

บทที่ 3

วิธีการดำเนินงานโครงการ

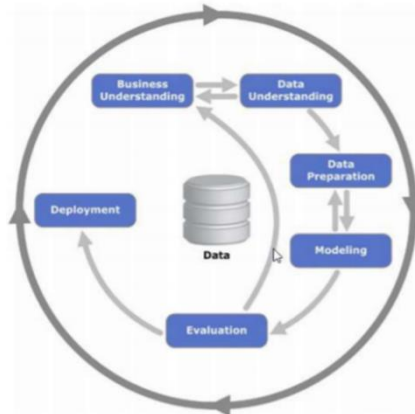
โครงการเรื่องนี้จัดทำได้นำเสนอเทคนิคการทำเหมืองข้อมูล (Data mining) ด้วยวิธีการจำแนกข้อมูล (Data classification) รูปแบบของแผนภูมิต้นไม้ (Decision tree) เพื่อวิเคราะห์หาข้อมูลสารสนเทศที่เกิดขึ้นจากข้อมูลอาจารย์และนักศึกษา คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ นำรูปแบบที่ให้ผลพยากรณ์ที่แม่นยำที่สุดมาเพื่อวางแผนการเพิ่มอัตราบุคลากรสายวิชาการของคณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ เมื่อเสร็จสิ้นจากกระบวนการวิเคราะห์ข้อมูลแล้วจะเป็นการออกแบบเว็บไซต์ และออกแบบรูปแบบการแสดงผลและบทสรุปจากวิธีการดำเนินงาน

3.1 การวิเคราะห์ข้อมูลด้วย ซีอาร์ไอเอสพี-ดีเอ็ม (CRISP-DM)

3.2 การออกแบบเว็บไซต์

3.3 บทสรุป

3.1 การวิเคราะห์ข้อมูลด้วย CRISP-DM



ภาพที่ 3.1 กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

ที่มา : sajeegm301.blogspot.com (ม.ป.ป.)

กระบวนการวิเคราะห์ข้อมูล (Cross Industry Standard Process for Data Mining หรือ (CRISP-DM) พัฒนาขึ้นในปี ค.ศ. 1936 โดยความร่วมมือของ 3 บริษัทคือ Daimler Chrysler, SPSS และ NCR ซึ่งกระบวนการวิเคราะห์ข้อมูล CRISP-UM. จะเป็นดังรูปแบบในรูปภาพที่ 3.2 แต่ละขั้นตอนในรูปจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือ ขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกระบวนการนั้น ๆ ประกอบด้วย 6 ขั้นตอนหลัก ดังนี้

3.1.1 รู้จักและเข้าใจในธุรกิจ (Business Understanding) เป็นขั้นตอนแรกของกระบวนการที่มุ่งเน้นไปที่การทำความเข้าใจกระบวนการทางธุรกิจโดยรวม คณะผู้จัดทำได้ทำความเข้าใจกับปัญหาให้อยู่ในรูปของการวิเคราะห์ข้อมูลทางดาด้า ไม่นิ่ง และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลของข้อมูลจากสำนักส่งเสริมวิชาการและงานทะเบียน กองบริหารงานบุคคล สำนักงานประกันคุณภาพการศึกษา และกองพัฒนานักศึกษา โดยการวิเคราะห์ข้อมูลในประเด็นนี้ คือ ได้แก่ข้อมูล อาจารย์ ผู้รับผิดชอบหลักสูตร อาจารย์ผู้สอนจำนวนอาจารย์ จำนวนภาระการสอนในหลักสูตร จำนวนภาระการสอนเฉลี่ยต่อคนอาจารย์เกษียณในปีถัดไป ซึ่งเก็บอยู่ในฐานข้อมูลบุคลากร และข้อมูลอีกส่วนเป็นข้อมูลนักศึกษาประกอบด้วย จำนวนนักศึกษาคงอยู่ จำนวนผู้สมัคร จำนวนรับสมัครนักศึกษา สัดส่วนนักศึกษาต่ออาจารย์ อัตราการแข่งขัน โดยมีการบันทึกข้อมูลในระบบทะเบียนกลางในส่วนงานรับสมัครนักศึกษา ของมหาวิทยาลัยราชชมงคลล้านนา เชียงใหม่ นำไปสู่การวางแผนประชาสัมพันธ์ปีการศึกษาต่อไปในอนาคต จำนวนทั้งหมด 46 รายการ ทำให้ไม่สามารถทำความเข้าใจกับข้อมูลที่มีความซับซ้อนได้อย่างรวดเร็ว

3.1.2 จัดเก็บและรวบรวมข้อมูลให้ครบ (Data Understanding) ขั้นตอนการจัดเก็บและรวบรวมข้อมูลจากกองบริหารงานบุคคล สำนักส่งเสริมวิชาการและงานทะเบียน และสำนักงานประกันคุณภาพการศึกษา ตลอดจนการพิจารณาตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลที่ได้รับ โดยพิจารณาข้อมูลทั้งหมดว่าส่วนใดของ ข้อมูลที่จำเป็นต่อการวิเคราะห์ข้อมูลให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้คณะผู้จัดทำ ทำการรวบรวมข้อมูลเพื่อตรวจสอบรายละเอียดข้อมูล ได้แก่ อาจารย์ผู้รับผิดชอบหลักสูตร อาจารย์ผู้สอน จำนวนอาจารย์ จำนวนภาระการสอนในหลักสูตร จำนวนภาระการสอนเฉลี่ยต่อคนอาจารย์เกษียณในปีถัดไป ที่ได้ข้อมูลมาจากกองบริหารงานบุคคล จำนวนสมัครของนักศึกษา 4 ปี และเทียบโอนในรอบโควตา จำนวนสมัครของนักศึกษา 4 ปี และเทียบโอนในรอบรับตรง ของปีการศึกษา 2558 ถึง 2563 จำนวนสมัคร TCAS ในปีการศึกษา 2561 ถึง 2563 จำนวนรับนักศึกษา และ

จำนวนนักศึกษาคงอยู่ ที่ได้มาจากสำนักส่งเสริมวิชาการและงานทะเบียน มี 46 รายการ
ประกอบไปด้วย 14 แอดทริบิวส์ ประกอบด้วย ปีการศึกษา หลักสูตร วิชาเอก อาจารย์
ผู้รับผิดชอบหลักสูตร อาจารย์ผู้สอน จำนวนอาจารย์ จำนวนภาระการสอนในหลักสูตร/จำนวน
Section จำนวนภาระการสอนเฉลี่ยต่อคน สัดส่วนนักศึกษาต่ออาจารย์ อาจารย์เกษียณในปี
ถัดไป อัตราการแข่งขัน จำนวนนักศึกษาคงอยู่ จำนวนสมัครจำนวนรับสมัคร

ตารางที่ 3.1 ข้อมูลนักศึกษาและอาจารย์

ปีการศึกษา	A	B	C	D	H	K	N	Q	T	U	V
	ปีการศึกษา	หลักสูตร	วิชาเอก	อาจารย์ผู้รับผิดชอบหลักสูตร	จำนวนการลงทะเบียนก่อน	สัดส่วนนักศึกษาต่ออาจารย์	อาจารย์เกษียณในปีถัดไป	อัตราการแข่งขัน	จ.น.นักศึกษาอยู่	จ.น.สมัคร	จ.น.รับ
3	2558	บธ.ม.การจัดการ	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	5	16.86	38.27	0	4.09	842	1,430	350
4	2558	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	6.50	14.38	0	0.98	230	120	122
5	2558	บธ.ม.การตลาด	บธ.ม.การตลาด - การตลาด	5	11.38	60.54	1	2.71	787	941	347
6	2558	บธ.ม.ภาษาอังกฤษธุรกิจ	บธ.ม.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	5	25.17	83.83	0	2.30	503	378	164
7	2558	บธ.ม.ระบบสารสนเทศทางคอมพิวเตอร์	บธ.ม.ระบบสารสนเทศทางคอมพิวเตอร์	5	14.33	53.83	0	1.76	646	490	278
8	2558	บช.ม.การบัญชี	บช.ม.การบัญชี	5	11.12	58.59	0	5.18	996	1,910	369
9	2558	ศส.ม.การท่องเที่ยวและการโรงแรม	ศส.ม.การท่องเที่ยวและการโรงแรม	5	21.10	42.20	0	3.16	422	458	145
10	2558	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	10.00	31.60	0	3.65	158	299	82
11	2559	บธ.ม.การจัดการ	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	5	13.45	36.82	0	4.22	810	1,098	260
12	2559	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	5.88	14.81	0	1.39	237	124	89
13	2559	บธ.ม.การตลาด	บธ.ม.การตลาด - การตลาด	5	6.17	47.00	1	2.44	846	1,126	462
14	2559	บธ.ม.ภาษาอังกฤษธุรกิจ	บธ.ม.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	5	20.20	89.60	0	2.34	448	314	134
15	2559	บธ.ม.ระบบสารสนเทศทางคอมพิวเตอร์	บธ.ม.ระบบสารสนเทศทางคอมพิวเตอร์	5	16.42	50.92	0	1.89	611	469	248
16	2559	บช.ม.การบัญชี	บช.ม.การบัญชี	5	15.50	60.69	1	6.35	971	1,645	259
17	2559	ศส.ม.การท่องเที่ยวและการโรงแรม	ศส.ม.การท่องเที่ยวและการโรงแรม	5	11.23	31.00	1	2.83	403	354	125
18	2559	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	3.18	9.55	0	3.90	210	335	86
19	2560	บธ.ม.บริหารธุรกิจ	บธ.ม.บริหารธุรกิจ	15	17.13	61.65	0	1.61	1911	1687	1050
20	2560	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.80	43.80	0	0.74	219	80	108
21	2560	บธ.ม.ระบบสารสนเทศทางธุรกิจ	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	13.93	37.00	0	0.94	518	260	276
22	2560	บช.ม.การบัญชี	บช.ม.การบัญชี	5	13.67	55.60	3	3.61	834	975	270
23	2560	ศส.ม.การท่องเที่ยวและการบริการ	ศส.ม.การท่องเที่ยวและการบริการ	5	17.83	29.50	0	2.91	354	471	162
24	2560	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	4.64	15.07	0	1.76	211	190	108
25	2561	บธ.ม.บริหารธุรกิจ	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	15	14.00	55.76	0	2.92	1840	1710	585
26	2561	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	20.60	38.40	0	1.23	192	74	60
27	2561	บธ.ม.ระบบสารสนเทศทางธุรกิจ	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	15.79	49.36	0	1.49	691	358	240
28	2561	บช.ม.การบัญชี	บช.ม.การบัญชี	5	11.00	62.08	0	5.86	807	879	150
29	2561	ศส.ม.การท่องเที่ยวและการบริการ	ศส.ม.การท่องเที่ยวและการบริการ	5	13.64	30.18	0	4.90	332	588	120
30	2561	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	8.22	22.89	0	2.97	206	178	60
31	2562	บธ.ม.บริหารธุรกิจ	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	6	16.20	65.10	3	2.89	651	521	180
32	2562	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	19.20	33.00	0	3.40	165	102	30
33	2562	บธ.ม.บริหารธุรกิจ	บธ.ม.การตลาด - การตลาด	3	16.38	79.75	0	4.53	638	815	180
34	2562	บธ.ม.บริหารธุรกิจ	บธ.ม.การตลาด - การจัดการธุรกิจค้าปลีก	3	25.33	56.67	0	1.38	170	83	60
35	2562	บธ.ม.บริหารธุรกิจ	บธ.ม.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	3	11.86	49.00	0	1.86	343	223	120
36	2562	บธ.ม.ระบบสารสนเทศทางธุรกิจ	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	13.27	29.33	0	0	440	0	0
37	2562	บช.ม.การบัญชี	บช.ม.การบัญชี	5	8.86	45.57	0	9.14	638	823	90
38	2562	ศส.ม.การท่องเที่ยวและการบริการ	ศส.ม.การท่องเที่ยวและการบริการ	5	11.91	26.64	0	9.72	293	875	90
39	2562	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	11.22	25.89	1	4.65	233	279	60

