

以微型教材與資料探勘為基礎的數位學習推薦系統 與學習成效探討

黃聖博

南臺科技大學資訊管理系
ma490106@stust.edu.tw

鄭鈺霖

南臺科技大學資訊管理系
jackjeng@stust.edu.tw

摘要

每位學生的學習能力不相同，因此在學習上成效也會有所不同。本研究建置一套數位學習推薦平台，主要推薦的資訊有三種，一種是與教材相關影片，另一種是學習路徑，最後是前三名學生的學習路徑及學術文章搜尋而來的補充教材，其中學習路徑主要是透過系統收集資訊，也就是閱讀了哪些教材，教材又可以細分為完整版教材及精簡版，在收集到一定的資料後，透過資料探勘找出適合的學習路徑推薦，進而實現個人化的輔助學習。

本研究的結果顯示，推薦系統能夠提高學習的成效，像是閱讀時間逐漸減少，然而為了有更明顯的數據，以測驗成績來評估學習成效，其中測驗的部份也以多份試卷隨機抽取，因此也避免背答案的可能，以結果而言，成績有了明顯的進步。

關鍵詞：數位推薦系統、資料探勘、學習路徑

1. 前言

過去在學習新的知識只能到學校、補習班，上課的方式都是老師在講台上教課，黑板上寫著許多課程的重點，底下的學生則是忙的把黑板上的東西全部的抄起來，回到家在慢慢的消化上課的內容，但等到回到家後，對上課的內容印象就模糊的了許多，也因此又需要花費更多的時間在學習上。

現今各大院校都有專屬的數位學習平台，數位化也使得學習上更加的多樣化，以數學解題來說，不在局限課本上的文字說明及公式，透過圖片、影片的方式，讓學生比較了解，這也成為老師傳遞資訊的地方，老師只需將教材放置平台上，所有的學生都可以閱讀，讓學生在學習上也有更多的選擇。儘管目前學生已經有許多數位平台可以選擇，仍有一些問題還是沒有改善，以平台的角色來說主要是用來輔助學習，特別是專業科目上學習更加的困難，

每個學生程度不一樣，學習速度更不會相同，也因此”個人化”的輔助學習就顯得相當重要。數位學習推薦系統在個人化方面提供很多幫助，藉由數位學習平台收集每個學生的學習歷程，例如：閱讀教材的時間長短、閱讀喜好（影片教學或文字說明），並透過資料探勘演算法找出每一位學生適合的數位教材，並由系統主動推薦給學生。

為了提供更多個人化的輔助，並且能讓學生利用較有效率的方式獲得學習概念，因此本研究提出的個人化協助的學習平台，適時的提供學生幫助，讓學習平台不在只是單純的繳交作業看看上課內容，讓每位學生能夠獲得專屬於自己的學習方法。

2. 文獻探討

2.1 數位學習推薦系統

現今大學科目種類繁多，專業科目也各不相同，這也造成學生在學習上相當不容易，然而以傳統教學的方式對學生而言學習上會造成相當大的負擔，因為課本只會有文字及圖片，就微積分來說只有圖片與文字說明還是難以理解，所以教學方式勢必要改變，也因此數位學習也開始漸漸的興起。

數位學習（e-learning）為近年來相當常見的教學方式，主要是透過網路、互動電視、光碟來傳遞學習的資訊，也可以稱為遠距教學。藉由數位化的特性，改變了以往的教學方式，如下表所示，其中數位學習最重要的一點就是”個人化”的功能，也就是所謂的適性化教學，這也是目前更多學者在討論的義題，因此為了實現個人化學習本研究將結合推薦系統。

隨著技術的進步，推薦已經存在各種網頁，然而推薦技術能應用的地方相當的多，以購物網站來說，例如 Amazon、博客來就有屬於自己的推薦機制，在消費者看商品時，同時透過推薦系統推薦給消費者可能會想要購買的商品，也因此原本打算只買一樣的商品，但因

為推薦系統讓消費者看到多的商品，使消費者常常比預期的多買了好幾樣的商品，也因此銷售業績也大幅度的提高，各大網站紛紛加入屬於自己專屬的推薦系統，然而在推薦之前最重要的就是收集使用者的使用網站資訊，可分為內容過濾（Content-based filtering）與協同過濾（Collaborative Filtering）。

