

การพัฒนาาระบบสนับสนุนการตัดสินใจในการเลือกสมัครในสาขาวิชา โดยใช้เทคนิคต้นไม้ตัดสินใจ

The Development of a Decision Support System to Apply for the Undergraduate Program Using Decision Tree Techniques

อนันต์ ปินะเต¹

Anan Pinate¹

Received: 6 December 2015; Accepted: 11 March 2016

บทคัดย่อ

การรับสมัครบุคคลเข้าศึกษาในระดับปริญญาตรี มหาวิทยาลัยมหาสารคาม มีกระบวนการในการคัดเลือกผู้สมัครที่มีคุณสมบัติตรงตามที่มหาวิทยาลัยกำหนด ซึ่งการแนะนำการศึกษาจึงเป็นกระบวนการที่สำคัญ ในการประชาสัมพันธ์ข้อมูลการสมัครให้กับผู้สมัครทราบ อาทิเช่น ข้อมูลการเลือกสาขาวิชา โดยทั่วไปผู้สมัครจะเลือกสาขาวิชาโดยยึดหลักตามความชอบ ความรู้สึกของตนเองเป็นส่วนใหญ่ โดยไม่ได้คำนึงถึงความรู้และทักษะด้านต่างๆ ของตนเอง ส่งผลให้เมื่อเข้ามาศึกษาในสาขาวิชานั้นแล้วเกิดปัญหาผลการเรียนที่ตกต่ำ ไม่ผ่านตามเกณฑ์ที่มหาวิทยาลัยกำหนด อันเป็นผลทำให้ต้องพ้นสภาพการเป็นนิสิตตั้งนั้นการประชาสัมพันธ์ข้อมูลเพื่อให้ผู้สมัครได้ใช้ประกอบการตัดสินใจในการเลือกสาขาวิชานั้นจึงเป็นสิ่งสำคัญในกระบวนการแนะนำการศึกษาต่อ งานวิจัยนี้จึงได้นำเสนอการใช้เทคนิคต้นไม้ตัดสินใจ C4.5 เพื่อค้นหากฎการตัดสินใจในการเลือกสาขาวิชาและนำกฎการตัดสินใจที่ได้มาพัฒนาเป็นระบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชา เพื่อให้ผู้ที่สมัครเข้าศึกษาในระดับปริญญาตรี ได้ทำการทดลองการเลือกสาขาวิชาก่อนการเลือกสมัครจริง ผู้สมัครจะได้ทราบถึงสาขาวิชาที่เหมาะสมกับความรู้และทักษะของผู้สมัครเองเมื่อเข้ามาศึกษาในสาขาวิชานั้น จากการวิจัยสามารถสร้างเป็นกฎการตัดสินใจได้จำนวนทั้งสิ้น 333 กฎจากทั้งหมด 51 สาขาวิชาที่ทำการทดลอง และสามารถสร้างระบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชาได้

คำสำคัญ : เทคนิคต้นไม้ตัดสินใจ กฎการตัดสินใจ

Abstract

When recruiting an individual to study at the undergraduate level in Mahasarakham University, there is a selection process to qualify applicants in line with university regulations. Thus, a study guide becomes an important source of information for the applicant It explains the disciplines and selection information. Generally, disciplines will be selected based on the applicants' choices, mostly without taking into consideration the individuals' knowledge and various skills. As a result, applicants will fail courses at the university if they do not pass the criteria set forth by the university. How to select disciplines becomes a really important process. This research presents the use of decision tree with C4.5 technique to search for decision rules in selecting fields of study and to improve decision rules to become a decision support system in order to help applicants be able to experiment with selecting fields of study before actual application. The applicants will realize which fields of study are compatible with their knowledge and skills. From the research, there were 333 decision rules from 51 experimental fields of study and decision.

Keywords : decision tree technique, decision rule

¹ นักวิชาการคอมพิวเตอร์ กองบริการการศึกษา มหาวิทยาลัยมหาสารคาม อำเภอกันทรวิชัย จังหวัดมหาสารคาม 44150

¹ Computer Technical Officer, Division of Academic Affair, Mahasarakham University, Kantharawichai District, MahaSarakhm 44150 Thailand.