3.1.3 เตรียมข้อมูลให้พร้อมใช้งาน (Data Preparation) ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาและเลือกไว้ ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้ โดยการทำให้เป็นข้อมูลที่ถูกต้อง (Data cleaning) มักใช้เวลาค่อนข้างมาก โดยมีขั้นตอนดังนี้

3.1.3.1 ทำการคัดเลือกข้อมูล (Data Selection) คือการคัดเลือกข้อมูลที่เหมาะสม เพื่อนำมาใช้ในการวิเคราะห์ข้อมูล

คณะผู้จัดทำได้ทำการคัดเลือกข้อมูลและทำการ Data Cleaning ข้อมูลเฉพาะข้อมูลที่จำเป็นในการวิเคราะห์ภาพรวม จำนวน 9 แอตทริบิวต์ ประกอบด้วย ปีการศึกษา หลักสูตร วิชาเอก อาจารย์ผู้รับผิดชอบหลักสูตร จำนวนภาระการสอนในหลักสูตร จำนวนภาระการสอนเฉลี่ยต่อคน สัดส่วนนักศึกษาต่ออาจารย์ อาจารย์เกษียณ อัตราการแข่งขัน

ตารางที่ 3.2 ข้อมูลที่ผ่านการ Data Cleaning

2	ปีการศึกษา	หลักสูตร	วิชาเอก	อาจารย์ผู้รับผิดชอบหลักสูตร	จำนวนการการสอนสัปดาห์	สัดส่วนนักศึกษาต่ออาจารย์	อาจารย์เกษียณในปีถัดไป	อัตราการแข่งขัน
3	2558	บธ.บ.การจัดการ	บธ.บ.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	5	16.86	38.27	0	4.09
4	2558	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	6.50	14.38	0	0.98
5	2558	บธ.บ.การตลาด	บธ.บ.การตลาด - การตลาด	5	11.38	60.54	1	2.71
6	2558	บธ.บ.ภาษาอังกฤษธุรกิจ	บธ.บ.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	5	25.17	83.83	0	2.30
7	2558	บธ.บ.ระบบสารสนเทศทางคอมพิวเตอร์	บธ.บ.ระบบสารสนเทศทางคอมพิวเตอร์	5	14.33	53.83	0	1.76
8	2558	บข.บ.การบัญชี	บข.บ.การบัญชี	5	11.12	58.59	0	5.18
9	2558	ศส.บ.การท่องเที่ยวและการโรงแรม	ศส.บ.การท่องเที่ยวและการโรงแรม	5	21.10	42.20	0	3.16
10	2558	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	10.00	31.60	0	3.65
11	2559	บธ.บ.การจัดการ	บธ.บ.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	5	13.45	36.82	0	4.22
12	2559	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	5.88	14.81	0	1.39
13	2559	บธ.บ.การตลาด	บธ.บ.การตลาด - การตลาด	5	6.17	47.00	1	2.44
14	2559	บธ.บ.ภาษาอังกฤษธุรกิจ	บธ.บ.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	5	20.20	89.60	0	2.34
15	2559	บธ.บ.ระบบสารสนเทศทางคอมพิวเตอร์	บธ.บ.ระบบสารสนเทศทางคอมพิวเตอร์	5	16.42	50.92	0	1.89
16	2559	บข.บ.การบัญชี	บข.บ.การบัญชี	5	15.50	60.69	1	6.35
17	2559	ศส.บ.การท่องเที่ยวและการโรงแรม	ศส.บ.การท่องเที่ยวและการโรงแรม	5	11.23	31.00	1	2.83
18	2559	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	3.18	9.55	0	3.90
19	2560	บธ.บ.บริหารธุรกิจ	บธ.บ.บริหารธุรกิจ	15	17.13	61.65	0	1.61
20	2560	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.80	43.80	0	0.74
21	2560	บธ.บ.ระบบสารสนเทศทางธุรกิจ	บธ.บ.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	13.93	37.00	0	0.94
22	2560	บข.บ.การบัญชี	บข.บ.การบัญชี	5	13.67	55.60	3	3.61
23	2560	ศส.บ.การท่องเที่ยวและการบริการ	ศส.บ.การท่องเที่ยวและการบริการ	5	17.83	29.50	0	2.91
24	2560	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	4.64	15.07	0	1.76
25	2561	บธ.บ.บริหารธุรกิจ	บธ.บ.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	15	14.00	55.76	0	2.92
26	2561	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	20.60	38.40	0	1.23
27	2561	บธ.บ.ระบบสารสนเทศทางธุรกิจ	บธ.บ.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	15.79	49.36	0	1.49
28	2561	บข.บ.การบัญชี	บข.บ.การบัญชี	5	11.00	62.08	0	5.86
29	2561	ศส.บ.การท่องเที่ยวและการบริการ	ศส.บ.การท่องเที่ยวและการบริการ	5	13.64	30.18	0	4.90
30	2561	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	8.22	22.89	0	2.97
31	2562	บธ.บ.บริหารธุรกิจ	บธ.บ.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	6	16.20	65.10	3	2.89
32	2562	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.บ.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	19.20	33.00	0	3.40
33	2562	บธ.บ.บริหารธุรกิจ	บธ.บ.การตลาด - การตลาด	3	16.38	79.75	0	4.53
34	2562	บธ.บ.บริหารธุรกิจ	บธ.บ.การตลาด - การจัดการธุรกิจค้าปลีก	3	25.33	56.67	0	1.38
35	2562	บธ.บ.บริหารธุรกิจ	บธ.บ.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	3	11.86	49.00	0	1.86
36	2562	บธ.บ.ระบบสารสนเทศทางธุรกิจ	บธ.บ.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	13.27	29.33	0	0
37	2562	บข.บ.การบัญชี	บข.บ.การบัญชี	5	8.86	45.57	0	9.14
38	2562	ศส.บ.การท่องเที่ยวและการบริการ	ศส.บ.การท่องเที่ยวและการบริการ	5	11.91	26.64	0	9.72
39	2562	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศส.บ.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	11.22	25.89	1	4.65

3.1.3.2 ทำการกลั่นกรองข้อมูล (Data Cleaning) คือการทำความสะอาดข้อมูล เป็นกระบวนการตรวจสอบและการแก้ไข หรือลบรายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูล ซึ่งเป็นหลักสำคัญของฐานข้อมูล ทางคณะผู้จัดทำได้ดำเนินการ ดังนี้

ข้อมูลวิชาเอก มหาวิทยาลัยราชภัฏวชิรเวศน์ขอนแก่น คณะผู้จัดทำได้ทำการตรวจสอบข้อมูล จำนวนนักศึกษาในปีการศึกษา 2560 และ 2561 เป็นข้อมูลที่ต้องนำมาจัดกลุ่มให้มาเป็น หลักสูตรเดียวกัน เนื่องจากยังมีข้อมูลที่ไม่ชัดเจน จึงรวมวิชาเอก บธ.บ.การจัดการ บธ.บ. การตลาด และ บธ.บ.ภาษาอังกฤษธุรกิจ นำมาจัดกลุ่มใหม่เป็น บธ.บ.บริหารธุรกิจ

ตารางที่ 3.3 การตรวจสอบข้อมูลที่เป็นหลักสูตรเดียวกัน

18	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	6	8	14	260	18.57	36.93	0	2.23	517	721	324
19	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	0	5	84	16.80	43.80	0	0.74	219	80	108
20	บธ.ม.การตลาด - การตลาด	3	4	7	166	23.71	86.57	0	1.38	606	586	426
21	บธ.ม.การตลาด - การจัดการธุรกิจค้าปลีก	3	0	3	6	2.00	0.00	0	1.70		143	84
22	บธ.ม.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	3	4	7	99	14.14	42.43	0	1.10	297	237	216
23	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	8	14	195	13.93	37.00	0	0.94	518	260	276
24	บธ.ม.การบัญชี	5	10	15	205	13.67	55.60	3	3.61	834	975	270
25	ศต.ม.การท่องเที่ยวและการบริการ	5	7	12	214	17.83	29.50	0	2.91	354	471	162
26	ศต.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	9	14	65	4.64	15.07	0	1.76	211	190	108
27	บธ.ม.บริหารธุรกิจ			0		#DIV/0!	#DIV/0!		#DIV/0!	491		
28	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	6	8	14	202	14.43	18.71	0	3.28	262	688	210
29	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	0	5	103	20.60	38.40	0	1.23	192	74	60
30	บธ.ม.การตลาด - การตลาด	3	4	7	128	18.29	52.14	0	2.69	365	565	210
31	บธ.ม.การตลาด - การจัดการธุรกิจค้าปลีก	3	2	5	37	7.40	0.00	0	2.95		221	75
32	บธ.ม.บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	3	4	7	95	13.57	29.14	0	2.62	204	236	90
33	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	8	14	221	15.79	49.36	0	1.49	691	358	240
34	บธ.ม.การบัญชี	5	8	13	143	11.00	62.08	0	5.86	807	879	150
35	ศต.ม.การท่องเที่ยวและการบริการ	5	6	11	150	13.64	30.18	0	4.90	332	588	120
36	ศต.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	4	9	74	8.22	22.89	0	2.97	206	178	60
37	บธ.ม.บริหารธุรกิจ			0		#DIV/0!	#DIV/0!		#DIV/0!	1,009		

