

บทความวิจัย

การทำเหมืองข้อมูลของระบบการจัดการเรียนออนไลน์: กรณีศึกษาของระบบ ATutor

สุณี รักษาเกียรติศักดิ์*

บทคัดย่อ

การทำเหมืองข้อมูล (data mining) เป็นกระบวนการสร้างความรู้ (knowledge) จากฐานข้อมูลประมวลผลรายการ (transactional database) เพื่อใช้ในการบริหาร การตัดสินใจ และ การวางแผน เชิงกลยุทธ์ วัตถุประสงค์ของงานวิจัยนี้เพื่อทำเหมืองข้อมูลของระบบจัดการเรียนการสอนออนไลน์ ATutor โดยกำหนดโจทย์ในการทำเหมืองข้อมูล เพื่อสร้างความรู้เกี่ยวกับวิชาที่อยู่ในระบบ ได้แก่ การนำเสนอค่าสถิติของวิชา (statistics) การจัดกลุ่มวิชา (clustering) ว่ามีการใช้งานอยู่ในระดับใด การสร้างโมเดลสำหรับการทำนายระดับการใช้งานของวิชา (classification) และการพยากรณ์การเข้าใช้วิชา จากรากฐานประจำของวิชา (regression) ซอฟต์แวร์ที่ใช้ในการดำเนินการเป็นซอฟต์แวร์โอเพนซอร์สทั้งหมด ได้แก่ Pentaho open suit ใช้คอมโพเนนต์ Data Integration ชื่อ SPOON เพื่อจัดเตรียมข้อมูล บนฐานข้อมูล MySQL โปรแกรม R เพื่อวิเคราะห์สถิติเบื้องต้นสำหรับอธิบายลักษณะข้อมูล และ Weka เพื่อสร้างโมเดลการทำเหมืองข้อมูล ผลการวิเคราะห์ที่สามารถนำไปใช้ในการรายงานสถานะการใช้งาน ATutor ในการจัดการเรียนการสอน ในที่ประชุมผู้บริหารมหาวิทยาลัยเพื่อกระตุนให้ผู้บริหารของหน่วยงาน กำหนดกลยุทธ์ สร้างแรงจูงใจให้อาจารย์มีการใช้งานมากขึ้น

คำสำคัญ: การทำเหมืองข้อมูล อีเลิร์นนิ่ง เอติวเตอร์ แอลเอ็มเอส

Data Mining in Learning Management System: ATutor Case Study

Sunee Raksakietisak*

ABSTRACT

Data mining is a process of knowledge discovery from transactional database for the purpose of management, decision making and strategic planning. The objectives of this paper are to do data mining from Learning Management System “ATutor”. The problem tasks for mining include statistics of course characteristics, clustering of level of course usage, building classification model for level of course usage, and building regression model predicting course access from course characteristics. Open source software were used in data mingng include SPOON Data Integration of Pentaho open suit for data preparation from MySQL database, R program for descriptive statistics and Weka for data modeling. The results from data mining give status report of ATutor system usage in teaching and learning for each faculty and that the university administrators can use the results for strategic management and planning to motivate the faculty members in order to increase the usage.

Keywords: data mining, e-learning, ATutor LMS

บทนำ

ระบบ ATutor เป็นระบบจัดการเรียนออนไลน์ (Learning Management System หรือ LMS) โอเพนซอร์สที่พัฒนาโดย ATRC (The Adaptive Technology Resource Center) มหาวิทยาลัยโตรอนโต ประเทศแคนาดา ระบบได้มีการพัฒนาต่อเนื่องตั้งแต่ปี พ.ศ. 2545 (ค.ศ. 2002) [1]

มหาวิทยาลัยครินทร์วิโรฒ หรือ มศว. ได้ปรับแต่งและใช้งานตั้งแต่ภาคการศึกษาที่ 1/2546 [2] ได้มีการพัฒนาอุดมการวิเคราะห์ข้อมูลให้กับระบบการทดสอบออนไลน์ของ ATutor [3]

วัตถุประสงค์ของการวิจัย เพื่อประยุกต์ใช้หลักการทำเหมืองข้อมูล (data mining) สร้างความรู้ (knowledge) จากฐานข้อมูลประมวลผลรายการ (transaction) ของระบบจัดการการเรียนแบบออนไลน์ ATutor ของมหาวิทยาลัยครินทร์วิโรฒเพื่อให้ได้สารสนเทศที่ผู้บริหารมหาวิทยาลัยสามารถใช้กำหนดแผนกลยุทธ์ที่จะส่งเสริม กระตุ้น สร้างแรงจูงใจ ให้นิวยางและคณาจารย์ใช้การเรียนการสอนแบบออนไลน์ผสมผสานกับการเรียนการสอนแบบปกติให้เกิดเป็นการจัดการเรียนการสอนแบบไฮบริด (hybrid) มากขึ้น [4]

กรอบแนวคิดเชิงทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

การทำเหมืองข้อมูล เป็นกระบวนการสร้างความรู้จากฐานข้อมูลประมวลผลรายการ เพื่อใช้ในการบริหาร การตัดสินใจ และการวางแผนเชิงกลยุทธ์ โดยใช้วิธีทางสถิติ (statistics) หรือวิธีการเรียนรู้ของเครื่อง (machine learning) โดยการพัฒนาอัลกอริทึมส์ที่สร้างตัวแบบการเรียนรู้หรือโมเดลจากข้อมูลจำนวนมาก (training set) และประเมินตัวแบบว่ามีความถูกต้องเพียงใดจากข้อมูลอีกชุดหนึ่ง (test set) [5, 6]

ประเภทของโจทย์/ปัญหา ที่ทำเหมืองข้อมูล (Problem task)

จากการต้องการของความรู้ที่ต้องการจากข้อมูลในฐานข้อมูลที่มีจำนวนมาก ผู้ทำเหมืองข้อมูล ต้องสามารถที่จะแปลงความต้องการให้ชัดว่าเป็นประเภทโจทย์/ปัญหาแบบใด เพื่อที่จะใช้วิธีการได้ถูกต้อง ได้แก่

- การสรุปข้อมูล (descriptive)

เป็นการสรุปข้อมูลให้ได้สารสนเทศที่นำไปใช้ได้ โดยนำเสนอด้วยสถิติ (statistics) เช่น จำนวนและร้อยละ ค่ากลางและค่าการกระจาย หรือค่าเบอร์เซ็นไทล์ต่างๆ อาจจะแสดงผลลัพธ์เป็นแบบกราฟ (visualization)

- การจำแนกประเภท (classification)

เป็นการสร้างโมเดลสำหรับการจำแนกประเภทข้อมูลจากคุณสมบัติหรือแอตทริบิวต์ (attribute) ของข้อมูล โดยต้องรู้ประเภทของข้อมูล (class) ผลลัพธ์ที่ได้จะเป็นกฎเพื่อรับประทานของข้อมูลจาก แอตทริบิวต์ของข้อมูล ตัวอย่างอัลกอริทึมส์ เช่น ID3, J48 เป็นต้น ความรู้ที่ได้จะเป็นกฎสำหรับจำแนกประเภท