บทนำ

มหาวิทยาลัยมหาสารคาม เป็นสถาบันอุดมศึกษาของรัฐ ที่มีการจัดการเรียนการสอนในระดับปริญญาตรี และบัณฑิตศึกษาระบบการการรับสมัครเพื่อคัดเลือกบุคคลเข้าศึกษาระดับปริญญาตรี เป็นกระบวนการที่สำคัญเพื่อคัดเลือกผู้ที่จะเข้าศึกษาให้ตรงตามคุณสมบัติของสาขาวิชาที่กำหนด ขั้นตอนการรับสมัครบุคคลเข้าศึกษาในระดับปริญญาตรี มีอยู่หลายขั้นตอนด้วยกันการแนะนำประชาสัมพันธ์ข้อมูลเกี่ยวกับการรับเข้าศึกษาเป็นขั้นตอนที่สำคัญอีกขั้นตอนหนึ่งรวมถึงการให้ข้อมูลประกอบการตัดสินใจในการเลือกเรียนในสาขาวิชาของผู้สมัครให้ตรงตามทักษะ และความรู้ของผู้สมัครเองถือเป็นข้อมูลที่สำคัญยิ่ง จากปัญหาที่สำคัญของการรับสมัครเข้าศึกษาในสาขาวิชาต่างๆ โดยส่วนใหญ่ผู้สมัครจะเลือกสมัครในสาขาวิชาตามความชอบ สมัครตามเพื่อน หรือผู้ปกครอง โดยไม่ได้คำนึงถึงความรู้ และทักษะที่ตนเองมี ส่งผลกระทบเมื่อเข้ามาศึกษาในสาขาวิชาดังกล่าวแล้วทำให้ผลการเรียนตกต่ำ เกิดปัญหาผลการเรียนไม่ผ่านตามเกณฑ์ที่มหาวิทยาลัยกำหนด การฟื้นฟูสภาพการเป็นนิสิต การขอย้ายสาขาวิชา และการผลิตบัณฑิตในสาขาวิชาไม่สูงเท่าที่ควร จากข้อมูลรายงานประจำปีของกองทะเบียนและประมวลผล มหาวิทยาลัยมหาสารคาม พบว่ามีจำนวนนิสิตที่ฟื้นฟูสภาพการเป็นนิสิตเพิ่มขึ้นทุกปีการศึกษา จากข้อมูลปีการศึกษา 2552 - 2555 มีนิสิตที่ฟื้นฟูสภาพร้อยละ 5.22, 10.30, 10.32, 10.48 ตามลำดับ² จากปัญหาส่งผลกระทบต่อผู้เรียน ได้แก่ เสียเวลาในการเรียน เสียค่าใช้จ่ายในการเรียน เป็นต้น จากปัญหาดังกล่าวมหาวิทยาลัยมีความต้องการที่จะแก้ไขปัญหาที่เกิดขึ้น ผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบสนับสนุนการตัดสินใจในการเลือกเรียนของผู้สมัคร โดยให้ผู้สมัครทำการทดลองความรู้ และทักษะที่ผู้สมัครมีอยู่เพื่อให้ทราบว่าจะสาขาวิชาที่เหมาะสมกับตนเอง โดยผู้วิจัยได้นำข้อมูลประวัตินิสิตที่สำเร็จการศึกษา และข้อมูลประวัติการสมัครเข้าเรียนของนิสิต มหาวิทยาลัยมหาสารคาม มาทำการทดลองโดยวิธีเทคนิคต้นไม้ตัดสินใจ อัลกอริทึม C4.5 (Decision Tree C4.5) เพื่อหาแบบจำลองและกฎการตัดสินใจเพื่อนำกฎที่ได้มาพัฒนาเป็นระบบสนับสนุนการตัดสินใจในการเลือกสมัครในสาขาวิชา

วัตถุประสงค์

เพื่อสร้างแบบจำลองและกฎการตัดสินใจด้วยวิธีต้นไม้ตัดสินใจ อัลกอริทึม C4.5 และพัฒนาระบบสนับสนุนการตัดสินใจในการเลือกสมัครเรียนในสาขาวิชาในระดับปริญญาตรี มหาวิทยาลัยมหาสารคาม

ทบทวนวรรณกรรม

งานวิจัยนี้ได้ศึกษาทฤษฎี และวิธีการดำเนินการทดลองด้วยวิธีต้นไม้ตัดสินใจ และการวัดประสิทธิภาพของผลการทดลองโดยมีรายละเอียด ดังนี้

ต้นไม้ตัดสินใจ

วิธีต้นไม้ตัดสินใจเป็นโมเดลที่มีรูปแบบที่ได้รับความนิยม โครงสร้างต้นไม้ตัดสินใจเป็นแบบลำดับชั้นโดยมีการตัดสินใจ ซึ่งประกอบด้วย โหนดที่ใช้ในการตัดสินใจ (Decision Node) และโหนดใบ (Terminal node) แต่ละโหนดตัดสินใจนั้นจะมีการสร้างฟังก์ชันที่เอาไว้สำหรับทดสอบทางเลือก $f_m(x)$ จากการป้อนข้อมูลเข้า (Input) จะทดสอบตามทางเลือกไปเรื่อยๆ ไปจนถึงโหนดใบ จะได้คำตอบในที่สุดรายละเอียดดัง (Figure 1)

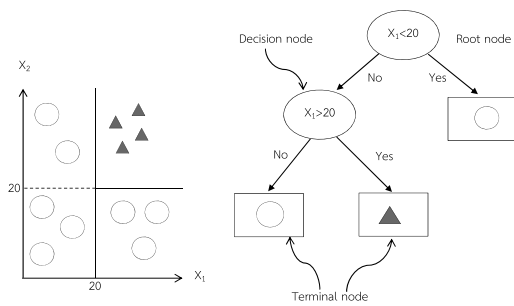


Figure 1 ตัวอย่างต้นไม้ที่สร้างจากชุดการสอน

อัลกอริทึม C4.5

เป็นอัลกอริทึม ที่พัฒนามาจากอัลกอริทึม ID3⁴ เป็นอัลกอริทึมในการจำแนกประเภทข้อมูลใช้หลักการสร้างต้นไม้โดยคัดเลือกคุณลักษณะที่สำคัญที่สุดมาเป็นโหนดราก (Root Node) โดยใช้ค่า Gain Ratio ที่สูงที่สุดเป็นโหนดราก และโหนดถัดไป และต้องหาค่า Entropy, Information Gain และ Split Information มีวิธีการการหาค่า ดังนี้