ตารางที่ 3.4 การจัดกลุ่มข้อมูล

20	2560	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.80	43.80	0	0.74	219	80	108
21	2560	บธ.ม.ระบบสารสนเทศทางธุรกิจ	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	13.93	37.00	0	0.94	518	260	276
22	2560	บธ.ม.การบัญชี	บธ.ม.การบัญชี	5	13.67	55.60	3	3.61	834	975	270
23	2560	ศต.ม.การท่องเที่ยวและการบริการ	ศต.ม.การท่องเที่ยวและการบริการ	5	17.83	29.50	0	2.91	354	471	162
24	2560	ศต.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	ศต.ม.ภาษาอังกฤษเพื่อการสื่อสารสากล	5	4.64	15.07	0	1.76	211	190	108
25	2561	บธ.ม.บริหารธุรกิจ	บธ.ม.การจัดการ - การจัดการทั่วไป,การจัดการสำนักงาน	15	14.00	55.76	0	2.92	1840	1710	585
26	2561	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	บธ.ม.การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	20.60	38.40	0	1.23	192	74	60
27	2561	บธ.ม.ระบบสารสนเทศทางธุรกิจ	บธ.ม.ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	15.79	49.36	0	1.49	691	358	240
28	2561	บธ.ม.การบัญชี	บธ.ม.การบัญชี	5	11.00	62.08	0	5.86	807	879	150
29	2561	ศต.ม.การท่องเที่ยวและการบริการ	ศต.ม.การท่องเที่ยวและการบริการ	5	13.64	30.18	0	4.90	332	588	120

จากข้อมูลที่ได้นำมาจัดหมวดหมู่แล้วนำมาวิเคราะห์เพื่อให้ได้เกรดรวมสุดท้ายแล้วนำเกรดมาวิเคราะห์ วางแผนการเพิ่มอัตราบุคลากรสายวิชาการของคณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ ทั้งหมด 100% โดยแบ่งเป็นในแต่ละหมวดดังนี้

1) อาจารย์ผู้รับผิดชอบหลักสูตร	16.67%
2) จำนวนภาระการสอนในหลักสูตร/จำนวน Section	16.67%
3) จำนวนภาระการสอนเฉลี่ยต่อคน	16.67%
4) สัดส่วนนักศึกษาต่ออาจารย์	16.67%
5) อาจารย์เกษียณในปีถัดไป	16.67%
6) อัตราการแข่งขัน	16.67%

ตารางที่ 3.5 การให้คะแนนในแต่ละหมวด

2	วิชาเอก	อาจารย์ที่ปรึกษา	คะแนน	จำนวนการการสอนในหลักสูตร/จำนวน Section	เกณฑ์	คะแนน	จำนวนการการสอนเฉลี่ยต่อคน	เกณฑ์	คะแนน	สัดส่วนนักศึกษาต่ออาจารย์	เกณฑ์	คะแนน	อาจารย์ที่ปรึกษาไปศึกษาไป	เกณฑ์	คะแนน	อาจารย์ที่ปรึกษา	เกณฑ์	คะแนน	ZScore	T-Score	เกรด	
3	สอน การจัดการ - การจัดการทั่วไป การจัดการสำนักงาน	5	16.67	371	2	16.67	16.86	2	11.11	38.27	1	5.56	0	1	8.33	4.09	2	11.11	69.44	0.56605	55.6605	B
4	สอน การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.67	104	1	8.33	6.50	1	5.56	14.38	1	5.56	0	1	8.33	0.98	1	5.56	50.00	-1.62995	39.7049	D
5	สอน การตลาด - การตลาด	5	16.67	148	1	8.33	11.38	1	5.56	60.54	2	11.11	1	2	16.67	2.71	1	5.56	63.89	-0.6138	49.38621	C
6	สอน บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	5	16.67	151	1	8.33	25.17	3	16.67	83.83	3	16.67	0	1	8.33	2.30	1	5.56	72.22	0.879765	58.79765	B
7	สอน ระบบสารสนเทศทางคอมพิวเตอร์	5	16.67	172	2	16.67	14.33	2	11.11	53.83	2	11.11	0	1	8.33	1.76	1	5.56	69.44	0.56605	55.6605	B
8	สอน การบัญชี	5	16.67	189	2	16.67	11.12	1	5.56	58.59	2	11.11	0	1	8.33	5.18	3	16.67	75.00	1.193479	61.93479	B
9	สอน การท่องเที่ยวและการโรงแรม	5	16.67	211	2	16.67	21.10	3	16.67	42.20	1	5.56	0	1	8.33	3.16	2	11.11	75.00	1.193479	61.93479	B
10	สอน ภาษาอังกฤษเพื่อการสื่อสารสากล	5	16.67	50	1	8.33	10.00	1	5.56	31.60	1	5.56	0	1	8.33	3.65	2	11.11	55.56	-1.00252	39.97478	D
11	สอน การจัดการ - การจัดการทั่วไป การจัดการสำนักงาน	5	16.67	296	2	16.67	13.45	1	5.56	36.82	1	5.56	0	1	8.33	4.22	2	11.11	63.89	-0.6138	49.38621	C
12	สอน การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.67	94	1	8.33	5.88	1	5.56	14.81	1	5.56	0	1	8.33	1.39	1	5.56	50.00	-1.62995	39.7049	D
13	สอน การตลาด - การตลาด	5	16.67	111	1	8.33	6.17	1	5.56	47.00	2	11.11	1	2	16.67	2.44	1	5.56	63.89	-0.6138	49.38621	C
14	สอน บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	5	16.67	101	1	8.33	20.20	3	16.67	89.60	3	16.67	0	1	8.33	2.34	1	5.56	72.22	0.879765	58.79765	B
15	สอน ระบบสารสนเทศทางคอมพิวเตอร์	5	16.67	197	2	16.67	16.42	2	11.11	50.92	2	11.11	0	1	8.33	1.89	1	5.56	69.44	0.56605	55.6605	B
16	สอน การบัญชี	5	16.67	248	2	16.67	15.50	2	11.11	60.69	2	11.11	1	2	16.67	6.35	3	16.67	88.89	2.762051	77.62051	A
17	สอน การท่องเที่ยวและการโรงแรม	5	16.67	146	1	8.33	11.23	1	5.56	31.00	1	5.56	1	2	16.67	2.83	1	5.56	58.33	-0.68881	43.11192	D
18	สอน ภาษาอังกฤษเพื่อการสื่อสารสากล	5	16.67	70	1	8.33	3.18	1	5.56	9.55	1	5.56	0	1	8.33	3.90	2	11.11	55.56	-1.00252	39.97478	D
19	สอน บริหารธุรกิจ	15	16.67	531	2	16.67	17.13	2	11.11	61.65	2	11.11	0	1	8.33	1.61	1	5.56	69.44	0.56605	55.6605	B
20	สอน การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.67	84	1	8.33	16.80	2	11.11	43.80	1	5.56	0	1	8.33	0.74	1	5.56	55.56	-1.00252	39.97478	D
21	สอน ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	16.67	195	2	16.67	13.93	1	5.56	37.00	1	5.56	0	1	8.33	0.94	1	5.56	58.33	-0.68881	43.11192	D
22	สอน การบัญชี	5	16.67	205	2	16.67	13.67	1	5.56	55.60	2	11.11	3	2	16.67	3.61	2	11.11	77.78	1.507193	65.07193	B
23	สอน การท่องเที่ยวและการบริการ	5	16.67	214	2	16.67	17.83	2	11.11	29.50	1	5.56	0	1	8.33	2.91	1	5.56	63.89	-0.6138	49.38621	C
24	สอน ภาษาอังกฤษเพื่อการสื่อสารสากล	5	16.67	65	1	8.33	4.64	1	5.56	15.07	1	5.56	0	1	8.33	1.76	1	5.56	50.00	-1.62995	39.7049	D
25	สอน บริหารธุรกิจ	15	16.67	462	2	16.67	14.00	2	11.11	55.76	2	11.11	0	1	8.33	2.92	1	5.56	69.44	0.56605	55.6605	B
26	สอน การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.67	103	1	8.33	20.60	3	16.67	38.40	1	5.56	0	1	8.33	1.23	1	5.56	61.11	-0.37509	46.24907	C
27	สอน ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	16.67	221	2	16.67	15.79	2	11.11	49.36	2	11.11	0	1	8.33	1.49	1	5.56	69.44	0.56605	55.6605	B
28	สอน การบัญชี	5	16.67	143	1	8.33	11.00	1	5.56	62.08	2	11.11	0	1	8.33	5.86	3	16.67	66.67	0.252336	52.52336	C
29	สอน การท่องเที่ยวและการบริการ	5	16.67	150	1	8.33	13.64	1	5.56	30.18	1	5.56	0	1	8.33	4.90	2	11.11	55.56	-1.00252	39.97478	D
30	สอน ภาษาอังกฤษเพื่อการสื่อสารสากล	5	16.67	74	1	8.33	8.22	1	5.56	22.89	1	5.56	0	1	8.33	2.97	1	5.56	50.00	-1.62995	39.7049	D
31	สอน การจัดการ - การจัดการทั่วไป การจัดการสำนักงาน	6	16.67	162	2	16.67	16.20	2	11.11	65.10	3	16.67	3	2	16.67	2.89	1	5.56	83.33	2.134622	1.34622	A
32	สอน การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.67	96	1	8.33	19.20	3	16.67	33.00	1	5.56	0	1	8.33	3.40	2	11.11	66.67	0.252336	52.52336	C
33	สอน การตลาด - การตลาด	3	16.67	131	1	8.33	16.38	2	11.11	79.75	3	16.67	0	1	8.33	4.53	2	11.11	72.22	0.879765	58.79765	B
34	สอน การตลาด - การจัดการธุรกิจฉบับเล็ก	3	16.67	76	1	8.33	25.33	3	16.67	56.67	2	11.11	0	1	8.33	1.38	1	5.56	66.67	0.252336	52.52336	C
35	สอน บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	3	16.67	83	1	8.33	11.86	1	5.56	49.00	2	11.11	0	1	8.33	1.86	1	5.56	55.56	-1.00252	39.97478	D
36	สอน ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	16.67	199	2	16.67	13.27	1	5.56	29.33	1	5.56	0	1	8.33	0	1	5.56	58.33	-0.68881	43.11192	D
37	สอน การบัญชี	5	16.67	124	1	8.33	8.86	1	5.56	45.57	2	11.11	0	1	8.33	9.14	5	16.67	66.67	0.252336	52.52336	C
38	สอน การท่องเที่ยวและการบริการ	5	16.67	131	1	8.33	11.91	1	5.56	26.64	1	5.56	0	1	8.33	9.72	3	16.67	61.11	-0.37509	46.24907	C
39	สอน ภาษาอังกฤษเพื่อการสื่อสารสากล	5	16.67	101	1	8.33	11.22	1	5.56	25.89	1	5.56	1	2	16.67	4.65	2	11.11	63.89	-0.6138	49.38621	C
40	สอน การจัดการ - การจัดการทั่วไป การจัดการสำนักงาน	6	16.67	152	1	8.33	15.20	2	11.11	60.80	2	11.11	0	1	8.33	4.39	2	11.11	66.67	0.252336	52.52336	C
41	สอน การจัดการธุรกิจระหว่างประเทศ(นานาชาติ)	5	16.67	74	1	8.33	12.33	1	5.56	29.83	1	5.56	0	1	8.33	3.03	2	11.11	55.56	-1.00252	39.97478	D
42	สอน การตลาด - การตลาด	3	16.67	115	1	8.33	14.38	2	11.11	73.00	3	16.67	1	2	16.67	2.97	1	5.56	75.00	1.193479	61.93479	B
43	สอน การตลาด - การจัดการธุรกิจฉบับเล็ก	3	16.67	94	1	8.33	31.33	3	16.67	71.33	3	16.67	0	1	8.33	0.55	1	5.56	72.22	0.879765	58.79765	B
44	สอน บริหารธุรกิจ - ภาษาอังกฤษธุรกิจ	3	16.67	85	1	8.33	12.14	1	5.56	47.57	2	11.11	0	1	8.33	2.42	1	5.56	55.56	-1.00252	39.97478	D
45	สอน ระบบสารสนเทศทางธุรกิจ - การจัดการสารสนเทศ	6	16.67	165	2	16.67	11.00	1	5.56	24.73	1	5.56	0	1	8.33	0.77	1	5.56	58.33	-0.68881	43.11192	D
46	สอน การบัญชี	5	16.67	109	1	8.33	8.38	1	5.56	40.85	1	5.56	0	1	8.33	4.46	2	11.11	55.56	-1.00252	39.97478	D
47	สอน การท่องเที่ยวและการบริการ	5	16.67	141	1	8.33	14.10	2	11.11	28.40	1	5.56	0	1	8.33	5.99	3	16.67	66.67	0.252336	52.52336	C

