登入後，進入課程畫面，未點選進入任何單元教材前，如圖 3 顯示出課程單元架構的樹狀結構物件，允許學習者可不按順序任意點選學習單元之教材節點；而點選任何單元教材開始瀏覽後，就可以自由瀏覽教材，其測驗畫面如圖 4 所示，只要利用『Continue →』循序控制按鈕，就可換到下一題測驗，直到完成十五題單元測驗試題後，結束此次受測。



圖 3：進入課程畫面



圖 4：單元測驗畫面

3.3 實驗對象

在「布林邏輯查詢」課程之教材主要教學對象是以南台科技大學數位設計學院大學部學生為實驗對象，以 SCORM 2004 RTE 為實驗平台，利用大學基礎的計算機概論課程中的「布林邏輯查詢」課程之教材為例，提供給教師在教材內容編排、教學策略制定及教學活動設計之參考。因此，本研究之研究樣本為資訊傳播系的大學部學生 22 人，扣除最佳分數和最差分數的受測者，總共為 20 位受測者。

3.4 實驗工具

本研究所採用之研究工具共有 4 大類，分別是：SCORM 2004 RTE、學習成就測驗、學習路徑探勘工具，茲分別說明如下：

(一) SCORM 2004 RTE

SCORM 2004 執行環境（Real Time Environment；RTE），在過去的網路教學中，教學平台與教材內容間的溝通方式不同，不僅所使用的程式不同，各個應用程式介面（Application Program Interface；API）的定義不相同之問題，SCORM 提出 Run-Time Environment 以 JavaScript 溝通教學平台與教材，使得教學平台以及教材之間有了共同溝通的介面，以達到網路教學共享的目的。

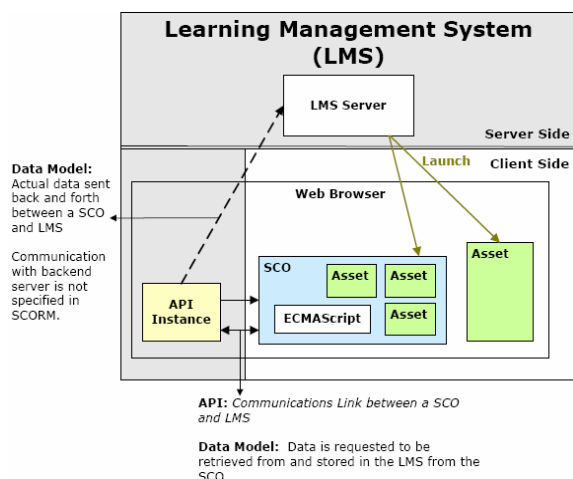


圖 5：RTE 架構概念圖（ADL, 2004）

SCORM 2004 RTE 標準的執行環境如圖 5 所示，圖中的上部份為伺服器端，即學習管理系統。LMS Server 為提供使用者所需之教材，當 Client 下載教材之後，其 API Instance 會隨著教材一起被下載至 Client。API Instance 便是居間作為學生與教學平台溝通的橋樑，教材透過 API，要求 API Instance 對教學平台發出對應的指令，以用來掌握管理學習者的學習進度與狀況，例如將學習者本次的考試成績傳給 LMS Server、或根據學習者之評量表現取得下一個適當教材等。在執行環境運作期間，學習者將渾然不覺其背後的指令運作，僅需專注於自己的學習即可。

SCORM 2004 RTE 為 ADL 所制定的標準化線上教材播放系統，教材需符合 SCORM 2004 之規範，才可在 RTE 上播放。而 RTE 提供 SampleRTE.mdb 及 GlobalObjectives.mdb 兩個資料庫，用以記錄匯入的課程、課程單元資料、使用者帳號和密碼以及課程的瀏覽情況，但並不足以提供本研究進行資料探勘所需之資料欄位。因此，經本研究修改 RTE 中的 LMSCMIServlet.java 及 sequencingEngine.jsp 兩個應用程式介面，在 RTE 的後端資料庫 SampleRTE.mdb 中新增 Portfolio 資料表，以記錄資料探勘所需的學習歷程資料欄位。而在安裝 SCORM 2004 RTE 套件後，會自動安裝 Tomcat Server，即個人小型網路作業系統；在啟動 Tomcat Server 後，即可啟用 SCORM 2004 RTE。