การหาค่า Entropy เป็นสมการที่ใช้ในการหาความสัมพันธ์ของข้อมูล (Entropy Measure) รายละเอียดดังสมการที่ 1

$$Entropy(s) = \sum_{i=1}^c -P_i \log_2 P_i \tag{1}$$

โดย s คือ Attribute หรือคุณลักษณะ ที่นำมาวัดค่า Entropy P_i คือ สัดส่วนของจำนวนสมาชิกในกลุ่ม i เทียบกับจำนวนสมาชิกทั้งหมดของกลุ่มตัวอย่าง

การหาค่า Information Gain เป็นสมการที่ใช้ในการหาค่าสารสนเทศก่อนนำไปใช้ในการหาค่ามาตรฐานอัตราส่วนเกน (Gain Ratio) รายละเอียดดังสมการที่ 2

$$Gain(S, A) = Entropy(s) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

โดย A คือ คุณลักษณะ A

$|S_v|$ คือ จำนวนสมาชิกของคุณลักษณะ A ที่มีค่า V

$|S|$ คือ จำนวนสมาชิกของกลุ่มตัวอย่าง

การหาค่า Split Information เป็นสมการที่ใช้ในการหาค่าสารสนเทศของการแบ่งแยก รายละเอียดดังสมการที่ 3

$$Split\ Information(S, A) = - \sum_{i=1}^n \frac{|S_i|}{S} \log_2 \frac{|S_i|}{S} \quad (3)$$

โดย S_i คือ จำนวนของจำนวนสมาชิกในกลุ่ม i

การหาค่า Gain Ratio เป็นสมการที่เพิ่มขึ้นจากอัลกอริทึม ID3 เพื่อลดความลำเอียงของข้อมูล รายละเอียดดังสมการ 4

$$Gain\ Ratio(S, A) = \frac{Gain(S, A)}{Split\ Information(S, A)} \quad (4)$$

การวัดประสิทธิภาพ

ในการวัดประสิทธิภาพของแบบจำลองในแต่ละข้อมูล ที่ทำการทดลองตามสาขาวิชา โดยใช้ค่าความถูกต้องของแบบจำลอง (Accuracy) ค่าความแม่นยำของแบบจำลอง (Precision) ค่าความระลึกของแบบจำลอง (Recall) และค่าความเหวี่ยงของแบบจำลอง (F-Measure) การวัดประสิทธิภาพของการจำแนกข้อมูลตามแนวคิดทางด้านการค้นคืนสารสนเทศ ซึ่งการวัดค่าประสิทธิภาพของแบบจำลองนั้นจะอาศัยตาราง Confusion Matrix ในการคำนวณค่ารายละเอียดดัง (Figure 2)

		Predicted Class	
		C_1	C_2
Actual	C_1	TP	FN
	C_2	FP	TN

Figure 2 Confusion Matrix

TP คือ จำนวนข้อมูลแบบจำลองจำแนกกลุ่ม C_1 และคำตอบเป็นกลุ่ม C_1

TN คือ จำนวนข้อมูลแบบจำลองจำแนกกลุ่ม C_2 และคำตอบเป็นกลุ่ม C_2

FP คือ จำนวนข้อมูลแบบจำลองจำแนกกลุ่ม C_1 และคำตอบเป็นกลุ่ม C_2

FN คือ จำนวนข้อมูลแบบจำลองจำแนกกลุ่ม C_2 และคำตอบเป็นกลุ่ม C_1

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F-Measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (8)$$

วัตถุประสงค์และวิธีการศึกษา

การวิจัยครั้งนี้เป็นการพัฒนาระบบสนับสนุนการตัดสินใจในการเลือกสมัครในสาขาวิชา โดยใช้เทคนิคต้นไม้ตัดสินใจ อัลกอริทึม C4.5 จากข้อมูลที่ใช้ในการทดลองประกอบด้วย ข้อมูลประวัติการศึกษาของผู้ที่สำเร็จการศึกษาในระดับปริญญาตรี มหาวิทยาลัยมหาสารคาม ปีการศึกษา 2549 – 2554 และข้อมูลประวัติการสมัครเข้าศึกษาในระดับปริญญาตรีโดยข้อมูลทั้งสองนี้ต้องเป็นข้อมูลที่มีความสัมพันธ์กันโดยมีข้อมูลที่ทำกรทดลอง 18,221 ชุดข้อมูลรวมทั้งหมด 51 สาขาวิชาการทดลองทำการเตรียมข้อมูลในแต่ละสาขาวิชาให้อยู่ในรูปแบบที่สามารถนำไปใช้กับโปรแกรม WEKA (Waikato Environment for Knowledge Analysis) ซึ่งประกอบด้วยการแบ่งข้อมูลออกเป็น 2 กลุ่ม คือ ข้อมูลทดสอบ (Data Testing) และข้อมูลเรียนรู้ (Data Training) โดยใช้หลักการแบ่งข้อมูลแบบ 10-fold Cross Validation จากนั้นทำการทดลองโดยใช้เทคนิคต้นไม้ตัดสินใจ อัลกอริทึม C4.5 (Decision Tree C4.5) เพื่อหาแบบจำลอง และทำการวัดประสิทธิภาพแบบจำลอง (Evaluation) จากแบบจำลองสามารถสร้างเป็นกฎการตัดสินใจ (Decision Rule) และนำกฎการตัดสินใจที่ได้มาพัฒนาเป็นระบบสนับสนุนการตัดสินใจในการเลือกสมัครในสาขาวิชา เพื่อให้ผู้สมัครได้ทำการทดสอบก่อนการเลือกสาขาวิชานั้น