3.1.3.3 การแปลงรูปแบบของข้อมูล (Data Transformation) เป็นขั้นตอนการแปลงข้อมูล โดยการจัดกลุ่มข้อมูลที่เป็นตัวเลขให้อยู่ในกลุ่มต่าง ๆ ตามหลักทางสถิติและกลุ่มตัวอย่าง จากข้อมูลที่ใช้ในการทดลองผู้จัดทำได้ทำการแทนค่าให้กับข้อมูลก่อนการวิเคราะห์ทดลองดังต่อไปนี้

- 1) ข้อมูลคะแนนอาจารย์ผู้รับผิดชอบหลักสูตร (teacher)
 - คะแนน 16.67 แทนค่า full (อาจารย์ผู้รับผิดชอบหลักสูตรครบ)
- 2) ข้อมูลเกณฑ์จำนวนภาระการสอนในหลักสูตร/จำนวน Section (section)
 - เกณฑ์ 1 แทนค่า few (มีจำนวนภาระการสอนในหลักสูตรน้อย)
 - เกณฑ์ 2 แทนค่า many (มีจำนวนภาระการสอนในหลักสูตรเยอะ)
- 3) ข้อมูลเกณฑ์จำนวนภาระการสอนเฉลี่ยต่อคน (noSubject)
 - เกณฑ์ 1 แทนค่า few (มีจำนวนภาระการสอนเฉลี่ยน้อย)
 - เกณฑ์ 2 แทนค่า middle (มีจำนวนภาระการสอนเฉลี่ยปานกลาง)
 - เกณฑ์ 3 แทนค่า many (มีจำนวนภาระการสอนเฉลี่ยน้อยเยอะ)
- 4) ข้อมูลเกณฑ์สัดส่วนนักศึกษาต่ออาจารย์ (noStudent)
 - เกณฑ์ 1 แทนค่า rightAmount (มีสัดส่วนนักศึกษาต่ออาจารย์พอดี)
 - เกณฑ์ 2 แทนค่า alot (มีสัดส่วนนักศึกษาต่ออาจารย์เยอะ)
 - เกณฑ์ 3 แทนค่า somany (มีสัดส่วนนักศึกษาต่ออาจารย์เยอะมาก)
- 5) ข้อมูลเกณฑ์อาจารย์เกษียณในปีถัดไป (retire)
 - เกณฑ์ 1 แทนค่า doNotHave (ไม่มีอาจารย์เกษียณ)
 - เกณฑ์ 2 แทนค่า have (มีอาจารย์เกษียณ)
- 6) ข้อมูลเกณฑ์อัตราการแข่งขัน (competitive)
 - เกณฑ์ 1 แทนค่า low (มีอัตราการแข่งขันต่ำ)
 - เกณฑ์ 2 แทนค่า moderate (มีอัตราการแข่งขันปานกลาง)
 - เกณฑ์ 3 แทนค่า high (มีอัตราการแข่งขันสูง)

7) ข้อมูลเกรดที่ได้ทำการคำนวณแล้ว (grade)

เกรด A และ B แทนค่า add (ควรเพิ่มอาจารย์)

เกรด C และ D แทนค่า notAdd (ไม่ควรเพิ่มอาจารย์)

ตารางที่ 3.6 เปลี่ยนข้อมูลเป็นภาษาอังกฤษ

teacher	section	noSubject	noStudent	retire	competitive	grade
full	many	middle	rightAmount	doNotHave	moderate	add
full	few	few	rightAmount	doNotHave	low	notAdd
full	few	few	alot	have	low	notAdd
full	few	many	somany	doNotHave	low	add
full	many	middle	alot	doNotHave	low	add
full	many	few	alot	doNotHave	high	add
full	many	many	rightAmount	doNotHave	moderate	add
full	few	few	rightAmount	doNotHave	moderate	notAdd
full	many	few	rightAmount	doNotHave	moderate	notAdd
full	few	few	rightAmount	doNotHave	low	notAdd
full	few	few	alot	have	low	notAdd
full	few	many	somany	doNotHave	low	add
full	many	middle	alot	doNotHave	low	add
full	many	middle	alot	have	high	add
full	few	few	rightAmount	have	low	notAdd
full	few	few	rightAmount	doNotHave	moderate	notAdd
full	many	middle	alot	doNotHave	low	add
full	few	middle	rightAmount	doNotHave	low	notAdd
full	many	few	rightAmount	doNotHave	low	notAdd
full	many	few	alot	have	moderate	add
full	many	middle	rightAmount	doNotHave	low	notAdd
full	few	few	rightAmount	doNotHave	low	notAdd
full	many	middle	alot	doNotHave	low	add
full	few	many	rightAmount	doNotHave	low	notAdd
full	many	middle	alot	doNotHave	low	add
full	few	few	alot	doNotHave	high	notAdd
full	few	few	rightAmount	doNotHave	moderate	notAdd
full	few	few	rightAmount	doNotHave	low	notAdd
full	many	middle	somany	have	low	add

3.1.4 สร้างแบบจำลอง (Modeling) ขั้นตอนการสร้างตัวแบบทางคณิตศาสตร์ และสถิติ เพื่อการวิเคราะห์ข้อมูล โดยสามารถใช้เทคนิควิธีการต่าง ๆ อาทิ การจำแนก(Classification) ต้นไม้ตัดสินใจ (Decision Tree)

คณะผู้จัดทำวิเคราะห์ข้อมูลด้วยการทำเหมืองข้อมูลแบบ Decision Tree เพื่อใช้สนับสนุนการตัดสินใจการเพิ่มอัตราบุคลากรสายวิชาการ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ นำไปสู่การวางแผนประชาสัมพันธ์ต่อไปในอนาคต

3.1.5. การสร้างโมเดล Decision tree ด้วยการคำนวณการสร้างโมเดล Decision tree จะทำการคัดเลือกแอตทริบิวต์ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็นโหนดบนสุดของ tree (root node) หลังจากนั้นก็จะหาแอตทริบิวต์ถัดไปเรื่อย ๆ ในการหาความสัมพันธ์ของแอตทริบิวต์นี้ จะใช้ตัววัด ที่เรียกว่า Information Gain(IG) ค่านี้คำนวณได้จากสมการดังนี้

ตารางที่ 3.7 สมการการหาความสัมพันธ์ของแอตทริบิวต์

การคำนวณ	สมการ
Entropy	$\text{entropy}(c1) = -p(c1) \log p(c1)$ และ $p(c2)$ คือ ค่าความน่าจะเป็นของ $c1$
Information Gain	$\text{IG}(\text{parent}, \text{child}) = \text{entropy}(\text{parent}) - [p(c1) \times \text{entropy}(c1) + p(c2) \times \text{entropy}(C2) + \dots]$

การคำนวณค่าแต่ละแอตทริบิวต์เทียบกับคลาสเพื่อหาแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น root ของ decision tree กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็นควรรเพิ่ม และไม่ควรรเพิ่ม ดังนี้

1) คำนวณค่า IG ของแอตทริบิวต์ teacher จากข้อมูลสามารถคำนวณค่า IG ได้ ดังนี้

$$\begin{aligned} \text{entropy}(\text{parent}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.348 \times \log_2(0.348) + 0.652 \times \log_2(0.652)] \\ &= 0.932 \end{aligned}$$

$$\begin{aligned} \text{entropy}(\text{ผล} = \text{full}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.348 \times \log_2(0.348) + 0.652 \times \log_2(0.652)] \\ &= 0.932 \end{aligned}$$

$$\begin{aligned} \text{IG}(\text{parent}, \text{child}) &= \text{entropy}(\text{parent}) - [p(\text{ผล} = \text{full}) \times \text{entropy}(\text{ผล} = \text{full})] \\ &= 0.932 - [1 \times 0.932] \\ &= 0.932 - [0.932] \\ &= 0.932 - 0.932 \\ &= 0 \end{aligned}$$