SCORM 2004 RTE 可依所登入之帳號及設定之權限，給予不同的功能，主要分為『學習者』與『管理者』二個角色，學習者有「登錄我的課程」、「瀏覽我的課程」、「變更密碼」與「登出」四項功能。而管理者所具備之功能，主要包含「匯入課程」、「新增使用者」、「清空所有資料庫」、「刪除使用者」四項功能。

(二) 學習成就測驗

學習成就測驗主要目的在評量使用者對「布林

邏輯查詢」課程之教材內各個單元的理解程度。協助學習者如何在網路上龐大的資訊量中，利用「網路查詢」的學習模式，依個人對資料需求，快速且精準的找尋到自己所要的資料，進而縮短網路查詢資料的時間。本研究之課程測驗共有十五題試題，乃是「布林邏輯查詢」課程之教材內容，而測驗分數可以做為學生在瀏覽教材後的學習成效評估。

(三) 學習路徑探勘工具

本研究採用 SQL Server 2005 關聯規則的探勘方法來進行學習歷程資料的探勘分析，透過已完成資料前處理的學習歷程檔案，轉換出每位學習者進行教材瀏覽的學習路徑型樣。

3.5 實驗實施

在進行實驗的過程，主要分為「準備階段」及「實施階段」，分別說明如下：

(一)準備階段：本階段主要在建置 RTE，安裝 SCORM RTE 1.3.3 與 J2SDK1.4.2 09，之後匯入課程，並為每位受測者建立其帳號及密碼，即完成準備階段。

(二)實施階段：先說明如何登入 RTE、如何瀏覽教材及何時實施課程後測，接下來受測者使用個人帳號及密碼登入教材播放環境，並熟悉操作方式後，正式進行課程之教材瀏覽，以評估其學習成效。

3.6 資料收集

修改後的RTE其後端資料庫所收集的學習歷程資料表Portfolio，包含的資料欄位如表2所示：

表2：Portfolio資料表欄位說明表

欄位名稱	資料類型	說明
識別碼	自動編號	資料紀錄編號
UserID	數字	學習者帳號
ScoID	文字	單元名稱
Numattempts	數字	單元教材閱覽次數
UserName	文字	學習者姓名
DateTime	日期／時間	單元教材開始閱覽日期及時間

4.研究結果與討論

為了發掘出較具意義的學習歷程路徑型樣，本研究配合課程學習成就測驗進行較佳的學習序列型樣的挖掘，將學習成效良好的受測者之學習序列型樣推薦給學習成效不佳的學習者參考，也可提供給尚未進行的受測者，或教師在進行教學策略規劃時參考。本研究利用學習成效良好之學習者的學習歷程資料作為學習路徑型樣探勘的資料來源，進行資料前處理，取出有效學習紀錄，再透過SQL Server 2005關聯規則當作資料探勘的探勘技術，對有效的學習紀錄進行學習歷程探勘，分析尋找出最有效的學習路徑。其分析方法分為「學習紀錄資料前處理」及「學習路徑型樣探勘」，作法分別說明

如下：

4.1 學習紀錄資料前處理

學習紀錄資料前處理包含「紀錄有效學習歷程」及「轉換為學習者瀏覽序列」兩個項工作說明如下：

(一)紀錄有效學習歷程

在瀏覽教材規則上，當學習者進入閱覽單元教材後，可以自由瀏覽所有單元教材內容，才能再選擇其他子單元教材閱覽，因此，透過SCORM 2004 RET可於後端資料庫的Portfolio資料表中，記錄學習者已閱覽了哪些單元及每個單元的後測結果。

Portfolio 資料表以UserID及DateTime兩個欄位分別為主要鍵及次要鍵排序後，可將同一學習者之學習紀錄排在同一段落，並以時間先後順序排列。如圖6所示：