- การพยากรณ์ (prediction)

เป็นการสร้างโมเดลสำหรับการพยากรณ์แอตทริบิวต์ตัวใดตัวหนึ่งจากแอตทริบิวต์ตัวอื่น ตัวอย่างอัลกอริทึม เช่น Multiple Linear Regression ความรู้ที่ได้จะเป็นสมการพยากรณ์

- การแบ่งกลุ่มข้อมูล (clustering)

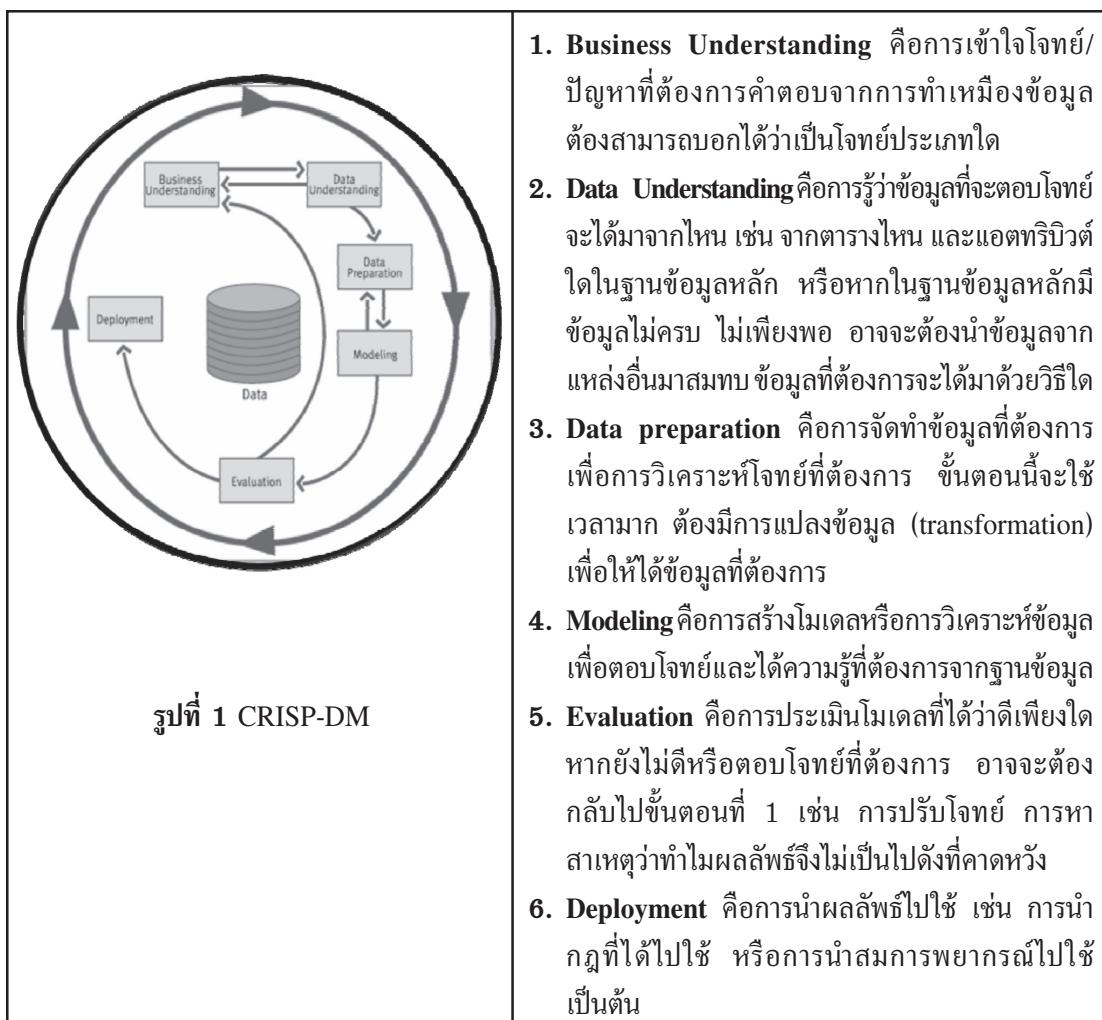
เป็นการแบ่งข้อมูลที่มีลักษณะคล้ายกันออกเป็นกลุ่มโดยใช้แอดทริบิวต์ของข้อมูลเป็นตัวแบ่ง ตัวอย่างอัลกอริทึม เช่น Simple K-Mean ผลลัพธ์ที่ได้จะเป็นการจัดกลุ่มข้อมูลได้ centroid ของกลุ่มเป็นค่าเฉลี่ย (mean) ของแอดทริบิวต์ต่างๆ

- การหาความสัมพันธ์ของข้อมูล (Association)

เป็นการหาความสัมพันธ์ของแอดทริบิวต์ของข้อมูลว่า แอดทริบิวต์ใดมีความสัมพันธ์กับแอดทริบิวต์ใด ตัวอย่างต้นแบบ คือ โจทย์ที่เรียกว่า การวิเคราะห์ตระกร้าสินค้า (Market Basket Analysis) เพื่อพิจารณาลินค้าที่ผู้ซื้อมักจะซื้อพร้อมกัน ตัวอย่างอัลกอริทึม เช่น Apriori ผลลัพธ์ที่ได้จะเป็นกฎความสัมพันธ์แบบ IF ... THEN

มาตรฐานกระบวนการทำเหมืองข้อมูล

CRISP-DM [7] หรือ CRoss Industry Standard Process เป็นเว็บไซต์อ้างอิงสำหรับมาตรฐานกระบวนการทำเหมืองข้อมูล ดังแสดงในรูปที่ 1 มี 6 ขั้นตอน ได้แก่



การทำเหมืองข้อมูลระบบ LMS

มีงานวิจัยที่ทำเหมืองข้อมูลเพื่อค้นหาความรู้จากการจัดการเรียนการสอนในระบบที่หลากหลาย [8] ทั้งระบบที่เป็นเชิงพาณิชย์ และระบบที่เป็นโอเพนซอร์ส เช่น Moodle [9, 10] แต่เท่าที่ทราบยังไม่มีการดำเนินการทำเหมืองข้อมูลในระบบ ATutor ดังนั้นการดำเนินการนี้จะเป็นประโยชน์ต่อผู้ใช้ระบบ ATutor ในการจัดการเรียนการสอนออนไลน์ ซึ่งเป็นระบบหนึ่งที่มีการใช้มากในประเทศไทย [11]

ซอฟต์แวร์ที่ใช้ในการทำเหมืองข้อมูล

ซอฟต์แวร์ที่ใช้ในการดำเนินการเกี่ยวกับการทำเหมืองข้อมูลอาจจะต้องใช้หลายตัวประกอบกัน [12] ทั้งซอฟต์แวร์ที่ช่วยในการเตรียมข้อมูล ซอฟต์แวร์ทางสถิติ และซอฟต์แวร์ทางการเรียนรู้ของเครื่อง ซอฟต์แวร์ที่ใช้ในงานวิจัยนี้เป็นซอฟต์แวร์โอเพนซอร์สทั้งหมด ได้แก่