2) คำนวณค่า IG ของแอตทริบิวต์ section จากข้อมูลสามารถคำนวณค่า IG ได้ ดังนี้

$$\begin{aligned} \text{entropy}(\text{parent}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.348 \times \log_2(0.348) + 0.652 \times \log_2(0.652)] \\ &= 0.932 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = few)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.167 \times \log_2 (0.167) + 0.833 \times \log_2 (0.833)] \\
 &= 0.650
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = many)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.688 \times \log_2 (0.688) + 0.313 \times \log_2 (0.313)] \\
 &= 0.896
 \end{aligned}$$

$$\begin{aligned}
 \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล = few}) \times \text{entropy(ผล = few)} \\
 &+ p(\text{ผล = many}) \times \text{entropy(ผล = many)}] \\
 &= 0.932 - [0.652 \times 0.650 + 0.348 \times 0.896] \\
 &= 0.932 - [0.424 + 0.312] \\
 &= 0.932 - 0.736 \\
 &= 0.197
 \end{aligned}$$

3) คำนวณค่า IG ของแอตทริบิวต์ noSubject จากข้อมูลสามารถคำนวณค่า IG ได้ ดังนี้

$$\begin{aligned}
 \text{entropy (parent)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.348 \times \log_2 (0.348) + 0.652 \times \log_2 (0.652)] \\
 &= 0.932
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = few)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.08 \times \log_2 (0.08) + 0.92 \times \log_2 (0.92)] \\
 &= 0.402
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = many)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.571 \times \log_2 (0.571) + 0.429 \times \log_2 (0.429)] \\
 &= 0.985
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = middle)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.714 \times \log_2 (0.714) + 0.286 \times \log_2 (0.286)] \\
 &= 0.863
 \end{aligned}$$

$$\begin{aligned}
 \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล} = \text{few}) \times \text{entropy}(\text{ผล} = \text{few}) \\
 + p(\text{ผล} = \text{many}) \times \text{entropy}(\text{ผล} = \text{many}) + p(\text{ผล} = \text{middle}) \times \text{entropy}(\text{ผล} = \text{middle})] \\
 &= 0.932 - [0.543 \times 0.402 + 0.152 \times 0.985 + 0.304 \\
 \times 0.863] \\
 &= 0.932 - [0.219 + 0.150 + 0.263] \\
 &= 0.932 - 0.631 \\
 &= 0.301
 \end{aligned}$$

4) คำนวณค่า IG ของแอตทริบิวต์ noStudent จากข้อมูลสามารถคำนวณค่า IG ได้ ดังนี้

$$\begin{aligned}
 \text{entropy (parent)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.348 \times \log_2 (0.348) + 0.652 \times \log_2 (0.652)] \\
 &= 0.932
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล} = \text{alot)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2 (0.5) + 0.5 \times \log_2 (0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล} = \text{rightAmount)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.083 \times \log_2 (0.083) + 0.917 \times \log_2 (0.917)] \\
 &= 0.414
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล} = \text{somany)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[1 \times \log_2 (1) + 0 \times \log_2 (0)] \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล} = \text{few}) \times \text{entropy}(\text{ผล} = \text{few}) \\
 + p(\text{ผล} = \text{rightAmount}) \times \text{entropy}(\text{ผล} = \text{rightAmount}) + p(\text{ผล} = \text{somany}) \times \text{entropy}(\text{ผล} = \\
 \text{somany})] \\
 &= 0.932 - [0.348 \times 1 + 0.522 \times 0.414 + 0.130 \times 0] \\
 &= 0.932 - [0.348 + 0.216 + 0]
 \end{aligned}$$

$$= 0.932 - 0.564$$

$$= 0.368$$

5) คำนวณค่า IG ของแอตทริบิวต์ retire จากข้อมูลสามารถคำนวณค่า IG ได้ ดังนี้

$$\begin{aligned} \text{entropy (parent)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.348 \times \log_2 (0.348) + 0.652 \times \log_2 (0.652)] \\ &= 0.932 \end{aligned}$$

$$\begin{aligned} \text{entropy (ผล = doNotHave)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.324 \times \log_2 (0.324) + 0.676 \times \log_2 (0.676)] \\ &= 0.909 \end{aligned}$$

$$\begin{aligned} \text{entropy (ผล = have)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.444 \times \log_2 (0.444) + 0.556 \times \log_2 (0.556)] \\ &= 0.991 \end{aligned}$$

$$\begin{aligned} \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล = doNotHave}) \times \text{entropy}(\text{ผล} \\ &= \text{doNotHave}) + p(\text{ผล = have}) \times \text{entropy}(\text{ผล = have})] \\ &= 0.932 - [0.804 \times 0.909 + 0.196 \times 0.991] \\ &= 0.932 - [0.731 + 0.194] \\ &= 0.932 - 0.925 \\ &= 0.007 \end{aligned}$$

6) คำนวณค่า IG ของแอตทริบิวต์ competitive จากข้อมูลสามารถคำนวณค่า IG ได้ ดังนี้

$$\begin{aligned} \text{entropy (parent)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.348 \times \log_2 (0.348) + 0.652 \times \log_2 (0.652)] \\ &= 0.932 \end{aligned}$$

$$\begin{aligned} \text{entropy (ผล = high)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.333 \times \log_2 (0.333) + 0.667 \times \log_2 (0.667)] \\ &= 0.918 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = low)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.385 \times \log_2 (0.385) + 0.615 \times \log_2 (0.615)] \\
 &= 0.961
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = moderate)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.286 \times \log_2 (0.286) + 0.714 \times \log_2 (0.714)] \\
 &= 0.863
 \end{aligned}$$

$$\begin{aligned}
 \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล = high}) \times \text{entropy(ผล = high)} + p(\text{ผล = low}) \times \text{entropy(ผล = low)} + p(\text{ผล = moderate}) \times \text{entropy(ผล = moderate)}] \\
 &= 0.932 - [0.130 \times 0.918 + 0.565 \times 0.961 + 0.304 \times 0.863] \\
 &= 0.932 - [0.120 + 0.543 + 0.263] \\
 &= 0.932 - 0.926 \\
 &= 0.006
 \end{aligned}$$

จากการคำนวณค่า IG ของทุกแอตทริบิวต์พบว่าค่า IG ของแอตทริบิวต์ สัดส่วนนักศึกษาต่ออาจารย์ (noStudent) มีค่ามากที่สุด (0.368) ดังนั้นจึงเลือกแอตทริบิวต์ สัดส่วนนักศึกษาต่ออาจารย์ขึ้นมาเป็นโหนด root และจะต้องทำการแตกกิ่งจาก โหนด root ออกไปจนข้อมูลในแต่ละโหนดมีคลาสคำตอบเดียวกัน และผู้วิเคราะห์จึงทำการสร้างโหนดในระดับถัดไป

การคำนวณค่าแต่ละแอตทริบิวต์ในระดับที่ 2 ต่อจากโหนด root เพื่อหาค่า IG ที่มากที่สุดของแอตทริบิวต์ dlot (มีสัดส่วนนักศึกษาต่ออาจารย์เยอะ) กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็นควรเพิ่มอาจารย์(add) และไม่ควรเพิ่มอาจารย์(notAdd) ดังนี้

1) คำนวณค่า IG ของแอตทริบิวต์ dlot ต่อ section ได้ดังนี้

$$\begin{aligned}
 \text{entropy (parent)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2 (0.5) + 0.5 \times \log_2 (0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = few)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= - [0 \times \log_2 (0) + 1 \times \log_2 (1)] \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = many)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= - [1 \times \log_2 (1) + 0 \times \log_2 (0)] \\
 &= 2
 \end{aligned}$$

$$\begin{aligned}
 \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล = few}) \times \text{entropy}(\text{ผล = few}) \\
 &+ p(\text{ผล = many}) \times \text{entropy}(\text{ผล = many})] \\
 &= 1 - [0.5 \times 0 + 0.5 \times 0] \\
 &= 1 - [0 + 0] \\
 &= 1 - 0 \\
 &= 1
 \end{aligned}$$

2) คำนวณค่า IG ของแอตทริบิวต์ alot ต่อ noSubject ได้ดังนี้

$$\begin{aligned}
 \text{entropy (parent)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2 (0.5) + 0.5 \times \log_2 (0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = few)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.25 \times \log_2 (0.25) + 0.75 \times \log_2 (0.75)] \\
 &= 0.811
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = many)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0 \times \log_2 (0) + 1 \times \log_2 (1)] \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy (ผล = middle)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.857 \times \log_2 (0.857) + 0.143 \times \log_2 (0.143)] \\
 &= 0.592
 \end{aligned}$$

$$\begin{aligned}
 IG(\text{parent, child}) &= \text{entropy}(\text{parent}) - [p(\text{ผล} = \text{few}) \times \text{entropy}(\text{ผล} = \text{few}) \\
 &+ p(\text{ผล} = \text{many}) \times \text{entropy}(\text{ผล} = \text{many}) + p(\text{ผล} = \text{middle}) \times \text{entropy}(\text{ผล} = \text{middle})] \\
 &= 1 - [0.5 \times 0.811 + 0.063 \times 0 + 0.438 \times 0.592] \\
 &= 1 - [0.406 + 0 + 0.259] \\
 &= 1 - 0.664 \\
 &= 0.336
 \end{aligned}$$

3) คำนวณค่า IG ของแอตทริบิวต์ alot ต่อ retire ได้ดังนี้

$$\begin{aligned}
 \text{entropy}(\text{parent}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2(0.5) + 0.5 \times \log_2(0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy}(\text{ผล} = \text{doNotHave}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2(0.5) + 0.5 \times \log_2(0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy}(\text{ผล} = \text{have}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2(0.5) + 0.5 \times \log_2(0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 IG(\text{parent, child}) &= \text{entropy}(\text{parent}) - [p(\text{ผล} = \text{doNotHave}) \times \text{entropy}(\text{ผล} \\
 &= \text{doNotHave}) + p(\text{ผล} = \text{have}) \times \text{entropy}(\text{ผล} = \text{have})] \\
 &= 1 - [0.75 \times 1 + 0.25 \times 1] \\
 &= 1 - [0.75 + 0.25] \\
 &= 1 - 1 \\
 &= 0
 \end{aligned}$$