圖 Portfolio：資料表						
識別碼	USERID	SCOID	NUMATTEMPTS	USERNAME	DATEIME	
436	49549	general203	1	38597;22914;3	2006/12/5 下午 02:43:12	
437	49549	general300	1	38597;22914;3	2006/12/5 下午 02:43:49	
438	49549	general301	1	38597;22914;3	2006/12/5 下午 02:43:59	
439	49549	general302	1	38597;22914;3	2006/12/5 下午 02:44:39	
440	49549	general303	1	38597;22914;3	2006/12/5 下午 02:45:04	
441	49549	general400	1	38597;22914;3	2006/12/5 下午 02:45:33	
442	49549	general401	1	38597;22914;3	2006/12/5 下午 02:45:42	
443	49549	general402	1	38597;22914;3	2006/12/5 下午 02:46:05	
444	49549	general403	2	38597;22914;3	2006/12/5 下午 02:46:26	
445	49549	general500	2	38597;22914;3	2006/12/5 下午 02:47:10	
446	49549	general501	2	38597;22914;3	2006/12/5 下午 02:47:28	
447	49549	finish	1	38597;22914;3	2006/12/5 下午 03:04:48	
448	49549		0	38597;22914;3	2006/12/5 下午 03:08:32	
449	49571	LogCoul00	2	24605;33841;2	2006/12/5 下午 03:10:30	
450	49571	LogCoul01	2	24605;33841;2	2006/12/5 下午 03:11:30	
451	49571	general200	1	24605;33841;2	2006/12/5 下午 03:11:40	
452	49571	general201	1	24605;33841;2	2006/12/5 下午 03:12:08	
453	49571	general202	1	24605;33841;2	2006/12/5 下午 03:12:24	
454	49571	general203	2	24605;33841;2	2006/12/5 下午 03:12:42	
455	49571	general300	1	24605;33841;2	2006/12/5 下午 03:13:09	
456	49571	general301	2	24605;33841;2	2006/12/5 下午 03:13:19	
457	49571	general302	1	24605;33841;2	2006/12/5 下午 03:13:28	

圖6：Portfolio資料表中有效學習節點紀錄

(二) 轉換為學習者瀏覽序列(User Session)

為了快速分析出所有紀錄的學習者瀏覽序列，所以在SCOID欄位中的教材網頁節點名稱，必須轉換為較簡短的字元代碼，因此建立一個對照表如表3所示，以縮短字串長度，加快比對速度。接下來將相同USERID的SCOID依時間前後順序排列出學習者的瀏覽路徑序列，轉換出每位學習者在課程當中的有效瀏覽路徑序列，稱為User Session，如圖7所示。

表3：課程單元名稱與代碼對照表

SCOID	ScName	ScContent	SCOID	ScName	ScContent
1	chapter1_G100	摘要	3.3	chapter1_G302	OR 教材內容-3
2.1	chapter1_G200	And 教材內容-1	4.1	chapter1_G400	Not 教材內容-1
2.2	chapter1_G201	And 教材內容-2	4.2	chapter1_G401	Not 教材內容-2
2.3	chapter1_G202	And 教材內容-3	4.3	chapter1_G402	Not 教材內容-3
2.4	chapter1_G203	And 教材內容-3	4.4	chapter1_G403	Not 教材內容-4
3.1	chapter1_G300	OR 教材內容-1	5.1	chapter1_G500	XOR 教材內容-1
3.2	chapter1_G301	OR 教材內容-2	5.2	chapter1_G501	XOR 教材內容-2

4	P20705	79	1,2,2,2,2,2,3,3,3,3,3,4,1,4,2,4,4,4,3,5,5,5,5,2
5	49518	100	1,2,3,2,4,3,3,4,4,3,4,5,1,5,2
6	49571	65	1,2,1,2,2,2,2,3,3,3,3,3,4,1,4,2,4,4,5,1,5,2
7	50105	72	1,2,1,2,2,2,2,3,3,3,3,3,4,1,4,2,4,4,4,3,5,1,5,2
8	4050145	87	1,2,1,2,2,2,2,3,3,3,3,3,4,1,4,2,4,4,4,5,1,5,2
9	50103	93	1,2,1,2,2,2,2,3,3,3,3,3,4,1,4,2,4,4,4,3,5,1,5,2
10	P90805	52	1,2,2,2,1,2,2,2,3,3,3,3,3,4,1,4,2,4,4,4,5,1,5,2