จากการวิจัยครั้งนี้ ผู้วิจัยได้แสดงขั้นตอน และกรอบแนวคิดเพื่อให้ทราบกระบวนการในการวิจัยในครั้งนี้โดยมีรายละเอียดดัง (Figure 3)

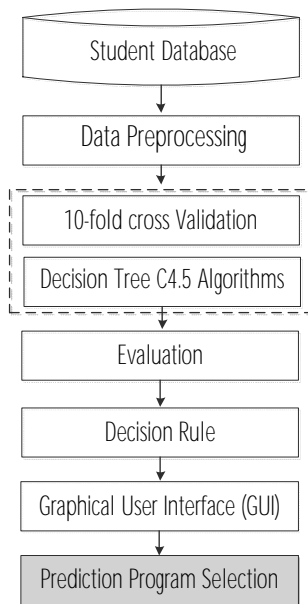


Figure 3 Conceptual framework

ข้อมูลที่ใช้ทำการทดลองนั้นจะแยกตามสาขาวิชาและการทดลองออกเป็นข้อมูลรายสาขาวิชา ข้อมูลในแต่ละสาขาวิชาประกอบด้วยข้อมูลตัวแปร โดยเลือกข้อมูลที่มีความสำคัญต่อการเลือกสมัครของสาขาวิชา ข้อมูลตัวแปรคุณลักษณะ (Attribute) ได้แก่ เพศ ผลการเรียนเฉลี่ยสะสม (GPAX) ผลการเรียนกลุ่มสาระการเรียนรู้ (GPA) ได้แก่ วิชาภาษาไทย วิชาคณิตศาสตร์ วิชาวิทยาศาสตร์ วิชาสังคมศึกษา วิชาสุขศึกษา วิชาศิลปะ วิชาการงานอาชีพ และวิชาภาษาต่างประเทศ ตัวแปรคำตอบหรือตัวแปรคลาส (Class) คือ ผลการเรียนนิสิตที่สำเร็จการศึกษาในระดับปริญญาตรีจากมหาวิทยาลัยมหาสารคาม รายละเอียดตัวแปรคุณลักษณะ (Attribute) และตัวแปรคำตอบ (Class) มีรายละเอียดดัง (Table 1) และตัวอย่างข้อมูลที่ใช้ในการทดลองในแต่ละสาขาวิชา มีรายละเอียดดัง (Table 2)

Table 1 Variable

Variable	Explanation
Sex	Sex
GPAX	Grade point average
GPA1	Grade point average Thai Language
GPA2	Grade point average Mathematics
GPA3	Grade point average Science
GPA4	Grade point average Social studies
GPA5	Grade point average Health education
GPA6	Grade point average Art
GPA7	Grade point average Home working
GPA8	Grade point average English
GPA	Grade point average in Mahasarakham University

Table 2 Data sample

Sex	GPAX	GPA1	GPA2	GPA3	GPA4	GPA5	GPA6	GPA7	GPA8	GPA
Female	3.35	3.50	3.00	3.26	3.68	3.75	3.50	3.62	3.32	3.52
Male	2.54	2.00	2.00	3.05	4.00	4.00	2.05	4.00	2.00	2.56
Male	2.40	2.55	1.54	2.60	2.52	3.57	3.36	2.50	2.33	2.35
Female	3.29	3.40	2.60	2.70	3.00	3.90	3.20	3.40	2.40	3.10
Female	3.00	3.35	1.70	3.16	3.17	3.40	3.50	3.45	3.31	2.74
Female	3.92	4.00	3.66	4.00	4.00	4.00	4.00	3.97	3.97	3.56
Male	3.51	3.71	3.05	3.42	3.94	4.00	4.00	3.50	3.60	3.20
Female	2.98	2.20	1.90	1.80	3.34	3.10	3.70	3.40	3.19	2.23
Male	3.71	4.00	3.32	3.56	4.00	4.00	3.90	4.00	3.76	3.64