4) คำนวณค่า IG ของแอตทริบิวต์ alot ต่อ competitive ได้ดังนี้

$$\begin{aligned}
 \text{entropy}(\text{parent}) &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\
 &= -[0.5 \times \log_2(0.5) + 0.5 \times \log_2(0.5)] \\
 &= 1
 \end{aligned}$$

$$\begin{aligned} \text{entropy (ผล = high)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.5 \times \log_2 (0.5) + 0.5 \times \log_2 (0.5)] \\ &= 1 \end{aligned}$$

$$\begin{aligned} \text{entropy (ผล = low)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.5 \times \log_2 (0.5) + 0.5 \times \log_2 (0.5)] \\ &= 1 \end{aligned}$$

$$\begin{aligned} \text{entropy (ผล = moderate)} &= -p(\text{add}) \times \log_2 p(\text{add}) + p(\text{notAdd}) \times \log_2 p(\text{notAdd}) \\ &= -[0.5 \times \log_2 (0.5) + 0.5 \times \log_2 (0.5)] \\ &= 1 \end{aligned}$$

$$\begin{aligned} \text{IG (parent, child)} &= \text{entropy (parent)} - [p(\text{ผล = high}) \times \text{entropy(ผล = high)} \\ &+ p(\text{ผล = low}) \times \text{entropy(ผล = low)} + p(\text{ผล = moderate}) \times \text{entropy(ผล = moderate)}] \\ &= 1 - [0.25 \times 1 + 0.625 \times 1 + 0.125 \times 1] \\ &= 1 - [0.25 + 0.625 + 0.125] \\ &= 1 - 1 \\ &= 0 \end{aligned}$$

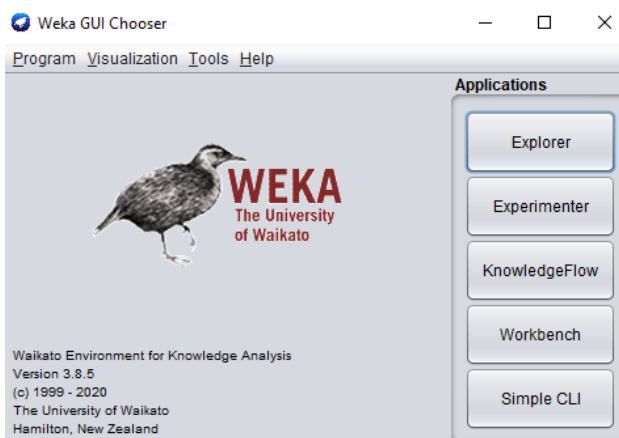
จากการคำนวณค่า IG ของแอดทรีวิวด์ dot (มีสัดส่วนนักศึกษาต่ออาจารย์เยอะ) ต่อจำนวนภาระการสอนในหลักสูตร จำนวนภาระการสอนเฉลี่ยต่อคน (noSubject) อาจารย์เกษียณในปีถัดไป (retire) และอัตราการแข่งขัน (competitive) ที่เป็นแอดทรีวิวด์สุดท้ายคณะผู้จัดทำพบว่าแอดทรีวิวด์ dot (มีสัดส่วนนักศึกษาต่ออาจารย์เยอะ) มีความสัมพันธ์กับจำนวนภาระการสอนในหลักสูตร/จำนวน Section มากที่สุดอยู่ที่ 1 และพบว่าข้อมูลในแต่ละโหนดมีคลาสคำตอบเดียวกันแล้ว คือ ผลลัพธ์เป็นควรเพิ่มอาจารย์ (add) และไม่ควรเพิ่มอาจารย์ (notAdd)

3.1.5 การประเมินผล (Evaluation) เป็นขั้นตอนก่อนนำผลลัพธ์ที่ได้จากขั้นตอนที่ 2 ไปใช้งานด้วยการวัดประสิทธิผลของผลลัพธ์ที่ได้กับวัตถุประสงค์ที่ตั้งไว้ในขั้นตอนแรก ว่ามีนัยสำคัญหรือความน่าเชื่อถือมากน้อยเพียงใด ด้วยการประเมินผลจากโปรแกรมว่าถูกต้องหรือไม่ คณะผู้จัดทำได้ทำการทดสอบโมเดล เพื่อวัดประสิทธิภาพที่ตรงกับความต้องการ ซึ่งการวัดประสิทธิภาพด้วยวิธี Self Consistency Test เหมาะสำหรับการใช้ในการทดสอบประสิทธิภาพ เพื่อดูแนวโน้มของโมเดลที่สร้างขึ้น และเมื่อนำข้อมูลมาทดสอบ (testing data) กับโปรแกรมที่ผู้วิเคราะห์

เลือกมาทดสอบกับข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดล Decision Tree จึงนำข้อมูลดังกล่าว มาทดสอบกับโปรแกรม Weka เวอร์ชัน 3.8.4

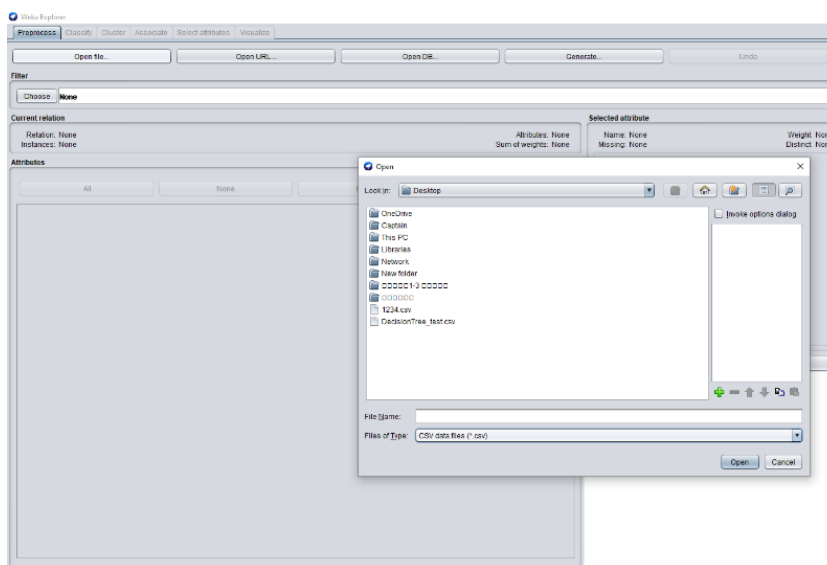
ซึ่งมีขั้นตอนการทำงาน ดังนี้

ขั้นตอนที่ 1 เปิดโปรแกรม Weka 3.8.4

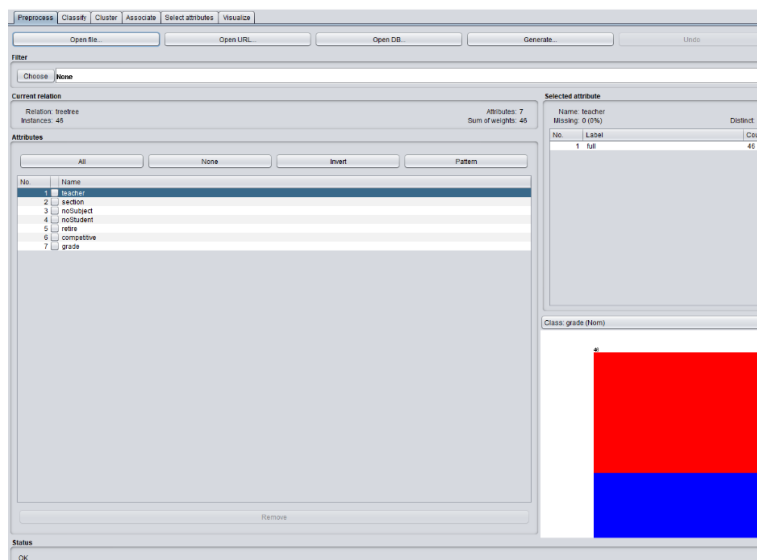


ภาพที่ 3.2 แสดงการเข้าใช้โปรแกรม Weka 3.8.4

ขั้นตอนที่ 2 นำเข้าข้อมูลที่ได้จัดเตรียมไว้ โดยเลือกที่ Application>>Explorer>>Open file เลือกไฟล์ข้อมูลที่ต้องการนำมาทดสอบ ตามภาพที่ 3.9 และหลังจาก นั้นโปรแกรมแสดงหน้าจอ ข้อมูล