內容過濾（Content-based filtering）

將收集的資訊內容，找出當中的關鍵字，以文件來說關鍵字通常會出現較多次，並去除一些與關鍵字無關的，依據關鍵字出現的次數給予不同的權重，最後依照關鍵字的權重推鍵給相關的資訊推薦給使用者。雖然內容過濾的作法相當容易理解，但仍有一問題，以收集資訊來說，就會有一些限制例如：音樂、電影就沒辦法收集到當中的資訊裡的內容。

協同過濾（Collaborative Filtering）

這也是目前較常見的推薦方式，主要以記錄使用者的使用經驗，將使用者分類，並與資料進行比較，找出與你相似的使用者（族群）的喜好，透過記錄的預測你的個人喜好(Burke, 2002)，但這樣的作法仍有一些問題，以下將會列出協同過濾的優缺點。

優點：

當收集的資料越多，推薦的資訊也會更加準確提供個人化的推薦

缺點：

新的使用者加入時，預測的資料會較不準確

當使用者的偏好突然改變，預測上就會不準確，也就需要重新收集資料與分析

綜合上述，傳統教學方式在學習成效上效果有限，以專業科目而言，只透過課本內容還是難以理解，但透過數位學習就能夠提供更多元化的學習方式，因此數位學習已經是目前相當重要的教學方式。數位學習雖然改善傳統教學方式，但以個人化學習而言，仍然沒有提供很好的幫助，為了解決這個問題，本研究將以數位學習推薦系統，適時的給予一些幫助，實現個人化學習的目標。

2.2 資料探勘（Data mining）

現今是一個資訊化及網路化的時代，在生活中也是學習相關，如：手機、電腦，然而這

也造成資料過多的問題，例如：到大賣場買商品，就會有一次的消費記錄及購買的明細，或者是以 Youtube 為例，就會收集使用者的資訊，像是查詢影片的記錄、觀看影片的時間、點擊了哪些影片，然而這些的動作都可以是一筆記錄，也因此對資料庫來說，儲存的資料量是相當的龐大，這些資訊對於一般人來說並無太大的用處，所以有了“資料探勘”這項技術，資料探勘主要是用來從大量的資料中，找出其中的潛在規則及有價值的知識，也就是從資料找尋新的知識 (Knowledge-Discovery in Database) (M Moradi & M R. Keyvanpour, 2015)，這也是資料探勘的一個步驟，簡單的說就像是從一大堆的書中，找出有幫助的資訊。

關聯法則（Association Rule）

關聯法則在資料探勘中是相當重要的技術，主要是用來評估資訊之間的相關性，並從這些關聯中找出有價值的資訊，應用的範圍也相當的廣，以醫院為例，較知名的就像是血液與心肌梗塞(G L Dong, K S Ryu & M Bashir et al, 2013)，二者之間看似並無直接的關聯，但經過研究發現血液對於心肌梗塞有相當大的影響。

資料探勘演算法

Apriori 在資料探勘中，是最常使用的演算法 (Agrawal, R. & Srikant, R., 1994)，也是最具有代表性的演算法之一，也因此許多學者以 Apriori 做為基礎去延伸，如：FP-growth，以下將介紹 Apriori 的主要做法。

名詞解釋：

最小支持度（Min Support）：在開始演算法前，會先定義一個門檻值（threshold），即為最小支持度，若低於支持度的門檻則不具有參考性。
最小信賴度（MIN Confidence）：與最小支持度意思相同，低於最小信賴度的則不具有參考性。

k-項目組（k-itemset）：假設某個項目組有 k 個項目，則稱為 k-項目組，例如{冰箱，電扇、冷氣機 } 為 3-項目組

候選 k-項目組（candidate k-itemsets）：集合所有 k-項目組的集合，即為候選 k-項目組

高頻 k-項目組：在 k-項目中在每個項目中，大於最小支持度，即為高頻 k-項目組

1. 使用者自行定義最小支持度與最小信賴度

2. 透過(k-1)- 高頻項目集，產生候選項目組
3. 搜尋資料庫，並計算所有項目集合中，其出現的總次數，只要大於等於最小支持度的，即為高頻度項目組
4. 重覆 2、3 步驟，直到無法在產生新的候選項目集合

Apriori 演算法學習上不難，在使用上經過許多學者引用及證實，所以算出的結果更具有說服力，但應用上仍會有些問題，新資料加入時，就需要全部重新計算一次，也因此資料量大時，即時運算就需要較長的時間。