- Pentaho open suit [13] ใช้คอมโพเนนต์ Data Integration ชื่อ SPOON เพื่อจัดเตรียมข้อมูล บนฐานข้อมูล MySQL ควบคู่กับ Microsoft Excel
- โปรแกรม R [14] เพื่อวิเคราะห์สถิติเบื้องต้นสำหรับอธิบายลักษณะข้อมูล
- Weka [15] เพื่อสร้างโมเดลการทำเหมืองข้อมูล

วิธีการศึกษา

ฐานข้อมูลที่ใช้ในการทำเหมืองข้อมูลเป็นฐานข้อมูลของระบบ ATutor ในหนึ่งปีการศึกษา ตั้งแต่วันที่ 25 ตุลาคม 2550 (2007-10-25) ถึงวันที่ 17 ตุลาคม 2551 (2008-10-17) ซึ่งมีข้อมูลของสองภาคการศึกษาคือ ภาคการศึกษาที่ 2/2550 และภาคการศึกษาที่ 1/2551 [16]

โจทย์ที่ต้องการคำตอบจากการทำเหมืองข้อมูล

ข้อ 1: จำนวนวิชาใน ATutor ของแต่ละคณะ (faculty) ในแต่ละภาคการศึกษา (semester) แยกตามประเภทวิชา (course_type) ว่าเป็นวิชาตามหลักสูตรของมหาวิทยาลัย หรือวิชาตามอธิบายด้วย

ข้อ 2: มีการใช้งานในลักษณะต่างๆ มากน้อยเพียงใด เช่น การสร้างเพจเนื้อหา การสร้างข้อสอบออนไลน์ การสร้างประกาศ การสร้างแหล่งเรียนรู้รายวิชา การมอบหมายงาน/การบ้าน เป็นต้น และผลของการเข้าใช้งานของนิสิต เช่น จำนวนนิสิตที่ลงทะเบียนเรียน จำนวนครั้งของการเข้าใช้วิชา เป็นต้น

ข้อ 3: ในรายวิชาทั้งหมดสามารถจัดกลุ่มวิชาเพื่อบอกสถานะของการใช้งานว่ามีการใช้งานมากน้อยเพียงใด เช่น เป็นวิชาที่ไม่ค่อยได้ใช้งาน วิชาที่มีการใช้งานในระดับปานกลาง หรือวิชาที่มีการใช้งานในระดับมาก เพื่อสามารถใช้เป็นเกณฑ์ในการสร้างแรงจูงใจในการให้การงาน

ข้อ 4: จำนวนอาจารย์ที่ใช้ ATutor ของแต่ละคณะ (faculty) ในแต่ละภาคการศึกษา (semester) แยกตามสถานะการใช้งาน (ActiveStatus)

ข้อ 5: จากข้อ 3 สร้างโมเดลของการจัดกลุ่มว่าการที่จะจำแนกวิชาว่ามีการใช้งานในระดับใดใช้เกณฑ์ใดมาจับ

ข้อ 6: แอตทริบิวต์หรือตัวแปรใดที่ส่งผลต่อการเข้าใช้งานของวิชาบ้าง

ฐานข้อมูลและแอตทริบิวต์ที่เกี่ยวข้อง

ข้อมูลในฐานข้อมูลที่ใช้ในการเตรียมข้อมูลสำหรับการวิเคราะห์แสดงในตารางที่ 1

ตารางที่ 1 ตารางข้อมูลที่ใช้ในการเตรียมข้อมูล

ชื่อตาราง	คำอธิบาย	แอตทริบิวต์สำคัญ	ระเบียน
AT_member	ผู้ใช้	member_id, email, first_name, last_name	26,553
AT_courses	วิชา	course_id, member_id, created_date, title	405
AT_content	เพจเนื้อหา	content_id, course_id	9,841
AT_tests	ข้อสอบ	test_id, course_id	527
AT_news	ประกาศ	news_id, course_id	1,072
AT_links	แหล่งวิทยาการ	link_id, cat_id	1,517
AT_links_categories		cat_id, owner_id*	506
AT_assignments	งาน/การบ้าน	assignment_id, course_id	254
AT_faq_topic	คำถามที่ถามบ่อย	topic_id, course_id	52
AT_glossary	อภิธานศัพท์	word_id, course_id	1,281
AT_polls	โพล	poll_id, course_id	86
AT_course_enrollment	นิสิตที่ลงทะเบียน	member_id, course_id	7,267
AT_course_stats	การเข้าใช้วิชา	course_id, login_date, guests, members	29,809
AT_member_track	การเข้าดูเพจเนื้อหาของนิสิต	member_id, course_id, content_id	38,305
AT_forums	กระดานสนทนา	forum_id, num_topics, num_posts	401
AT_forums_courses		forum_id, course_id	302
AT_blog_posts	บล็อก	post_id, owner_id*	412

*หมายเหตุ: เหนืออนกับ course_id

ตัวแปร/แอตทริบิวต์ที่ต้องการสำหรับตอบโจทย์

แอตทริบิวต์เกี่ยวกับวิชาที่ต้องการแบ่งเป็น 3 ส่วน ดัง

1. แอตทริบิวต์ของวิชา ได้แก่

- course_id: รหัสวิชา
- member_id: รหัสเจ้าของวิชา
- created_date: วันที่สร้างวิชา
- title: ชื่อวิชา
- email: อีเมลของเจ้าของวิชา ซึ่งจะกำหนดให้เป็นอีเมลที่ มศว. ออกให้เรียกว่าบัตรคุณภาพ
- first_name: ชื่อเจ้าของวิชา
- last_name: นามสกุลเจ้าของวิชา
- fac_id: รหัสหน่วยงานในระดับคณะ
- fac_name: ชื่อหน่วยงานในระดับคณะ
- semester: ภาคการศึกษาที่วิชานั้นเปิดใช้งาน (2/2550 หรือ 1/2551)
- course_type: ประเภทของวิชา มีค่าเป็น SWU หรือ GEN (SWU คือ วิชาตามหลักสูตรของมหาวิทยาลัย และ GEN คือวิชาตามอัธยาศัยที่ผู้สอนประยุกต์ใช้ ATutor)

แอตทริบิวต์ในส่วนนี้จะได้จากการ AT_course, AT_member โดยใช้ member_id เป็นตัวเชื่อมเพื่อนำค่าของ email (จะเป็น email ของ มศว. เท่านั้น ซึ่งเรียกว่าบัตรคุณภาพ) ไปแมปกับข้อมูลของฐานข้อมูลมหาวิทยาลัย (SUPREME) เพื่อดึงค่า fac_id, fac_name ให้กับ member_id นั้นๆ เพื่อที่จะทราบว่าวิชานั้นเป็นของหน่วยงานใด ส่วน semester ได้จาก created_date ตามเงื่อนไขต่อไปนี้ ถ้า created_date อยู่ระหว่าง 2007-10-25 ถึง 2008-02-28 semester = 2_2550 ที่เหลือจะเป็น semester = 1_2551 ส่วน course_type จะดูที่ title ถ้า title ขึ้นต้นด้วยรหัสตามหลักสูตรของมหาวิทยาลัย course_type = SWU ที่เหลือจะเป็น course_type = GEN