圖7：教材網頁節點瀏覽路徑序列

4.2 學習路徑型樣探勘

利用了學習成效良好之學習者之學習歷程資料，經過有效學習歷程紀錄的篩選，並且轉換為學習者瀏覽路徑序列之後，透過資料庫軟體SQL Server 2005中關聯規則當作資料探勘的技術，並對學習歷程紀錄進行分析，以下將使用SQL Server 2005內建資料庫設計關聯規則模型，如圖8、9、10所示：

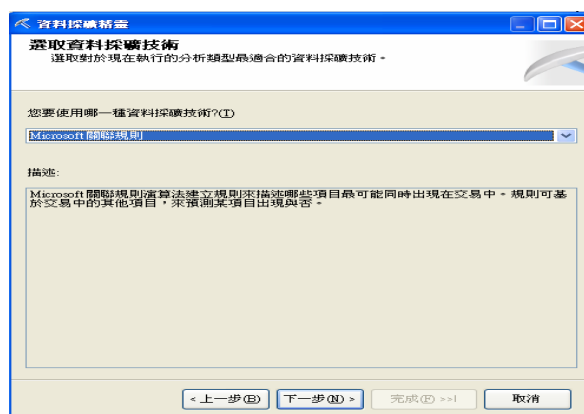


圖8：選取關聯規則演算法



圖9：設定案例資料表與巢狀資料表



圖 10：設定索引鍵、輸入與可預測變數

SQL Server 2005 關聯規則提供之參數如表 4 所示：

表 4：SQL Server 2005 關聯規則參數表

MINIMUM_ITEMSET_SIZE	一個物件組最少可包含之物件數量。
MAXIMUM_SUPPORT	最大支援門檻。物件組如果支援高於此參數將會被刪除。如果此參數為大於 1 之整數，代表絕對數量；如果是小於 1 之小數時，則代表佔總體的百分比。
MINIMUM_SUPPORT	最小支援門檻。物件組如果支援低於此參數將會被刪除。如果此參數為大於 1 之整數，代表絕對數量；如果是小於 1 之小數時，則代表佔總體的百分比。
MINIMUM_PROBABILITY	規則最小信心水準門檻。必須填入界於 1 到 0 的小數，信心水準低於此值的規則將會被刪除。
MINIMUM_IMPORTANCE	規則最小重要性門檻。當規則重要性低於此值時則會被刪除。

如果分析的資料是巨大的資料，以至於規則之支援會極小，而找不出符合最小支援們見之規則，因此，在建立預測模型時，可以先將最小支援門檻設小，以便演算法找出大多數的規則，接下來再透過最小支援門檻或是規則最小信心水準門檻逐步調高，以排除不重要之規則。

利用了學習成效良好之學習者之學習歷程資料，經過有效學習歷程紀錄的篩選，並且轉換為學習者瀏覽路徑序列之後，透過 SQL Server 2005 的關聯規則當作資料探勘的探勘技術對學習歷程紀錄進行分析，產生二個值得推薦之學習路徑型樣分別是：

1. [1]→ [2.1]→ [2.2]→ [2.3]→ [2.4]→ [3.1]→ [3.2]→ [3.3]→ [3.4]→ [4.1]→ [4.2]→ [4.3]→ [4.4]→ [5.1]→ [5.2]。
2. [1]→ [2.1]→ [2.2]→ [2.4]→ [2.3]→ [3.1]→ [3.2]→ [3.4]→ [3.3]→ [4.1]→ [4.2]→ [4.4]→ [4.3]→ [5.1]→ [5.2]。

5.結論與建議

本研究開發課程教材並符合 SCORM 2004 順序瀏覽規範 (Sequencing and Navigation; SN) 的布林邏輯查詢教材。受測者在整個課程的學習過程中，可依不同順序自由選擇其所欲學習的單元，不同的學習順序，可能形成不同的學習成效，為判斷學習者在其所選擇的學習單元中是否真正進行了有效的學習，用以判斷學習者是否於該單元進行了有效學習。以達到在「數位學習」最終目的，就是提供了不受時間與空間限制的另類學習機制，造成學習者可以有更多的學習時間與機會學習知識。