依據上述觀點可發現，資料探勘運用的領域相當的廣範，以數位學習而言，就能透過系統收集到許多的使用者資訊，像是學生閱讀的章節、閱讀的時間長短，並將每位的學生的資訊找出有意義的關聯，進而達到更好的學習成效。

2.3 學習路徑

依教材來說都有章節及順序性，特別是對於剛開學習新知識的學生更為重要，因為剛開始學習不可能就直接就學習到進階的東西，也因此推薦合適的學習路徑更能夠有效的提高學習的品質。(Alva Muhammad, Qingguo Zhou, Ghassan Beydoun, Dongming Xu & Jun Shen, 2016)，對於學生來說也是輔助學習的一種參考，以下將介紹學習路徑的是如何形成的。

學習路徑主要由透過數據追蹤 (data tracking) 學習者使用資料，以數位學習平台而言，閱讀教材也能夠算是一筆資料，例如閱讀哪些章節，也就是每次閱讀的畫面為一個節點，並將點與點之間以鏈結連接起來，所形成的就是學習路徑 (林紀慧, 1998)。依本研究而言，教材有完整版與精簡版，用於鏈結點與點之間的線，主要連接以外也是用來記錄學習路徑的順序性，例如閱讀第一章完整版後閱讀第二張精簡版，最後則在閱讀第三章精簡版，鏈結的順序也就是第一章完整版連接至第二章精簡版，第二章精簡版連接至第三章精簡版，這樣的順序也就是一筆完整的學習路徑。本研究將所有的可進行的學習路徑畫出來，如下圖 1 所示。

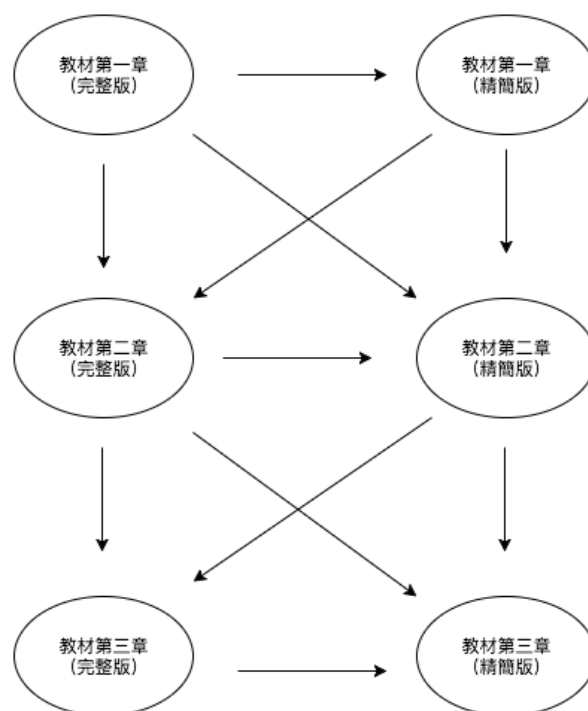


圖 1：本研究學習路徑圖

依據上述可以發現，對於學習新知識的學生而言，學習路徑就像是給予學習的方向，讓學生在學習不會感到迷惘，有效的利用學習路徑不但能減輕學生的負擔及學習時間，更能夠提升學習的成效，也因此本研究將結合數位推薦系統、資料探勘、學習路徑的特色，提供學生更有效的學習方式。

3.研究方法

教材將以微型教材為主，因為微型教材的特色就是短時間、較小單位的學習素材，在學習上也不容易造成學生過多的負擔。教材可在分為完整版與精簡版，一種是教師上傳後的完整教材，在研究中稱為完整版，另一種教材是將完整版教材經由摘要系統整理出的教材內容，稱為精簡版。

本研究透過收集學生使用情況，分為不同階段的推薦，收集的資料有閱讀的時間、閱讀的章節、閱讀微型教材內容 (完整版與精簡版)、測驗成績，然而推薦的內容可分為影片、學習路徑、學術文章，由於學術文章通常內容都相當的多，也因此可透過系統內的摘要系統簡化文章，進而縮短學習時間，本章將會先介紹系統流程，之後會在詳細介紹推薦的模組。