2. แอตทริบิวต์ในส่วนที่เป็นเนื้อหาที่ผู้สอนเป็นคนสร้าง ส่วนนี้จะเป็นส่วนที่เรียกว่าคอร์สแวร์ (courseware) ซึ่งผู้สอนสามารถที่จะทำการ Backup วิชาเมื่อถึงการสอนในแต่ละภาคการศึกษา และนำมา Restore สำหรับการสอนในภาคการศึกษาต่อไป โดยไม่ต้องสร้างใหม่ เพียงแต่ปรับปรุงบางส่วน ได้แก่

- content: จำนวนเพจเนื้อหา (Content Pages)
- announce: จำนวนประกาศ (Announcements/News)
- test: จำนวนข้อสอบ/แบบสอบถาม (Test & Surveys)
- link: จำนวนแหล่งวิทยาการรายวิชา (Links)
- assign: จำนวนงานที่มอบหมาย (Assignments)
- glossary: จำนวนอภิธานศัพท์ (Glossary)
- faq: จำนวนคำถามที่ถามบ่อย (FAQ)
- poll: จำนวนโพล (Polls)

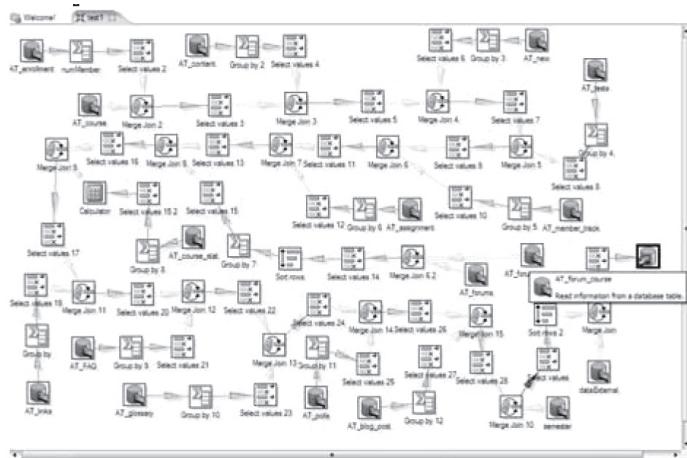
3. แอตทริบิวต์ในส่วนการใช้งานของผู้เรียน โดยแอตทริบิวต์ในส่วนนี้จะเป็นข้อมูลของการใช้งานวิชาการในการสอนภาคการศึกษานั่นๆ หากการสอนเสร็จสิ้นแล้วเมื่อ Backup วิชา และนำมา Restore เพื่อการสอนในภาคการศึกษาต่อไป ข้อมูลส่วนนี้จะไม่มา ได้แก่

- enroll: จำนวนผู้ลงทะเบียนเรียน
- coAccess: จำนวนครั้งของการเข้าใช้วิชา
- paAccess: จำนวนครั้งของการเข้าดูเพจเนื้อหา
- forumTopic: จำนวนหัวข้อ (กระทู้) ของ Forums
- forumPost: จำนวนการแสดงความคิดเห็นในหัวข้อต่างๆ
- blogPost: จำนวน Blogs ของวิชา

การจัดเตรียมข้อมูลเพื่อการวิเคราะห์

ดำเนินการเตรียมข้อมูลโดยใช้โปรแกรมในส่วนของ Data Integration ชื่อ SPOON ของ Pentaho open suit ซึ่งเป็นโปรแกรมในการทำ ETL (Extract Transform Loading) ที่เป็น GUI ที่นำ Input ที่ต้องการจากฐานข้อมูล ATutor (extract) มาทำการประมวลค่าที่ต้องการหรือทราบส์ฟอร์ม (transform) และนำผลลัพธ์สุดท้ายจัดเก็บลงฐานข้อมูล (loading) และสามารถจัดเก็บเป็นไฟล์ประมวลผล (executable file) ทำให้สามารถปรับเปลี่ยนในส่วนที่ต้องการได้ ไม่ต้องดำเนินการใหม่ทั้งหมด

ตัวอย่างการประมวลค่าแอตทริบิวต์ของวิชาที่ต้องการ เช่น นับจำนวนนิสิตที่ลงทะเบียนเรียน ในแต่ละวิชาในตาราง AT_course_enrollment และนำผลลัพธ์ที่ได้มา left join กับตาราง AT_course ก็จะได้แอตทริบิวต์ของจำนวนนิสิตที่ลงทะเบียนในแต่ละวิชา (Enroll) ในการทำเดินการแต่ละทราบส์ฟอร์ม เมนชัน (transformation) ได้มีการตรวจสอบความถูกต้องด้วยคำสั่ง SQL จากฐานข้อมูล การประมวลผลค่าที่ต้องการในกระบวนการ ETL ดังแสดงในรูปที่ 2



รูปที่ 2 ไฟล์ประมวลผลของการทำ ETL

ในการดำเนินการจะได้ไฟล์ข้อมูลวิชาทั้งหมด 405 รายการ/วิชา

การวิเคราะห์ข้อมูลเพื่อตอบโจทย์

ข้อ 1: ใช้ R ทำ Contingency Tables สำหรับตัวแปร semester และ course_type

ข้อ 2: ใช้ R ทำ Summaries > Frequency distributions ของตัวแปรใน 3.3 โดยเลือก เนพาระวิชาที่มี course_type เป็น SWU ซึ่งมีทั้งหมด 273 วิชา โดยหาค่า Mean, Min, Max, P10, P25, P50, P75, P90

ข้อ 3: ใช้ Weka ทำ clustering ด้วยอัลกอริทึม Simple K = Mean และนำผลลัพธ์ไปกำหนดกลุ่มให้กับวิชา และใช้ R ทำ Contingency Tables เพื่อแสดงสถานะการใช้งานของวิชา

ข้อ 4: กำหนดว่าอาจารย์คนใด Active หรือไม่โดยใช้เกณฑ์ว่า ถ้ามี 1 วิชาที่ Active จะถือว่า อาจารย์คนนั้น Active การแสดงค่าร้อยละได้นำข้อมูลของจำนวนอาจารย์ที่ปฏิบัติหน้าที่ของแต่ละคณะจาก SWU Info Service [17]

ข้อ 5: ใช้ Weka ทำ classify ด้วยอัลกอริทึม J48 เพื่อให้ได้กฎ โดยการกำหนดตัวแปร class จากผลการจัดกลุ่มของข้อ 3