ในการแทนค่าข้อมูลในการทดลอง เพื่อทำการทดลองในข้อมูลแต่ละสาขาวิชา เพื่อให้การวิเคราะห์ข้อมูลในการทดลองนั้น เพื่อให้ถูกต้อง และมีความแม่นยำในการทดลอง ข้อมูลต้องอยู่ในรูปแบบที่คอมพิวเตอร์เข้าใจถึงความหมายของข้อมูลนั้นเสียก่อน การวิจัยนี้ได้มีการแทนค่าข้อมูลให้กับข้อมูล ซึ่งประกอบด้วยข้อมูลผลการเรียนเฉลี่ยสะสม (GPAX) และผลการเรียนกลุ่มสาระการเรียนรู้ (GPA) ซึ่งข้อมูลทั้งสองส่วนเป็นข้อมูลคะแนนแบบมีทศนิยม ซึ่งก่อนทำการทดลองกับข้อมูลต้องทำการแทนค่าเพื่อลดการกระจายของข้อมูลซึ่งเป็นทศนิยมแบบต่อเนื่อง (Binning data) โดยผู้วิจัยได้แทนค่าข้อมูลโดยมีการแบ่งช่วงคะแนนผลการเรียนออกเป็นสองส่วน ได้แก่ ส่วนข้อมูลที่ตัวแปรคุณลักษณะ (Attribute) มีการแทน

ค่าช่วงคะแนนผลการเรียนออกเป็น 3 ระดับ คือ ช่วงคะแนนผลการเรียนระหว่าง 0.00 – 2.50 อยู่ในระดับต่ำ (Low) แทนค่า L, ช่วงคะแนนผลการเรียนระหว่าง 2.51 – 3.25 อยู่ในระดับกลาง (Medium) แทนค่า M และช่วงคะแนนผลการเรียนระหว่าง 3.26 – 4.00 อยู่ในระดับสูง (High) แทนค่า H และส่วนข้อมูลที่เป็นคำตอบ (Class) มีการแทนค่าช่วงคะแนนออกเป็น 3 ระดับ คือ ช่วงคะแนนผลการเรียน (GPA) ระหว่าง 2.00 – 2.74 เป็นระดับพอใช้ แทนค่า Fair, ช่วงคะแนนผลการเรียนระหว่าง 2.75 – 3.24 เป็นระดับดี แทนค่า Good และช่วงคะแนนผลการเรียนระหว่าง 3.25 – 4.00 เป็นระดับดีเยี่ยม แทนค่า Excellent ตัวอย่างการแทนค่าข้อมูลในการทดลอง ในข้อมูลแต่ละสาขาวิชาดัง (Table 3)

Table 3 Substitution information

Sex	GPAX	GPA1	GPA2	GPA3	GPA4	GPA5	GPA6	GPA7	GPA8	GPA
Female	H	H	M	H	H	H	H	H	H	Excellent
Male	M	L	L	M	H	H	L	H	L	Fair
Male	L	M	L	M	M	H	H	L	L	Fair
Female	H	H	M	M	M	H	M	H	L	Good
Female	M	H	L	M	M	H	H	H	H	Fair
Female	H	H	H	H	H	H	H	H	H	Excellent
Male	H	H	M	H	H	H	H	H	H	Good
Female	M	L	L	L	H	M	H	H	M	Fair
Male	H	H	H	H	H	H	H	H	H	Excellent

เมื่อแทนค่าให้กับข้อมูลทุกสาขาวิชาที่จะทำการทดลองแล้ว ทำการวิเคราะห์ข้อมูลโดยใช้โปรแกรม WEKA (Waikato Environment for Knowledge Analysis) การทดลองโดยทำการทดลองในข้อมูลที่ละสาขาวิชา รายละเอียดดัง (Figure 4)

จากนั้นทำการวิเคราะห์ข้อมูลโดยการเลือกวิธีการจำแนก (Classify) และเลือกวิธีต้นไม้ตัดสินใจ อัลกอริทึม C4.5 (J48) โปรแกรมจะทำการวิเคราะห์ข้อมูลทดลองรายละเอียดดัง (Figure 5)