在一開始學生新註冊閱讀上課教材時，由於收集的資訊不足，也因此本研究將收集的時限定為一週，然而在一開始閱讀推薦方式以 20 分鐘為推薦條件，超過 20 分則會找尋相關影片進行推薦，直到閱讀時間小於 20 分鐘。一週後收集資訊也比較足夠，透過 Apriori 演算法找出每一位專屬的學習路徑並推薦給學生，最後以測驗來進行評估學生是否有學習到，以是否有及格為標準，若不及格學生則會推薦二種不同的資訊，以前三名學生的學習路徑提攻給不及格的學生做為參考，及透過學術搜尋的相關文章，流程如下圖 2 所示。

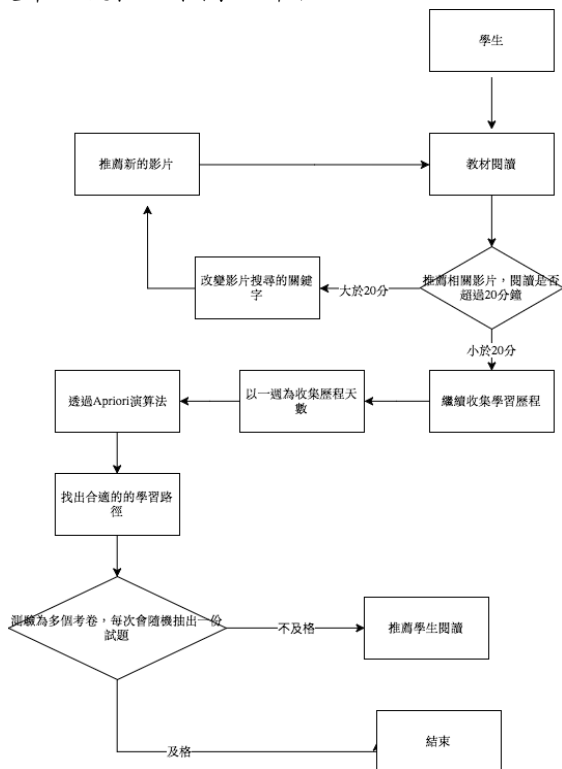


圖 2：系統流程圖

3.1 影片推薦

學生在閱讀微型教材可分為完整版與精簡版，閱讀時間則以 5 至 25 分鐘較為合適 (Allen, 1967)，因此本研究將閱讀時間超過 20 分定為太長，系統將會透過教材關鍵字去找尋相關的影片並推薦給學生，如下圖 3、4 所示。

推薦影片列表

標題
Easy Java: Check Box, Radio Buttons Part 1 - YouTube
04 在RadioGroup上建立事件Android APP開發教學吳老師3 - YouTube
AutoCAD How to Create Layouts - New Layout Tutorial - YouTube
05_如何將按鈕變圖檔的版型設計(APP教學吳老師提供) - YouTube
(15) Using pop up messages. Toast versus AlertDialog

12

圖 3：推薦影片列表

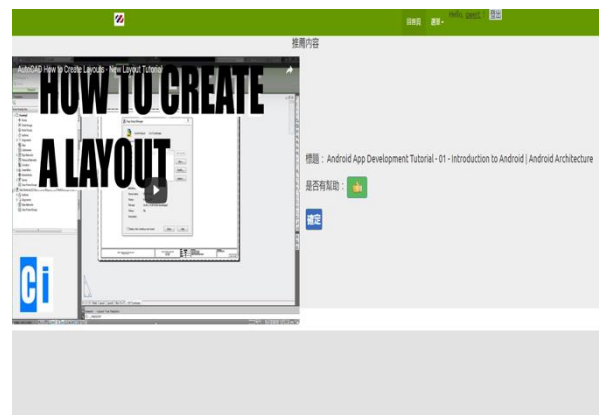


圖 4：推薦影片

3.2 學習路徑/學術文章推薦

另一方面學生看過影片閱讀時間仍然沒有減少，則更換不同的教材關鍵字，在平台收集資訊一週後，將會透過 Apriori 演算法找出關聯，推薦學習路徑主要可分為二個階段，每一章的教材定義：第一章教材（完整版）為 A、教材（精簡版）為 B、第二章教材（完整版）為 C、教材（精簡版）為 D...以此類推，算出來的結果會推薦給每一位學生，也因此每一位都有專屬於自己個人化的學習路徑，最後以測驗的方式評估學生是否有順利的獲取新知識，然而為了避免學生背答案的現象，測驗則以多份考卷中隨機抽取一份進行測驗，若成績不及格者，系統將會推薦一些輔助的資訊，例如：測驗前三名學生的學習路徑、透過 Google 學術搜尋的補充教材，學生也可以透過摘要系統讓閱讀上也更加便利，如圖 5、6 所示。