ข้อ 6: ใช้ Weka ทำ Linear Regression ของตัวแปร coAccess โดยมีตัวอธิบายหรือตัวพยากรณ์ที่เลือกมา ได้แก่ content, announce, test, link, assign, enroll ส่วนแอดทริบิวต์อื่นที่ไม่ได้เลือกมาเนื่องด้วยมีการใช้งานน้อย

ผลการศึกษา

ผลข้อ 1 แสดงในตารางที่ 2 ซึ่งจะเห็นได้ว่าในภาพรวมจำนวนวิชามากขึ้นจากการศึกษาหนี้ไปยังภาคการศึกษาถัดไป และได้สารสนเทศของการใช้งานของแต่ละหน่วยงานเพื่อที่ผู้บริหารสามารถนำไปใช้เชิงนโยบายได้

ตารางที่ 2 จำนวนและร้อยละของวิชาของแต่ละหน่วยงาน แยกตามประเภทวิชาและการศึกษา

คณะ/หน่วยงาน	ประเภท วิชา	ภาคการศึกษา		รวม
		2/2550	1/2551	
ไม่ทราบ	GEN	1 (3.3%)	29 (96.7%)	30 (100.0%)
	SWU	1 (16.7%)	5 (83.3%)	6 (100.0%)
	รวม	2 (5.6%)	34 (94.4%)	36 (8.9%)
คณะทันตแพทยศาสตร์	GEN	1 (33.3%)	2 (66.7%)	3 (100.0%)
	SWU	4 (100.0%)	0 (0.0%)	4 (100.0%)
	รวม	5 (71.4%)	2 (28.6%)	7 (1.7%)
คณะพยาบาลศาสตร์	GEN	0 (0.0%)	2 (100.0%)	2 (100.0%)
	SWU	4 (57.1%)	3 (42.9%)	7 (100.0%)
	รวม	4 (44.4%)	5 (55.6%)	9 (2.2%)

ตารางที่ 2 (ต่อ)

คณะ/หน่วยงาน	ประเภท วิชา	ภาคการศึกษา		รวม
		2/2550	1/2551	
คณะพลศึกษา	GEN	2 (100.0%)	0 (0.0%)	2 (100.0%)
	SWU	3 (75.0%)	1 (25.0%)	4 (100.0%)
	รวม	5 (83.3%)	1 (16.7%)	6 (1.5%)
คณะแพทยศาสตร์	GEN	8 (25.0%)	24 (75.0%)	32 (100.0%)
	SWU	16 (53.3%)	14 (46.7%)	30 (100.0%)
	รวม	24 (38.7%)	38 (61.3%)	62 (15.3%)
คณะเภสัชศาสตร์	GEN	0 (0.0%)	2 (100.0%)	2 (100.0%)
	SWU	1 (25.0%)	3 (75.0%)	4 (100.0%)
	รวม	1 (16.7%)	5 (83.3%)	6 (1.5%)
คณะมนุษยศาสตร์	GEN	0 (0.0%)	10 (100.0%)	10 (100.0%)
	SWU	18 (18.8%)	78 (81.3%)	96 (100.0%)
	รวม	18 (17.0%)	88 (83.0%)	106 (26.2%)
คณะวิทยาศาสตร์	GEN	5 (22.7%)	17 (77.3%)	22 (100.0%)
	SWU	36 (42.4%)	49 (57.6%)	85 (100.0%)
	รวม	41 (38.3%)	66 (61.7%)	107 (26.4%)
คณะวิศวกรรมศาสตร์	GEN	1 (25.0%)	3 (75.0%)	4 (100.0%)
	SWU	8 (57.1%)	6 (42.9%)	14 (100.0%)
	รวม	9 (50.0%)	9 (50.0%)	18 (4.4%)
คณะศิลปกรรมศาสตร์	GEN	0 (0.0%)	1 (100.0%)	1 (100.0%)
	SWU	1 (14.3%)	6 (85.7%)	7 (100.0%)
	รวม	1 (12.5%)	7 (87.5%)	8 (2.0%)
คณะพีಠ迦ศาสตร์	GEN	0 (0.0%)	2 (100.0%)	2 (100.0%)
	SWU	6 (42.9%)	8 (57.1%)	14 (100.0%)
	รวม	6 (37.5%)	10 (62.5%)	16 (4.0%)
คณะสังคมศาสตร์	GEN	0 (50.0%)	0 (50.0%)	0 (100.0%)
	SWU	0 (0.0%)	4 (100.0%)	4 (100.0%)
	รวม	0 (0.0%)	4 (100.0%)	4 (1.0%)

ตารางที่ 2 (ต่อ)

คณะ/หน่วยงาน	ประเภท วิชา	ภาคการศึกษา		รวม
		2/2550	1/2551	
บัณฑิตวิทยาลัย	GEN	0 (0.0%)	2 (100.0%)	2 (100.0%)
	SWU	0 (0.0%)	1 (100.0%)	1 (100.0%)
	รวม	0 (0.0%)	3 (100.0%)	3 (.7%)
สถาบันวิจัยพุทธกรรมศาสตร์	GEN	0 (50.0%)	0 (50.0%)	0 (100.0%)
	SWU	1 (33.3%)	2 (66.7%)	3 (100.0%)
	รวม	1 (33.3%)	2 (66.7%)	3 (.7%)
สำนักคอมพิวเตอร์	GEN	3 (23.1%)	10 (76.9%)	13 (100.0%)
	SWU	0 (0.0%)	0 (0.0%)	0 (100.0%)
	รวม	3 (23.1%)	10 (76.9%)	13 (3.2%)
สำนักงานอธิการบดี	GEN	1 (100.0%)	0 (0.0%)	1 (100.0%)
	SWU	0 (0.0%)	0 (0.0%)	0 (0.0%)
	รวม	1 (100.0%)	0 (0.0%)	1 (.2%)
รวมทุกหน่วยงาน		121 (29.9%)	284 (70.1%)	405 (100%)

หมายเหตุ: ประเภทวิชา SWU หมายถึง วิชาตามหลักสูตรของมหาวิทยาลัย

GEN หมายถึง วิชาอื่นๆ ที่ไม่ได้เป็นวิชาตามหลักสูตรของมหาวิทยาลัย หรือ
อาจจะเป็นวิชาที่ใช้ในการบริหารจัดการความรู้และการสื่อสาร

ผลข้อ 2 แสดงในตารางที่ 3 ซึ่งในภาพรวมมีการใช้งานมากในลำดับต้นๆ ได้แก่ การสร้างเพจเนื้อหา (content) การสร้างประกาศ (announce) การสร้างแบบทดสอบ (test) การสร้างแหล่งเรียนรู้รายวิชา (link) และการมอบหมายงาน (assign) ส่วนการใช้เครื่องมืออื่นในการจัดการการเรียนรู้ยังมีน้อย เช่น faq poll เป็นต้น ส่วนการเข้าใช้งานของนิสิตมากพอสมควร ที่ยังน้อยอยู่คือ log_post ซึ่งควรส่งเสริมให้มากขึ้น