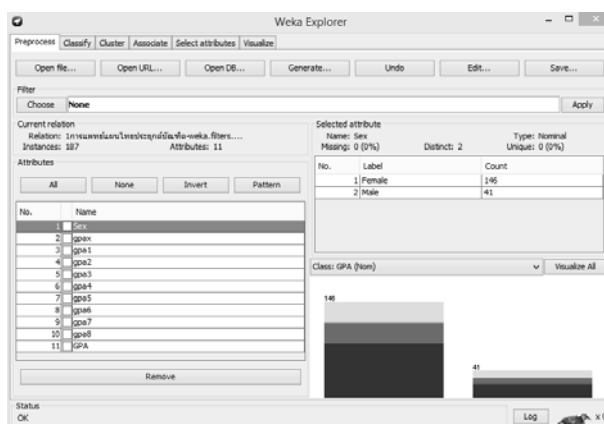


Figure 4 Program information WEKA

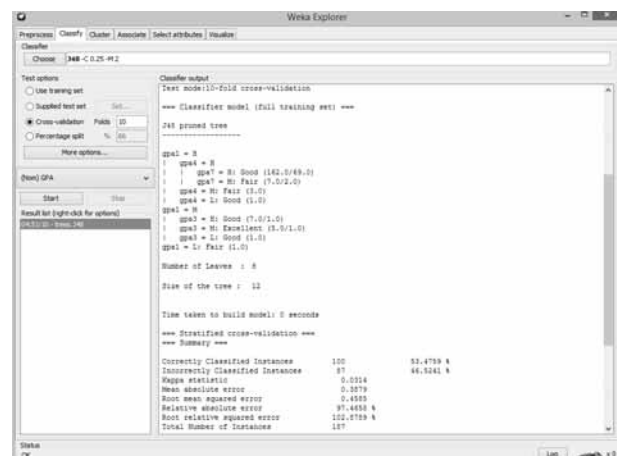


Figure 5 Analysis result

ผลการวิจัย

จากการทดลองโดยการวิเคราะห์ข้อมูลจากโปรแกรม WEKA โดยทำการทดลองข้อมูลในแต่ละสาขาวิชา ด้วยวิธีต้นไม้ตัดสินใจ (Decision tree) เพื่อหาแบบจำลองที่มีประสิทธิภาพ เพื่อสร้างเป็นกฎการตัดสินใจ (Decision Rule)

และนำกฎที่ได้ไปพัฒนาเป็นระบบสนับสนุนการตัดสินใจในการเลือกสมัครในสาขาวิชา ผลจากการทดลองข้อมูลแต่ละสาขาวิชา ได้ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความเหวี่ยง (F-Measure) รายละเอียดดัง (Table 4)

Table 4 Results

Major Name	Decision Tree (C45)				Rule Number
	Accuracy	Precision	Recall	F-Measure	
Bachelor of Applied Thai Traditional Medicine	81.28	0.833	0.813	0.818	10
Pharmaceutical Care	97.17	0.955	0.972	0.962	3
Bachelor of Nursing Science	80.06	0.818	0.801	0.800	5
Bachelor of Public Health	82.73	0.780	0.827	0.886	5
Bachelor of Science Program in Nutrition and Food Safety Management	79.11	0.822	0.791	0.798	7
Bachelor of Science Chemistry	78.06	0.746	0.781	0.740	4
Bachelor of Science Biology	71.08	0.748	0.711	0.720	7
Bachelor of Science Mathematic	77.70	0.709	0.777	0.727	5
Bachelor of Science Program in Food Technology	70.28	0.607	0.703	0.631	5
Bachelor of Science Program in Biotechnology	84.61	0.818	0.846	0.816	6
Bachelor of Science Program in Agricultural Technology	77.94	0.753	0.769	0.759	10
Bachelor of Science Program in Food Product Development	80.88	0.654	0.809	0.723	5
Bachelor of Science Program in Animal Science	77.14	0.709	0.761	0.711	5
Bachelor of Science Environmental Technology	71.71	0.721	0.717	0.717	8
Bachelor of Science Environmental and Resource Management	77.00	0.777	0.770	0.749	11
Bachelor of Arts Information Science	72.39	0.742	0.724	0.728	9
Bachelor of Science Information Technology	73.82	0.671	0.738	0.660	5
Bachelor of Science Computer Science	79.03	0.705	0.790	0.795	6
Bachelor of Communication Arts	78.09	0.684	0.781	0.706	8
Bachelor of Engineering	81.58	0.699	0.816	0.749	6
Bachelor of Architecture	71.00	0.685	0.710	0.779	7
Bachelor of Creative Arts	78.20	0.784	0.782	0.778	11
Bachelor of Accountancy	84.67	0.835	0.847	0.811	8
Bachelor of Marketing	83.57	0.767	0.836	0.764	9
Bachelor of Management	73.45	0.672	0.735	0.666	6
Bachelor of Business Computing	78.82	0.725	0.788	0.691	9
Bachelor of Financial Management	73.13	0.735	0.731	0.714	8
Bachelor of Human Resource Management	74.60	0.712	0.746	0.747	5
Bachelor of Business Economics	72.99	0.712	0.730	0.715	7
Bachelor of Tourism and Hotel Management	74.96	0.740	0.750	0.716	7
Bachelor of Music	71.68	0.703	0.717	0.696	7

Major Name	Decision Tree (C45)				Rule Number
	Accuracy	Precision	Recall	F-Measure	
Politics and Government	77.77	0.790	0.778	0.769	9
International Relations	77.47	0.740	0.755	0.745	6
Bachelor of Political Science	76.62	0.792	0.766	0.753	9
Bachelor of Arts Program in Thai	88.21	0.825	0.842	0.836	7
Bachelor of Arts Program in English	79.70	0.831	0.797	0.797	4
Bachelor of Arts Program in Chinese	70.00	0.615	0.700	0.648	7
Bachelor of Arts Program in Japanese	74.53	0.724	0.745	0.722	6
Bachelor of Arts Program in History	74.19	0.692	0.742	0.710	4
Bachelor of Arts Program in Community Development	88.85	0.844	0.849	0.845	6
Bachelor of Education Program in General Science	77.66	0.765	0.777	0.748	6
Bachelor of Education Program in Mathematics	75.00	0.800	0.750	0.744	5
Bachelor of Education Program in Social Studies	82.85	0.804	0.829	0.815	6
Bachelor of Education Program in English	84.22	0.824	0.842	0.797	6
Bachelor of Education Program in Thai Language	74.23	0.737	0.742	0.710	7
Bachelor of Education Program in Early Childhood Education	77.00	0.593	0.770	0.670	4
Bachelor of Education Program in Educational Technology	78.41	0.762	0.784	0.772	6
Bachelor of Science Program in Psychology	88.66	0.841	0.847	0.842	4
Bachelor of Science Program in Sport Science	78.73	0.796	0.787	0.763	4
Bachelor of Fine and Applied Arts Program in Visual Arts	80.00	0.793	0.800	0.785	5
Bachelor of Laws Program	81.49	0.740	0.815	0.744	8