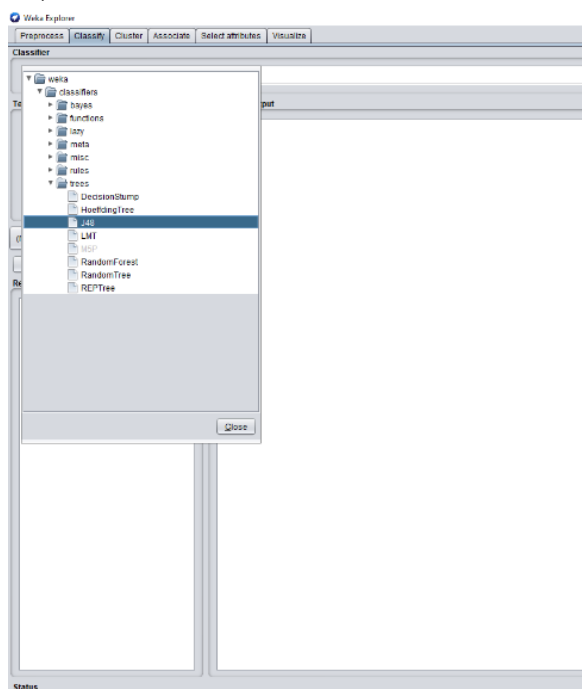


ภาพที่ 3.3 แสดงการนำเข้าไฟล์ข้อมูลเข้าโปรแกรม Weka 3.8.4

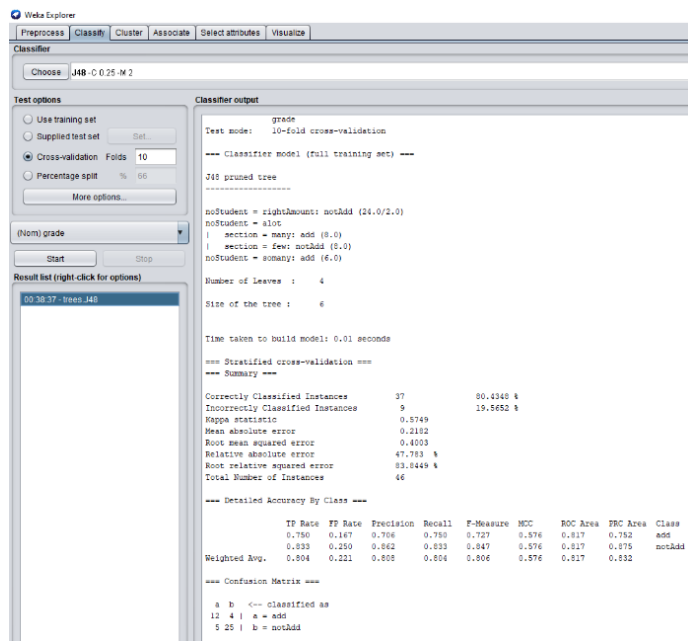


ภาพที่ 3.4 แสดงข้อมูลหลังจากนำเข้าโปรแกรม Weka 3.8.4

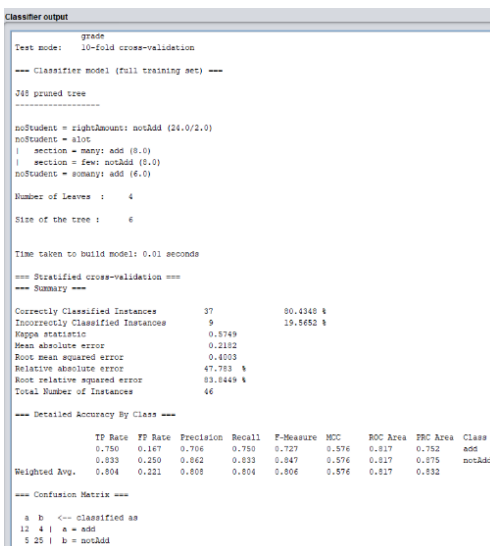
ขั้นตอนที่ 3 ดำเนินการเลือกเทคนิคการจัดกลุ่มข้อมูลแบบ Decision Tree โดย เลือกที่ Classification>>Choose>>tree และเลือกรูปแบบเป็น J48 ตามภาพที่ 3.11 จากนั้นทำการเลือก entropy (parent) แล้วกด ปุ่ม Start



ภาพที่ 3.5 แสดงการเลือกเทคนิคการจัดกลุ่มข้อมูลแบบ Decision Tree



ภาพที่ 3.6 แสดงหน้าจอฟลลัพท์ของโมเดลการจัดกลุ่มข้อมูลแบบ Decision Tree: J48



ภาพที่ 3.7 แสดงหน้าจอฟลลัพท์ของโมเดลการจัดกลุ่มข้อมูลแบบ Decision Tree: J48

จากผลลัพธ์การทดลองพบว่าเทคนิค Decision Tree: J48 ให้ผลลัพธ์การจำแนก ควรเพิ่มอาจารย์(Add) และไม่ควรมีอาจารย์(notAdd) มีความถูกต้องถึง 80.44%

```

      grade
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
-----

noStudent = rightAmount: notAdd (24.0/2.0)
noStudent = alot
| section = many: add (8.0)
| section = few: notAdd (8.0)
noStudent = somany: add (6.0)

Number of Leaves : 4
Size of the tree : 6

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances 37      80.4348 %
Incorrectly Classified Instances 9      19.5652 %
Kappa statistic 0.5749
Mean absolute error 0.2182
Root mean squared error 0.4003
Relative absolute error 47.783 %
Root relative squared error 83.8449 %
Total Number of Instances 46

=== Detailed Accuracy By Class ===

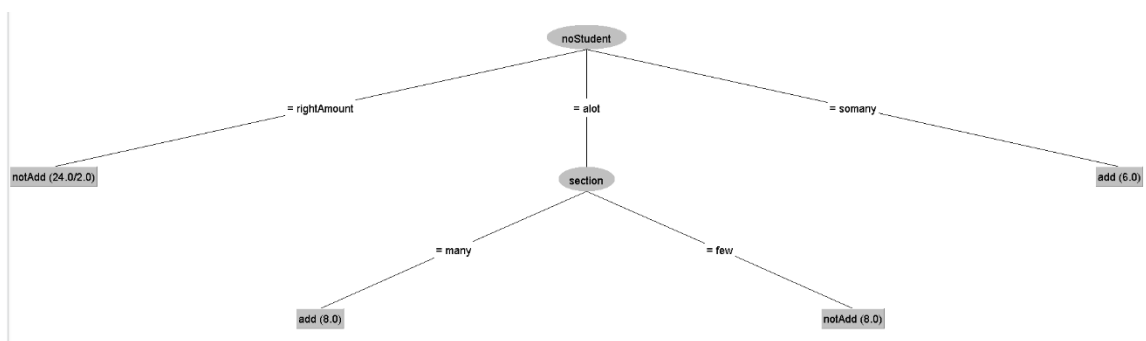
      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
-----
0.750  0.167  0.706  0.750  0.727  0.576  0.817  0.752  add
0.833  0.250  0.862  0.833  0.847  0.576  0.817  0.875  notAdd
Weighted Avg.  0.804  0.221  0.808  0.804  0.806  0.576  0.817  0.832

=== Confusion Matrix ===

 a b  <-- classified as
12 4 | a = add
 5 25 | b = notAdd

```

ภาพที่ 3.8 ผลลัพธ์จากการจำแนกกลุ่มแบบ Decision Tree: J48 ในโปรแกรม Weka



ภาพที่ 3.9 แสดงรูปแบบโมเดล Graph Decision Tree ในโปรแกรม Weka 3.8.4

ดังนั้น คณะผู้จัดทำจะใช้เทคนิคของการจำแนกกลุ่มแบบ Decision Tree: J48 มาใช้ในการศึกษา เนื่องจากให้ผลลัพธ์ของกฎที่สามารถทำนายได้จำนวน 4 กฎ ซึ่งสามารถนำไปใช้ในการแบ่งกลุ่มได้ตามเงื่อนไขได้ชัดเจน และสามารถนำกฎที่ได้สามารถนำไปวิเคราะห์กฎต่อไปได้ โดยสามารถจำแนกกฎที่ได้ ดังนี้

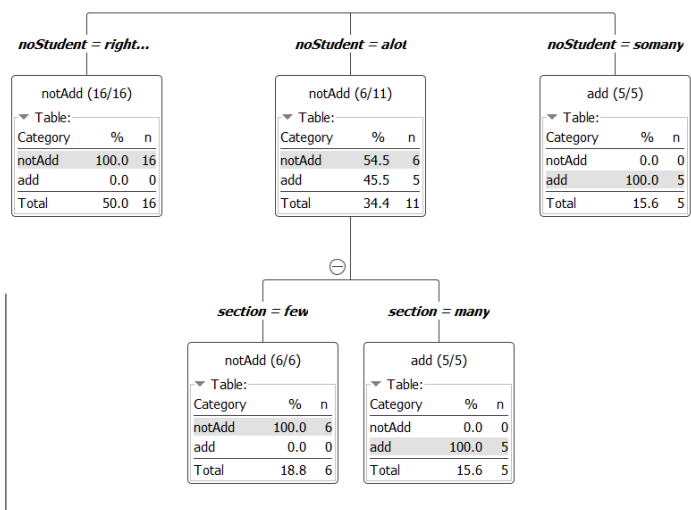
กฎข้อที่ 1 IF noStudent = rightAmount THEN grade = notAdd หมายความว่า ถ้าสัดส่วนนักศึกษาต่ออาจารย์ มีสัดส่วนนักศึกษาต่ออาจารย์พอดี ผลการพิจารณาพบว่า “ไม่ควรเพิ่มอาจารย์”

กฎข้อที่ 2 IF noStudent = somany THEN grade = add หมายความว่า ถ้าสัดส่วนนักศึกษาต่ออาจารย์ มีสัดส่วนนักศึกษาต่ออาจารย์เยอะมาก ผลการพิจารณาพบว่า “ควรเพิ่มอาจารย์”

กฎข้อที่ 3 IF noStudent = alot AND section = many THEN grade = add หมายความว่า ถ้าสัดส่วนนักศึกษาต่ออาจารย์ มีสัดส่วนนักศึกษาต่ออาจารย์เยอะ และจำนวนภาะการสอนในหลักสูตร/จำนวน Section มีจำนวนภาะการสอนในหลักสูตรเยอะ ผลการพิจารณาพบว่า “ควรเพิ่มอาจารย์”

กฎข้อที่ 4 IF noStudent = alot AND section = few THEN grade = notAdd หมายความว่า ถ้าสัดส่วนนักศึกษาต่ออาจารย์ มีสัดส่วนนักศึกษาต่ออาจารย์เยอะ และจำนวนภาะการสอนในหลักสูตร/จำนวน Section มีจำนวนภาะการสอนในหลักสูตรน้อย ผลการพิจารณาพบว่า “ไม่ควรเพิ่มอาจารย์”

ทางคณะผู้จัดทำยังได้นำโมเดลของข้อมูลที่เลือกใช้ มาเปรียบเทียบกับโมเดลที่สร้างด้วยโปรแกรม KNIME ซึ่งคณะผู้จัดทำพบว่าได้ผลลัพธ์จากการจำแนกกลุ่ม และรูปแบบแผนภาพโมเดลที่ตรงกัน



ภาพที่ 3.10 Decision Tree View แผนผังการตัดสินใจ ในโปรแกรม KNIME

หลังจากผู้วิเคราะห์เลือกการทดสอบประสิทธิภาพของ Model ด้วยวิธี Self Consistency Test หรือเรียกว่า Use Training Set เป็นวิธีการที่นำข้อมูลที่ใช้ในการสร้างโมเดล (model) และข้อมูลที่ใช้ในการทดสอบโมเดลเป็นข้อมูลชุดเดียวกัน ที่ได้ทำการคัดเลือกมาทั้งหมด ซึ่งผู้วิเคราะห์เลือกโปรแกรมที่ใช้นำเสนอ คือ โปรแกรม Weka 3.8.4 พบว่าการทดสอบประสิทธิภาพโมเดล Decision Tree (J48) พิจารณาได้ว่า โมเดลที่ถูกสร้างขึ้น มีค่าความถูกต้องเฉลี่ยในทุกโมเดล เท่ากับ 80.44% มีค่าการทำนายข้อมูลไม่ถูกต้องเท่ากับ 19.57% และมีค่าความคลาดเคลื่อน เท่ากับ 0.2182 และเมื่อพิจารณาสวนค่า Confusion Matrix ข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนาย แบ่งตามประเภทของผลการวิเคราะห์ ควรเพิ่มอาจารย์(add) และไม่ควรมเพิ่มอาจารย์(notAdd)