ผลข้อ 3 แสดงในตารางที่ 4, 5 โดยการทดลองจัดกลุ่มวิชาออกเป็น 2, 3, 4, 5 กลุ่ม ตามลำดับ และพบว่าความเหมาะสมสมอยู่ที่ 5 กลุ่ม (SSE = 20.32) คือ InActive, ActiveL1, ActiveL2, ActiveL3, ActiveL4, ActiveL5 โดย Centroid ของกลุ่มนี้ถูกแนะนำและแสดงในตารางที่ 4 และตารางที่ 5 แสดงจำนวนและร้อยละของวิชาของแต่ละหน่วยงานแยกตามสถานะการใช้งาน (InActive/Active) โดยวิชาที่อยู่ในกลุ่ม ActiveL1 ถึง ActiveL5 นับว่า Active

ตารางที่ 3 สถิติเบื้องต้นสำหรับตัวแปร/แอ็ตทริบิวต์ของวิชา

	Mean	Min	Max	P10	P25	P50	P75	P90
content	26.7	0	433	1.0	2.0	9.0	19.5	50.8
	14.34	0	50	1.0	2.0	9.0	19.5	49.6
announce	3.1	0	33	1.0	1.0	1.0	3.0	7.6
test	1.3	0	31	0.0	0.0	0.0	2.0	3.0
link	1.4	0	17	0.0	0.0	0.0	1.0	4.6
assign	0.8	0	19	0.0	0.0	0.0	0.0	2.0
glossary	4.1	0	270	0.0	0.0	0.0	0.0	1.0
faq	.2	0	7	0.0	0.0	0.0	0.0	0.0
poll	0.2	0	30	0.0	0.0	0.0	0.0	0.6
enroll	22.5	1	680	1.0	1.0	3.0	26.5	53.0
coAccess	576.9	0	30693	0.0	0.5	45.0	408.5	1219.2
paAccess	123.8	0	3794	0.0	0.0	0.0	69.0	371.2
forumTopic	4.4	0	392	0.0	0.0	0.0	0.0	3.6
topicicPost	31.6	0	1372	0.0	0.0	0.0	0.0	25.6
blogPost	1.0	0	42	0.0	0.0	0.0	1.0	3.0

หมายเหตุ: P90 ของ Content จะอยู่ที่ 50.8 ดังนั้นหากปรับส่วนที่เป็น outlier ให้กลับคืนมา P90 คือ 50 เมื่อปรับแล้วจะได้ค่าเฉลี่ยของ Content เป็น 14.34

ตารางที่ 4 ค่าเฉลี่ยของแต่ละแอตทริบิวต์ของวิชาจากการจัดกลุ่ม 5 กลุ่ม (InAct/ActL1/ActL2/ActL3/ActL4)

แอตทริบิวต์	All	InAct	ActL1	ActL2	ActL3	ActL4
จำนวนวิชา	273	144	36	73	8	12
	100%	53%	13%	27%	3%	4%
content	14.3	3.3	46.1	16.7	18	34.5
announce	3.1	1.6	3.7	4.7	5.6	7.9
test	1.3	0.5	3.2	1.4	1.1	4.5
link	1.4	0.3	1.6	1.7	2.5	11.3
assign	0.7	0.2	0.4	0.5	3.8	0.9
enroll	22.5	7.7	10.6	33.5	40.0	157.8
coAccess	576.9	108.3	194.2	649.8	2033.0	5935.4
paAccess	123.8	11.6	112.8	163.4	404.8	1075.4
forumTopic	4.4	0.2	1.8	4.4	10.6	58.9
topicicPost	31.6	1.4	3.8	57.4	76.3	291.0

หมายเหตุ: วิชาที่อยู่ในกลุ่ม ActL1, ActL2, ActL3, ActL4 คือวิชาที่นับว่าอยู่ในสถานะ Active

ตารางที่ 5 จำนวนและร้อยละของวิชาของแต่ละหน่วยงาน แยกตามสถานะการใช้ (InActive/Active) และภาคการศึกษา สำหรับวิชาตามหลักสูตรของมหาวิทยาลัย

คณะ/หน่วยงาน	สถานะ	ภาคการศึกษา 2/2550	ภาคการศึกษา 1/2551	รวม
คณะทันตแพทยศาสตร์	InActive	1 (100.0%)	0 (0.0%)	1 (100.0%)
	Active	3 (100.0%)	0 (0.0%)	3 (100.0%)
	รวม	4 (100.0%)	0 (0.0%)	4 (1.5%)
คณะพยาบาลศาสตร์	InActive	3 (50.0%)	3 (50.0%)	6 (100.0%)
	Active	1 (100.0%)	0 (0.0%)	1 (100.0%)
	รวม	4 (44.4%)	3 (55.6%)	7 (2.6%)
คณะพลศึกษา	InActive	2 (66.7%)	1 (33.3%)	3 (100.0%)
	Active	1 (100.0%)	0 (0.0%)	1 (100.0%)
	รวม	3 (75.0%)	1 (25.0%)	4 (1.5%)
คณะแพทยศาสตร์	InActive	9 (56.3%)	7 (43.8%)	16 (100.0%)
	Active	7 (50.0%)	7 (50.0%)	14 (100.0%)
	รวม	16 (53.3%)	14 (46.7%)	30 (11.0%)
คณะเภสัชศาสตร์	InActive	0 (0.0%)	1 (100.0%)	1 (100.0%)
	Active	1 (33.3%)	2 (66.7%)	3 (100.0%)
	รวม	1 (25.0%)	3 (75.0%)	4 (1.5%)
คณะมนุษยศาสตร์	InActive	6 (10.9%)	49 (89.1%)	55 (100.0%)
	Active	12 (29.3%)	29 (70.7%)	41 (100.0%)
	รวม	18 (18.8%)	78 (81.3%)	96 (35.2%)
คณะวิทยาศาสตร์	InActive	11 (32.4%)	23 (67.6%)	34 (100.0%)
	Active	25 (49.0%)	26 (51.0%)	51 (100.0%)
	รวม	36 (42.4%)	49 (57.6%)	85 (31.1%)
คณะวิศวกรรมศาสตร์	InActive	4 (50.0%)	4 (50.0%)	8 (100.0%)
	Active	4 (66.7%)	2 (33.3%)	6 (100.0%)
	รวม	8 (57.1%)	9 (50.0%)	14 (5.1%)
คณะศิลปกรรมศาสตร์	InActive	1 (25.0%)	3 (75.0%)	4 (100.0%)
	Active	0 (0.0%)	3 (100.0%)	3 (100.0%)
	รวม	1 (14.3%)	8 (85.7%)	7 (2.6%)

ตารางที่ 5 (ต่อ)