จากผลการทดลองในแต่ละสาขาวิชา ด้วยวิธีต้นไม้ตัดสินใจ (Decision tree) อัลกอริทึม C4.5 สามารถนำผลการวิเคราะห์ข้อมูลในแต่ละสาขาวิชามาสร้างเป็นกฎการตัดสินใจ (Decision rule) ที่ได้จากแบบจำลอง เพื่อพัฒนาเป็นระบบ

สนับสนุนการตัดสินใจในการเลือกสาขาวิชา ผลการทดลองสามารถสร้างเป็นต้นไม้ตัดสินใจในแต่ละสาขาวิชา ตัวอย่างต้นไม้ตัดสินใจสาขาวิชาการแพทย์แผนไทยประยุกต์บัณฑิตมีรายละเอียดดัง (Figure 5)

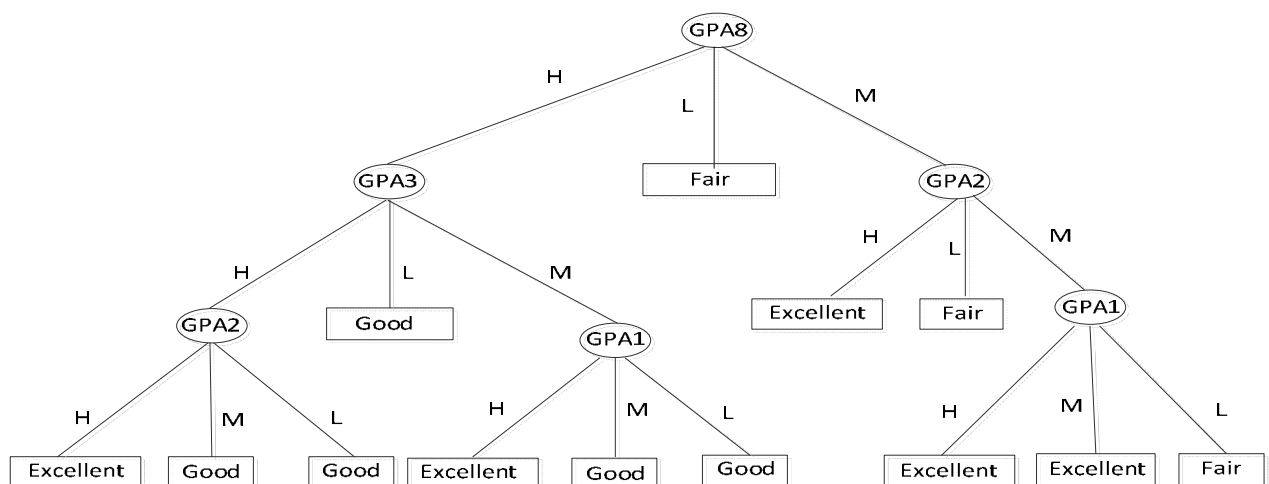


Figure 5 Example Tree Bachelor of Applied Thai Traditional Medicine

จากต้นไม้ตัดสินใจตัวอย่างการทดลองสาขาวิชาการแพทย์แผนไทยประยุกต์บัณฑิต สามารถสร้างเป็นกฎการตัดสินใจได้จำนวน 10 กฎรายละเอียดของแต่ละกฎมีดังนี้

Rule1 IF GPA8=H AND GPA3=H AND GPA2=H THEN GPA=Excellent

Rule2 IF GPA8=H AND GPA3=H AND GPA2=(M OR L) THEN GPA=Good

Rule3 IF GPA8=H AND GPA3=M AND GPA1=H THEN GPA=Excellent

Rule4 IF GPA8=H AND GPA3=M AND GPA1=(M OR L) THEN GPA=Good

Rule5 IF GPA8=H AND GPA3=L THEN GPA=Good

Rule6 IF GPA8=M AND GPA2=H THEN GPA=Excellent

Rule7 IF GPA8=M AND GPA2=L THEN GPA=Fair

Rule8 IF GPA8=M AND GPA2=M AND GPA1=H THEN GPA=Excellent

Rule9 IF GPA8=M AND GPA2=M AND GPA1=(M OR L) THEN GPA=Fair

Rule10 IF GPA8=L THEN GPA=Fair

จากการทดลองข้อมูลในแต่ละสาขาวิชาทั้งหมดจำนวน 51 สาขาวิชา สามารถสร้างเป็นกฎการตัดสินใจได้ทั้งสิ้น 333 กฎการตัดสินใจ สามารถนำกฎที่ได้มาพัฒนาเป็นระบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชาก่อนการสมัครเลือกในสาขาวิชานั้น รายละเอียดดัง (Figure 6)