參考文獻

- [1]李昇暉、林居鴻、馮志成、游合正、陳怡妃(2003)。符合 SCORM 標準的多媒體串流教材模型。網路教學系統平台與內容標準化研討會(WISCS 2003) 第 12 卷，專刊期 103-109。
- [2]林秀如(2004)。以資料探勘發掘學習錯誤概念之關聯的預防教學-以英語介係詞為例。新竹縣：中華大學資訊管理研究所碩士論文。
- [3]林孟鴻(2001)。應用歷程檔案建立網路學習輔助機制。高雄市：國立中山大學資訊管理研究所碩士論文。
- [4]林錦泓(2002)。透過多代理人實作 SCORM 為基礎的教案推薦系統。高雄市：國立高雄師範大學資訊教育研究所碩士論文。
- [5]倪家祥(2001)。以網站行為的歷程建立具時間性學習者模式。桃園縣：國立中央大學資訊工程研究所碩士論文。
- [6]莊宛毓(2005)。符合 SCORM 2004 SN 之適性化教材對學習成效之研究。高雄市：國立高雄師範大學資訊教育研究所碩士論文。
- [7]曾憲雄、蔡秀滿、蘇東興、曾秋蓉、王慶堯(2005)。資料探勘 Data Mining。旗標出版股份有限公司。
- [8]童宜慧、張基成(1996)。網路化學習歷程檔案系統之建構與評鑑——一個電子化的

真實性學習評量工具。遠距教學。第 13/14 期合刊，78-90。

[9]黃秋燕(2006)。符合 SCORM 2004 SN 規範之學習歷程分析研究—以資料探勘為工具。高雄市：國立高雄師範大學資訊教育研究所碩士論文。

[10]ADL (2004). *SCORM 2nd Edition Overview*. (July 22, 2004). Retrieved March 20, 2006, from the World Wide Web: <http://www.adlnet.org/downloads/>

[11]ADL (2004). *SCORM Content Aggregation Model*. Version 1.3.1. (July 22, 2004). Retrieved March 20, 2006, from the World Wide Web: <http://www.adlnet.org/downloads/>

[12]ADL (2004). *SCORM Run-Time Environment*. Version 1.3.1. (July 22, 2004). Retrieved March 20, 2006, from the World Wide Web: <http://www.adlnet.org/downloads/>

[13]ADL (2004). *SCORM Sequencing and Navigation*. Version 1.3.1. (July 22, 2004). Retrieved March 20, 2006, from the World Wide Web: <http://www.adlnet.org/downloads/>

[14]Agrawal R. and Srikant, R. (1995). *Mining Sequential Patterns*. Proceedings of Int'l Conference on Data Engineering(ICDE), Taipei, Taiwan, 3-14.

[15]Anderson, T., & Elloumi F.(2003).*Theory and Practice of Online Learning*. Athabasca University, from the World Wide Web: http://cde.athabasca.ca/online_book/ch1.html

[16]Anderson T., & Elloumi F. (2003).*Theory and Practice of Online Learning*. Athabasca University, from the World Wide Web: http://cde.athabasca.ca/online_book/ch1.html

[17]Berry, Michael J.A., & Gordon L. (1997). *Data Mining Techniques For Marketing, Sales, and Customer Support*. John Wiley & Sons, Inc., New York, NY.

[18]Chen M.S., Park J.S., & Yu P.S. (1996).

Data Mining for Path Traversal Patterns in a Web Environment. Proceedings of the 16th ICDCS, 385-392.

[19]Cooley R., Mobasher B., & Srivastava J. (1997). *Web Mining: Information and Pattern Discovery on the World Wide Web*. Proceedings of 9th IEEE Int'l Conference on Tools with Artificial Intelligence, 558-567.

[20]Hand D., Mannila H., & Smyth P. (2001). *Principles of Data Mining*. Massachusetts Institute of Technology.

[21]Han J., Pei J., & Yin Y. (2000). *Mining Frequent Patterns without Candidate Generation*. Proceedings of AMC SIGMOD, Int'l Conference Management of Data, 1-12.