推薦學習路徑列表

標題
請依序閱讀第一章(完整版)>第二章(精簡版)
請依序閱讀第四章(精簡版)>第五章(精簡版)>第六章(完整版)

圖 5：學習路徑推薦

紅字為不及格

首頁：成績查詢

考卷名稱	分數	及格分數	訊息
小考1	40	60	請依序閱讀第一章精簡版.../補充教材
小考2	60	60	
小考3	65	60	
小考4	60	60	
期末考	80	70	

圖 6：學習路徑（成績前三名學生的學習路徑）/補充教材（學術搜尋的文章）

本研究以南臺科大某班級為實驗對象，以下將列出三十位學生的實驗數據，章節均以第一章為主，表 1 為每位學生尚未觀看推薦影片前的數據，表 2 為閱讀時間超過 20 分鐘的學生，看過推薦影片後再次閱讀教材的數據，表 1、2 為從資料庫抓提出來的資料。

表 1 學生尚未看推薦影片的閱讀時間

	閱讀時間(分鐘)
學生 1	39
學生 2	19
學生 3	20
學生 4	40

學生 5	35
學生 6	42
學生 7	15
學生 8	18
學生 9	33
學生 10	36
學生 11	40
學生 12	42
學生 13	44
學生 14	18
學生 15	19
學生 16	20
學生 17	22
學生 18	25
學生 19	30
學生 20	32
學生 21	34
學生 22	40
學生 23	15
學生 24	18
學生 25	14
學生 26	39
學生 27	35
學生 28	41

學生 29	31
學生 30	40

表 2 再次閱讀的時間

	閱讀時間(分鐘)
學生 1	20
學生 4	19
學生 5	20
學生 6	25
學生 9	23
學生 10	18
學生 11	35
學生 12	40
學生 13	40
學生 14	15
學生 15	18
學生 16	15
學生 17	20
學生 18	22
學生 19	28
學生 20	30
學生 21	31
學生 22	33
學生 23	15
學生 24	15

學生 25	12
學生 26	35
學生 27	33
學生 28	40
學生 29	26
學生 30	34

3.3 學習路徑/學術搜尋推薦

在使用一週後，將系統所收集的資訊，透過 Apriori 演算法，計算出學習路徑並推薦給學生，讓每一位學生都能夠有個人專屬的學習路徑，做為學習上參考，最後透過測驗的方式進行評估每位學生的學習情況，測驗成績滿分為 100，及格分數為 60，以下將列出三十位學生的測驗成績，表 3 尚未看推薦內容前的測驗成績，與依照推薦的學習路徑閱讀教材後，再次測驗的成績，表 4 為閱讀不及格的學生，依照推薦前三名學生的學習路徑重新閱讀教材，或學術搜尋的文章後的再次測驗成績，表 3、4 為從資料庫提取出來的資料。

表 3 閱讀推薦內容前後成績比較表

	尚未看推薦內容前的測驗成績	依序推薦的學習路徑，閱讀教材後的測驗成績
學生 1	60	70
學生 2	55	65
學生 3	40	60
學生 4	45	50
學生 5	35	55
學生 6	50	60
學生 7	70	70
學生 8	55	50

學生 9	70	70
學生 10	70	70
學生 11	55	60
學生 12	60	60
學生 13	50	60
學生 14	70	75
學生 15	75	75
學生 16	60	60
學生 17	30	45
學生 18	55	65
學生 19	25	55
學生 20	30	50
學生 21	60	60
學生 22	30	35
學生 23	30	60
學生 24	60	85
學生 25	55	70
學生 26	30	50
學生 27	40	45
學生 28	40	60
學生 29	45	70
學生 30	50	50

表 4：閱讀推薦學習路徑後的成績

	尚未看推薦內容前的測驗成績	依序推薦的學習路徑，閱讀教材後的測驗成績	不及格的學生，依照推薦前三名學生的學習路徑，並重新閱讀教材/補充教材後
--	---------------	----------------------	-------------------------------------