Figure 6 Home program

เมื่อผู้ใช้กรอกข้อมูลเพื่อทำการทดสอบผ่านระบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชาแล้วระบบจะแสดงผลการทดสอบทุกสาขาวิชา และแสดงผลระดับผลการเรียนเมื่อเข้ามาศึกษาในสาขาวิชานั้น รายละเอียดดัง (Figure 7)

Figure 7 Test Results

อภิปรายผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองและกฎการตัดสินใจด้วยวิธีต้นไม้ตัดสินใจ อัลกอริทึม C4.5 และนำกฎการตัดสินใจที่ได้จากการวิเคราะห์ข้อมูลมาพัฒนาระบบสนับสนุนการตัดสินใจในการเลือกสมัครเรียนในสาขาวิชาในระดับปริญญาตรี มหาวิทยาลัยมหาสารคาม จากการทดลองพบว่าสามารถสร้างเป็นกฎการตัดสินใจได้ทั้งหมด 333 กฎการตัดสินใจจากข้อมูลในการทดลอง 51 สาขาวิชา

ข้อเสนอแนะ

งานวิจัยนี้ได้ทำการวิเคราะห์ปัจจัยเฉพาะผลการเรียนก่อนเข้าเรียนในมหาวิทยาลัย การเรียนในระดับมหาวิทยาลัยปัจจัยด้านอื่นๆ อาจจะมีผลต่อผลการเรียนของนิสิต อาทิเช่น การเอาใจใส่ในการเรียน การมีระเบียบวินัยในการเรียน เป็นต้น และการวิจัยนี้ไม่สามารถทดลองกับข้อมูลได้ทุกสาขาวิชาที่เปิดรับเข้าศึกษาอันเนื่องมาจากจำนวนข้อมูลที่ไม่เพียงพอในการทดลอง ในอนาคตผู้วิจัยมีแนวคิดที่จะทำการทดลองให้ครบทุกสาขาวิชาที่มหาวิทยาลัยรับเข้าศึกษา

กิตติกรรมประกาศ

โครงการวิจัยนี้ได้รับการสนับสนุนจากเงินทุนอุดหนุนการวิจัยจากงบประมาณเงินรายได้ ประจำปี 2559 มหาวิทยาลัยมหาสารคาม

เอกสารอ้างอิง

1. กองบริการการศึกษา มหาวิทยาลัยมหาสารคาม.ระเบียบ
การการสมัครคัดเลือกบุคคลเข้าศึกษาใน ระดับปริญญา
ตรี ระบบรับตรง ปี 2549-2554.
2. กองทะเบียนและประมวลผล มหาวิทยาลัย มหาสารคาม.
รายงานประจำปี 2552-2555
3. กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ, ธนาวินท์ รักธรรมานนท์.
การใช้เทคนิคด้าต้าไมนิ่งเพื่อพัฒนาคุณภาพการศึกษา
คณะวิศวกรรมศาสตร์; NECTEC Technical Journal,
Vol.3, No.11; 2001 July-October 2001.
4. ฉัตรเกล้า เจริญผล. เอกสารประกอบการสอนรายวิชา
Introduction to Data Mining 2013.
5. ชุตติมา อุดมะณี, ประสงค์ ปรานีตพลกรัง. การพัฒนาตัว
แบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์
สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาในระดับ
อุดมศึกษา; Journal of Information science and Tech-
nology, Vol.1, ISSUE 2, JUL-DEC 2010.
6. บุญเสริม กิจศิริกุล. รายงานวิจัยฉบับสมบูรณ์โครงการ
ย่อยที่ 7 อัลกอริทึมการทำเหมืองข้อมูล; ปี 2545.
7. ประเวศน์ วงษ์คำชัย, ใช้งาน AJAX และ PHP แบบมือ
อาชีพ; พิมพ์ครั้งที่1, กรุงเทพฯ : ไทย เจริญการพิมพ์,
2550.
8. อนันต์ ปิณะเต, ฉัตรเกล้า เจริญผล, แกมกาญจน์
สมประเสริฐศรี. การใช้เทคนิคเหมืองข้อมูลในการเลือก
กลุ่มสาขาวิชาที่เหมาะสมสำหรับการศึกษาต่อระดับปริญญาตรี;
วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม,
ปีที่ 33, ฉบับที่ 6, ประจำเดือนพฤศจิกายน - ธันวาคม
2557.
9. Han J, Kamber M, Data Mining Concepts and Tech-
niques; The Morgan Kaufmann Publishers, 2001.
10. Olson D, Shi Y. Introduction to Business Data Mining;
McGraw Hill International Edition, 2007.
11. Tomas B, Marcel J, Pavel K, Selection Representative
Data Sets; INTECH 2012.
12. Zdravko M, Ingrid R, An Introduction to the WEKA
Data Mining System.