[22]Spiliopoulou M., & Faulstich L.C. (1998). *WUM: A Tool for Web Utilization Analysis*. Proceedings of EDBT Workshop WebDB, 184-193.

附錄一：「布林邏輯查詢」學習成就測驗試題

01.在布林代數運算中，若 $A=1, B=1$ ，則 $A'+B=$ (A) 0 (B) 1

02.在布林代數運算中，若 $A=1, B=1$ ，則 $A \text{ and } B' \text{ xor } A =$ (A) 0 (B) 1

03.在布林代數運算中，若 $A=1, B=1$ ，則 $\text{NOT } A \text{ OR } B \text{ AND } A \text{ AND NOT } B =$ (A) 0 (B) 1

04.在布林代數運算中，若 $A=0, B=1$ ，則 $A \text{ OR NOT } B \text{ AND NOT } A \text{ OR } B =$ (A) 0 (B) 1

05.計算 $(1 \text{ XOR } 1) \text{ AND } (1 \text{ XOR } 0) \text{ XOR } (1 \text{ XOR } 1)$ 的結果為：(A) 0 (B) 1

06. 若 $A=1100, B=0110$ ，則 $A' \text{ XOR } B=?$ (A) 0110 (B) 1001 (C) 0101 (D) 1100

07.若 $A=1010, B=1110$ ，則 $A \text{ and } B'=?$ (A) 0110 (B) 1001 (C) 1010 (D) 1100

08.下列陳述何者為正確？(A)若 A,B 兩事件， $P(A \cap B) = P(A) + P(B) - P(A \cap B)$

(B)若 A,B 兩事件， $P(A \cup B) =$

$P(A) + P(B) - P(A \cup B)$

(C)若 A,B 兩事件， $P(A \text{ and } B) =$

$$P(A)+P(B)-P(A \cup B)$$

(D)若 A,B 兩事件， $P(A \text{ or } B) =$

$$P(A)+P(B)-P(A \cup B)$$

09.下列陳述何者為正確？(A)若 A,B 兩事件，

$$P(A \cap B) = P(A)+P(B)-P(A \cup B)$$

(B)若 A,B 兩事件， $P(A \cup B) =$

$$P(A)+P(B)-P(A \cup B)$$

(C)若 A,B 兩事件，則 $P(A \text{ or } B) = P(A)+P(B)+P(A \cap B)$

(D)若 A,B 兩事件，則 $P(A \text{ and } B) =$

$$P(A)+P(B)+P(A \cup B)$$

10.布林代數之邏輯運算中，下列何者有誤？

(A) $1+1=1$ (B) $0+1=0$ (C) $1*1=1$ (D) $1*0=0$

(E) $0 \oplus 1=0$ (F) $1 \oplus 1=1$

11.執行布林運算，下列何者為真？

(A) $1*1=0$ (B) $0*1=1$ (C) $1+1=0$ (D) $0+1=1$

(E) $0 \oplus 0=1$ (F) $1 \oplus 0=0$

12.邏輯閘 AND 的二個輸入端 X,Y 均為 1 時，則輸出端的值為：(A) 1 (B) 0 (C) 2 (D) 以上皆非。

13.若兩個輸入端輸入的值不同，則輸出為 1 的邏輯閘是：(A) NAND (B) NOR (C) XOR (D) OR。

14. 若邏輯閘 XOR 兩輸入端分別送入兩列信號 10000101 與 01000111，試問 XOR 的輸出結果為何？ (A) 11000111 (B) 00000101 (C) 11000010 (D) 00111101。

15.下列何者為 AND 之真值表：

A.

X	Y	F
0	0	0
0	1	1
1	0	1
1	1	0

B.

X	Y	F
0	0	0
0	1	0
1	0	0
1	1	1

C.

X	Y	F
0	0	0
0	1	1
1	0	1
1	1	1

附錄二：課程學習成就測驗成績

User ID(帳號)	Score(分數)	User ID(帳號)	Score(分數)
49549	73	50604	87
F80705	51	9102	74
50700	80	4050182	38
F20705	79	F30705	86
49518	100	4050121	81
49571	65	49553	59
50105	72	50506	86
4050145	87	4051102	58
4050195	72	F70405	87
50103	93	49552	79
F90805	52	9182	93