คณะ/หน่วยงาน	สถานะ	ภาคการศึกษา 2/2550	ภาคการศึกษา 1/2551	รวม
คณะศึกษาศาสตร์	InActive	3 (30.0%)	7 (70.0%)	10 (100.0%)
	Active	3 (75.0%)	1 (25.0%)	4 (100.0%)
	รวม	6 (42.9%)	8 (57.1%)	14 (5.1%)
คณะสังคมศาสตร์	InActive	0 (0.0%)	3 (100.0%)	3 (100.0%)
	Active	0 (0.0%)	1 (100.0%)	1 (100.0%)
	รวม	0 (0.0%)	4 (100.0%)	4 (1.5%)
บัณฑิตวิทยาลัย	InActive	0 (50.0%)	0 (50.0%)	0 (100.0%)
	Active	0 (0.0%)	1 (100.0%)	1 (100.0%)
	รวม	0 (50.0%)	1 (50.0%)	1 (0.4%)
สถาบันวิจัยพุทธกรรมศาสตร์	InActive	1 (33.3%)	2 (66.7%)	3 (100.0%)
	Active	0 (50.0%)	0 (50.0%)	0 (100.0%)
	รวม	1 (33.3%)	2 (66.7%)	3 (1.1%)
รวมทุกหน่วยงาน		98 (35.9%)	175 (64.1%)	273 (100%)

ผลข้อ 4 แสดงในตารางที่ 6 โดยมีการให้ข้อมูลว่าถ้าอาจารย์คนใดมีวิชาที่ Active อย่างน้อย 1 วิชาจะถือว่าเป็นอาจารย์ที่ Active ใน การใช้ระบบ ATutor

ตารางที่ 6 จำนวนและร้อยละของอาจารย์ที่ใช้ ATutor แยกตามคณะ และสถานะการใช้งาน (InActive/Active)

คณะ	อาจารย์ ทั้งหมด*	อาจารย์ที่ใช้		
		รวม	InActive	Active
คณะทันตแพทยศาสตร์	52	3 (5.8%)	0 (0.0%)	3 (100.0%)
คณะพยาบาลศาสตร์	17	5 (29.4%)	4 (80.0%)	1 (20.0%)
คณะพลศึกษา	80	4 (5.0%)	3 (75.0%)	1 (25.0%)
คณะแพทยศาสตร์**	197	17 (8.6%)	8 (47.1%)	9 (52.9%)
คณะเภสัชศาสตร์	34	3 (8.8%)	0 (0.0%)	3 (100.0%)
คณะมนุษยศาสตร์	113	42 (37.2%)	23 (54.8%)	19 (45.2%)
คณะวิทยาศาสตร์	140	30 (21.4%)	9 (30.0%)	21 (70.0%)
คณะวิศวกรรมศาสตร์	49	4 (8.2%)	1 (25.0%)	3 (75.0%)
คณะศิลปกรรมศาสตร์	56	4 (7.1)	2 (50.0%)	2 (50.0%)
คณะศึกษาศาสตร์	85	6 (7.0%)	3 (50.0%)	3 (50.0%)
คณะสังคมศาสตร์	140	2 (1.4%)	1 (50.0%)	1 (50.0%)
บัณฑิตวิทยาลัย	16	1 (6.3%)	0 (0.0%)	1 (100.0%)
สถาบันวิจัยพุทธกรรมศาสตร์	14	2 (14.3%)	2 (100.0%)	0 (0.0%)
รวม	993	123 (12.4%)	56 (45.5%)	67 (54.5%)

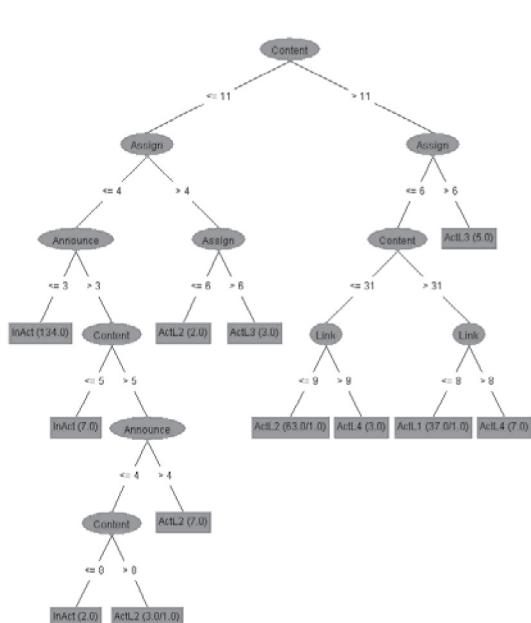
หมายเหตุ: * อาจารย์ที่ปฏิบัติหน้าที่ ณ ภาคการศึกษาที่ 1/2550

** รวมคณะสหเวชศาสตร์ (11 คน)

อาจารย์ที่มีสถานะ Active หมายถึง อาจารย์ที่มีวิชาที่ Active อย่างน้อย 1 วิชา

ผลข้อ 5 แสดงในรูปที่ 3 โดยโมเดลกู้ต้อง 94% ได้กู้ทั้งหมด 12 กู้ แต่กู้ที่ดีควรจะมีความน่าเชื่อถือได้ (confidence) สูง เช่น มากกว่า .90 (หรือ 90%) และมีข้อมูลที่เป็นจริงตามกู้ (support) มากกว่า .30 (หรือ 30%) ดังนั้นจากโมเดลที่ได้สามารถสรุปกฎสำหรับการจัดกลุ่มวิชาออกเป็น 5 ระดับได้ดังแสดงในตารางที่ 7 ผลนี้สามารถใช้เป็นกู้ในการให้อาจารย์ผู้สร้างวิชาทำนายได้ดังต่อไปนี้ เช่น หากมีการสร้างเพจนื้อหา (content) น้อยกว่าหรือเท่ากับ 11 เพจ และมีการมอบหมายงาน (assign) น้อยกว่าหรือเท่ากับ 4 และมีการประกาศ (announce) น้อยกว่าหรือเท่ากับ 3 ความน่าจะเป็นที่นิสิตจะเข้าใช้งานวิชาน้อย วิชานั้นก็จะจัดประเภทเป็น InActive

ตารางที่ 7 โมเดล (กฎ) ของการจำแนกกลุ่ม



กลุ่ม (จำนวนกฎ)	กฎที่มี Support สูงสุด	ถูก/ผิด	Supp	Conf
InAct (3)	Content < = 11 และ Assign < = 4 และ Announce < = 3	134/0	49	1.00
ActL1 (1)	Content > 31 และ Assign < = 6 และ Link < = 8	37/1	.14	.97
ActL2 (4)	11 < Content < = 31 และ Link < = 9	63/1	.23	.98
ActL3 (2)	Content > 11 และ Assign > 6	5/0	.02	1.00
ActL4 (2)	Content > 31 และ Assign < = 6 และ Link < = 8	7/0	.03	1.00

รูปที่ 3 Classification Tree

ผลข้อ 6 ได้ว่า แอ็ตทริบิวต์ที่มีนัยสำคัญมีเพียง 4 ตัว คือ enroll assign test และ link สมการพยากรณ์จาก Weka ที่ได้ คือ

$$\text{coAccess} = 31.34 * \text{enroll} + 73.84 * \text{assign} + -54.21 * \text{test} + 81.77 * \text{link} - 228.89$$