			的測驗成績
學生 1	60	70	
學生 2	55	65	
學生 3	40	60	
學生 4	45	50	60
學生 5	35	55	65
學生 6	50	60	
學生 7	70	70	
學生 8	55	50	70
學生 9	70	70	
學生 10	70	70	
學生 11	55	60	
學生 12	60	65	
學生 13	50	50	60
學生 14	70	70	
學生 15	75	75	
學生 16	60	70	
學生 17	30	50	60
學生 18	55	55	60
學生 19	25	40	70
學生 20	30	50	65
學生 21	60	65	70
學生 22	30	35	50
學生 23	30	50	60
學生 24	60	75	75
學生 25	55	60	60
學生 26	30	55	65
學生 27	40	60	60
學生 28	40	40	60
學生 29	45	55	60

學生 30	50	50	50
-------	----	----	----

推薦系統實現的個人化的學習，透過推薦的方式使學生在學習上能夠更加容易，以影片推薦來說確實能夠降低閱讀時間，在學習上也提供更多元的方式。以推薦學習路徑而言，使學生在學習上有學習路徑可以參考，更適合應用在期中、期末考前，讓學生在複習時不會再沒有方向。最後透過測驗評估學習情況，以收集的數據來看，仍有幾位學生學習情況不理想，也因此由系統推薦前三名學生的學習路徑或學術文章提供參考，讓成績不理想的學生能獲得教材以外的學習參考。

4.結論

本研究結合了數位學習平台、資料探勘、推薦系統，以數位學習平台收集每個學生的資訊，並藉由資料探勘找出適合的學習路徑，最後透過推薦系統在適當的時後給予輔助，例如：推薦相關影片、依照每個人情況不同算出的學習路徑、前三名的學生的學習路徑、Google 學術搜尋的相關文獻，透過推薦的方式實現個人化學習及輔助學習的二十大特色，另一方面教師可以透過測驗成績、閱讀時間、評估學生是否有順利獲取新的知識。

然而數位學習平台在收集資訊上仍有些問題，最容易遇到的情況就是，收集到很多無意義的資料，例如學生要閱讀的教材是第二章，但卻點選成第一章..等情況，以致於影響到資料探勘計算的結果。藉由以上經驗，因此本研究未來在開始使用系統前將進行教育訓練，由老師帶著學生實際的操作後，學生對系統也有基本的了解，也能夠減少收集到無意義的資料。在系統設計上，預計加入判斷資料是否有意義的功能，以教材閱讀而言，閱讀時間太過於短的就可能是無意義的資料，透過這樣的方式就能夠精準的收集資料，使得推薦的內容更加有用。

在實驗的數據收集上則有些限制，主要是因為時間上的不足，所以沒辦法進行太長時間的實驗，也因此才將數據收集的時間定為一週，但後續則會持續收集資料，並驗證方法的有效性。

預計本研究未來可以加入學習風格，由學習風格將每位學生的特性區分出來，例如：偏好閱讀影音、文字類型..等，並透過學習風格改善推薦的內容，讓學生一開始就能夠看到專屬於自己學習風格的推薦內容。

誌謝

本研究承蒙中華民國科技部 MOST 105-2511-S-218 -003 -MY2 計畫經費補助，謹此致謝。

參考文獻：

- [1] 林紀慧，”個別差異和學習路徑策略與電腦超本文學習成效研究”，*新竹師院學報*，11 期，頁 1-14，1998
- [2] Agrawal, R. and Srikant, R., “Fast Algorithms for Mining Association Rules in Large Database,” *Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 487-499, 1994
- [3] Alva Muhammad, Qingguo Zhou, Ghassan Beydoun, Dongming Xu, Jun Shen ,”Learning path adaptation in online learning systems”, *2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design* , pp. 421-426, 2016
- [4] Allen ,D. *Microteaching, a description. ERIC document* (ED 019 224), 1967
- [5] Dong Gyu Lee ,Kwang Sun Ryu,Mohamed Bashir,Jang-Whan Bae & Keun Ho Ryu , "Discovering Medical Knowledge using Association Rule Mining in Young Adults with Acute Myocardial Infarction.[J]", *Journal of Medical Systems*, vol. 37, no. 2, pp. 1-10, 2013.
- [6] Moradi, M & Keyvanpour, M R “An analytical review of XML association rules mining [J] “, *Artificial Intelligence Review*, Vol. 43, no. 2, pp. 277-300, 2015.
- [7] Robin Burke . “Hybrid Recommender Systems: Survey and Experiments”,*User Modeling and User-Adapted Interaction*, Vol.12, pp 331-370 , 2002 .