ได้ Correlation coefficient = 0.87

ให้ค่าสัมประสิทธิ์ของ test เป็นลบ ดังนั้น test จึงไม่เหมาะสมที่จะเป็นตัวพยากรณ์ จึงดำเนินการสร้างโมเดลใหม่โดยไม่เอา test ได้สมการดังนี้

$$\text{coAccess} = 31.10 * \text{enroll} + 73.41 * \text{assign} + 69.66 * \text{link} - 275.85$$

ได้ Correlation coefficient 0.87

โดยสรุปได้ว่า จำนวนนักเรียนที่ลงทะเบียนเรียน (enroll) จำนวนการบ้านที่ให้ส่งผ่านระบบ (assign) และจำนวนแหล่งวิชาการรายวิชา (link) เป็นแอ็ตทริบิวต์/ตัวแปรที่ส่งผลต่อจำนวนครั้งของการเข้าใช้งานของรายวิชา (coAccess)

อภิปรายผลการศึกษาและข้อเสนอแนะ

ผลการวิเคราะห์ที่สามารถนำไปใช้งานได้ดังนี้

1. ได้นำไปใช้ในการรายงานสถานะการใช้งานระบบ ATutor ในการจัดการเรียนการสอนในที่ประชุมผู้บริหารมหาวิทยาลัยเพื่อกระตุ้นให้ผู้บริหารของหน่วยงานกำหนดกลยุทธ์ สร้างแรงจูงใจให้อาจารย์ มีการใช้งานมากขึ้น เนื่องด้วยการมีการจัดการเรียนการสอนทางเครือข่ายคอมพิวเตอร์ (internet) เป็นสาระสำคัญของตัวบ่งชี้ที่ 2.2 (มีกระบวนการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ) ของการประกันคุณภาพการศึกษา มหาวิทยาลัยครินทร์วิโรฒ [18]

2. ผลลัพธ์จากการจัดกลุ่มวิชาในระดับต่างๆ สามารถนำไปใช้ในการให้การงานของการจัดการเรียนการสอนด้วย ATutor ของแต่ละวิชาได้

3. ทำให้ทราบว่าอาจารย์ที่ใช้งานในระดับ active เป็นครึ่งบ้าง และสามารถเชิญชวนให้เข้าร่วมโครงการพัฒนาอาจารย์เพื่อส่งเสริมการเรียนการสอนออนไลน์มีค่า. ได้เพื่อย้ายผลการใช้งานให้มากขึ้น

จากการทำเหมือนข้อมูลทำให้ผู้วิจัยซึ่งเป็นผู้ดูแลระบบเข้าใจโครงสร้างฐานข้อมูลของระบบอย่างลึกซึ้ง ทำให้มองเห็นแนวทางในการที่จะปรับแต่งระบบมากขึ้น เช่น การเชื่อมข้อมูลกับระบบสารสนเทศ หลักของมหาวิทยาลัย เพื่อให้สามารถนำข้อมูลเกี่ยวกับวิชาตามหลักสูตรของมหาวิทยาลัยเข้าสู่ระบบ ATutor เพื่อให้การระบุลักษณะวิชาถูกต้องมากขึ้น เช่น ประเภทของวิชาฯ เป็นวิชาตามหลักสูตรของมหาวิทยาลัย หรือวิชาตามอธิบาย วิชานี้เป็นของคณะใด สำหรับภาคการศึกษาใด เป็นต้น

เอกสารอ้างอิง

1. ATutor Learning Content Management System: Information: Available from: <http://www.atutor.ca/>. 11 August 2009.
2. Viyanon, W., Ruijrote, K., Raksakietisak, S., and Moulanont, S. 2003. An Implementation of an Open Source LCMS: Srinakharinwirot University Case Study. Proceedings of the 29th Congress on Science and Technology of Thailand. 20-22 October 2003. Golden Jubilee Convention Hall, Khon Kaen University. p. 9.
3. Raksakietisak, S. 2005. Test Analysis System in Open Source Learning Content Management System ATutor 1.4.3. Proceedings of the 1st National Conference on Computing and Information Technology. 24-25 May 2005. King Mongkut Institute of Technology North Bangkok. Bangkok. Thailand.
4. Cyber Education Center of Srinakharinwirot University. Available from: <http://cybered.swu.ac.th/>. 11 August 2009.
5. Han, J., and Kamber, M. 2006. Data Mining: Concepts and Techniques. 2nd Edition. The Morgan Kaufmann series in data management systems. Amsterdam. Elsevier.
6. Witten, I. H., and Frank, E. 2005. Data Mining : Practical Machine Learning Tools and Techniques. 2nd Edition. The Morgan Kaufmann Series in Data Management Systems. Amsterdam. Elsevier.

7. CRISP-DM-Home. Available from: <http://www.crisp-dm.org/>. 11 August 2009.
8. Romero, C., and Ventura, S. 2007. Educational Data Mining: A Survey from 1995 to 2005. *Expert Systems with Applications* 33(1): 135-146.
9. Romero, C., Ventura, S., and Garcia, E. 2008. Data Mining in Course Management Systems: Module case Study and Tutorial. *Computers and Education* 51(1): 368-384.
10. สุกัตรา วนกุ งานนิจ อาจอินทร์ และ นิตยา เกิดประสพ. 2005. ระบบการประเมินผล E-Learning โดยใช้วิธีการทำเหมืองข้อมูลการใช้เว็บ. The Joint Conference on Computer Science and Software Engineering: JCSSE2005. 17 November 2005. Dhamrong Buasri Auditorium Hall, Burapha University. Chonburi.
11. ข้อเปรียบเทียบ ATutor กับ Moodle. Available from: <http://www.radompon.com/webboard/index.php?topic=238.0>. 11 August 2009.
12. Data Mining Community's Top Resource. Available from: <http://www.kdnuggets.com/>. 11 August 2009.
13. Pentaho Commercial Open Source Business Intelligence: Pentaho. Available from: <http://www.pentaho.com/>. 11 August 2009.
14. The R Project for Statistical Computing. Available from: <http://www.r-project.org/>. 11 August 2009.
15. Weka 3-Data Mining with Open Source Machine Learning Software in Java. Available from: <http://www.cs.waikato.ac.nz/ml/weka/>. 11 August 2009.
16. SWU Course Web. Available from: <http://course.swu.ac.th>. 11 August 2009.
17. ศูนย์บริการข้อมูล มศว. Available from: <http://infoservice.swu.ac.th/>. 11 August 2009.
18. การประกันคุณภาพการศึกษา มหาวิทยาลัยคริสตินาวิโรฒ. Available from: [http://qa.swu.ac.th/Portals/50/files/manual_51\(new\).pdf](http://qa.swu.ac.th/Portals/50/files/manual_51(new).pdf). 11 August 2009.

ได้รับบทความวันที่ 16 สิงหาคม 2552
ยอมรับตีพิมพ์วันที่ 1 ตุลาคม 2552