ดังนั้น ผู้วิเคราะห์จึงเลือกใช้โมเดล Decision Tree: J48 ในการนำไปหาแนวทางการเพิ่มอัตราบุคลากรสายวิชาการ เพราะมีค่าความถูกต้องของโมเดล และค่าเฉลี่ย Confusion Matrix จากการวิเคราะห์ ควรเพิ่มอาจารย์(Add) และไม่ควรมเพิ่มอาจารย์(NotAdd)

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      37          80.4348 %
Incorrectly Classified Instances    9          19.5652 %
Kappa statistic                    0.5749
Mean absolute error                 0.2182
Root mean squared error            0.4003
Relative absolute error            47.783 %
Root relative squared error        83.8449 %
Total Number of Instances          46

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0.750   0.167   0.706     0.750   0.727     0.576   0.817    0.752    add
                0.833   0.250   0.862     0.833   0.847     0.576   0.817    0.875    notAdd
Weighted Avg.   0.804   0.221   0.808     0.804   0.806     0.576   0.817    0.832

=== Confusion Matrix ===

 a  b  <-- classified as
12  4 | a = add
 5 25 | b = notAdd

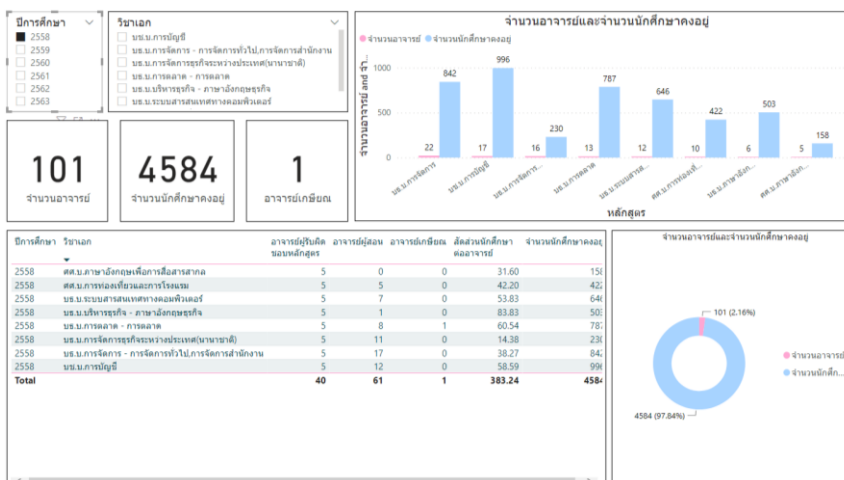
```

ภาพที่ 3.11 ผลลัพธ์จากการจำแนกกลุ่มแบบ Decision Tree: J48 ในโปรแกรม Weka

3.1.6 เผยแพร่ผลวิเคราะห์ (Deployment) ขั้นตอนการนำผลลัพธ์ที่ได้ไปใช้งานเป็นการทั่วไป อาจจัดทำเป็นรูปแบบของรายงาน (Report) หรือแผนภาพ (Dashboard) ที่พร้อม นำไปใช้ประโยชน์ในการวางแผน กำหนดกลยุทธ์ และดำเนินการต่าง ๆ ในทางธุรกิจ คณะผู้จัดทำนำผลข้อมูลที่ทำการวิเคราะห์แล้ว แสดงผลข้อมูลบน Web browser โดยใช้ชุดคำสั่ง HTML และ CSS3 ด้วยการนำเสนอข้อมูลแบบ Visualization เป็นการแสดงผล ข้อมูลในรูปแบบของภาพ โดยใช้โปรแกรม Power Bi ซึ่งทางคณะผู้จัดทำได้ยกตัวอย่างการ จัดทำเป็นรูปแบบของรายงาน (Report)

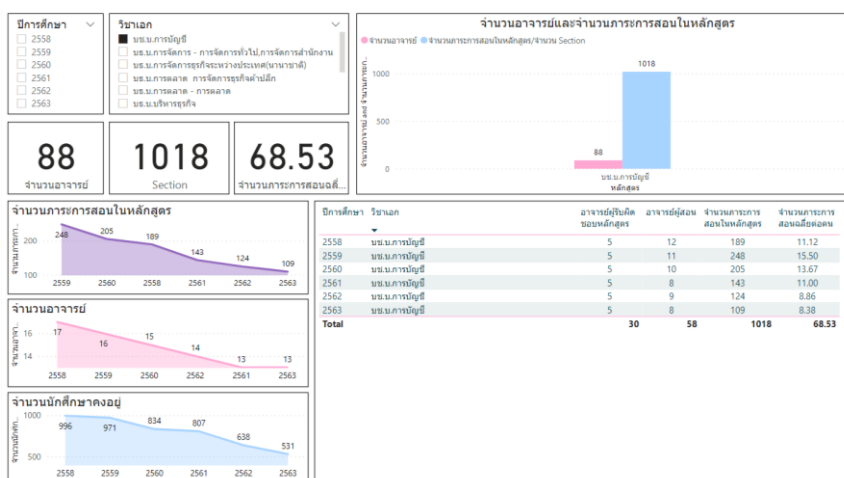
หรือแผนภาพ (Dashboard) กับข้อมูลนักศึกษาและอาจารย์ย้อนหลังปี 2558-2563 โดยแต่ละปีจะมีรูปแบบเดียวกัน

1) แดชบอร์ดของข้อมูลจำนวนอาจารย์และจำนวนนักศึกษาคงอยู่ที่ในปีการศึกษา 2558-2563 แสดงผลโดยแยกเป็น ปีการศึกษา วิชาเอก จำนวนอาจารย์ จำนวนนักศึกษาคงอยู่ อาจารย์เกษียณ ผู้ใช้สามารถกรองข้อมูลได้จากตัวกรอง



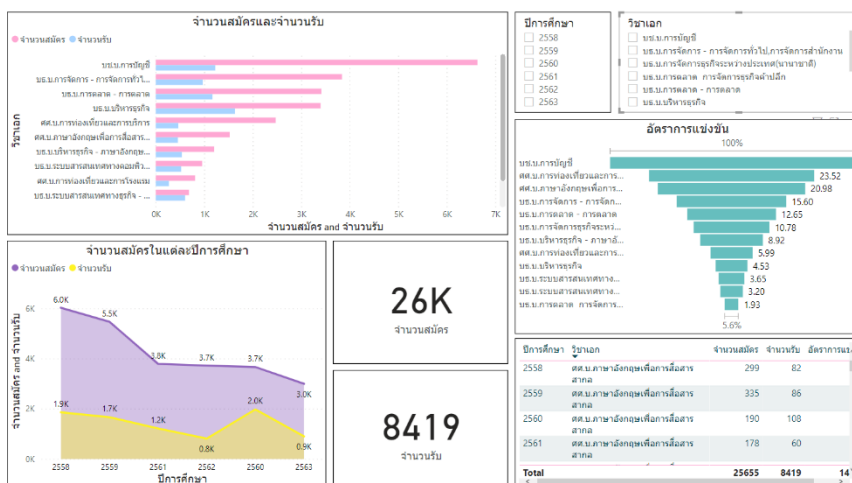
ภาพที่ 3.12 แดชบอร์ดของข้อมูลจำนวนอาจารย์และนักศึกษาคงอยู่

2) แดชบอร์ดของข้อมูลจำนวนอาจารย์และจำนวนภาระการสอนในหลักสูตรในปีการศึกษา 2558-2563 แสดงผลโดยแยกเป็น ปีการศึกษา วิชาเอก จำนวนอาจารย์ จำนวนภาระการสอนในหลักสูตร ผู้ใช้สามารถกรองข้อมูลได้จากตัวกรอง



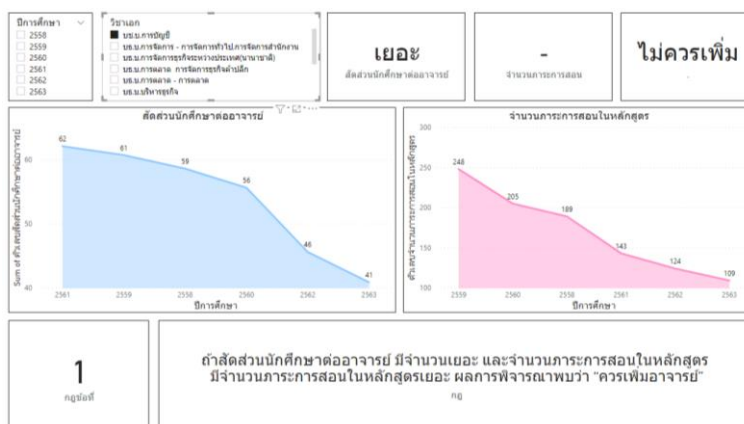
ภาพที่ 3.13 แดชบอร์ดของข้อมูลจำนวนอาจารย์และจำนวนภาระการสอนในหลักสูตร

3) แดชบอร์ดของข้อมูลจำนวนสมัครและจำนวนรับสมัครในปี 2558-2563 แสดงผลโดยแยกเป็น ปีการศึกษา วิชาเอก จำนวนสมัคร จำนวนรับสมัคร อัตราการแข่งขันของแต่ละหลักสูตร ผู้ใช้สามารถกรองข้อมูลได้จากตัวกรอง



ภาพที่ 3.14 แดชบอร์ดของข้อมูลจำนวนสมัครและจำนวนรับสมัคร

4) แดชบอร์ดผลการพยากรณ์จากการทำโมเดลต้นไม้ตัดสินใจ แสดงผลโดยแยกเป็น ปีการศึกษา วิชาเอก สัดส่วนนักศึกษาต่ออาจารย์ จำนวนภาระการสอนในหลักสูตร ผู้ใช้สามารถดูผลลัพธ์ของแต่ละหลักสูตรในแต่ละปีการศึกษาได้จากตัวกรอง



ภาพที่ 3.15 แดชบอร์ดผลการพยากรณ์จากการทำโมเดลต้นไม้ตัดสินใจ

3.2 การออกแบบเว็บไซต์

3.2.1 การออกแบบ Wireframe หน้าจอเว็บไซต์

1) หน้าของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บ



ภาพที่ 3.16 หน้าของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บ

2) หน้าของ Dashboard



ภาพที่ 3.17 หน้าของ Dashboard

3.3 บทสรุป

วัตถุประสงค์เพื่อวิเคราะห์ข้อมูลเพื่อสนับสนุนการตัดสินใจการเพิ่มอัตราบุคลากรสายวิชาการ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ ด้วยการนำเทคนิคเหมืองข้อมูล (Data mining) การจำแนกข้อมูล (Data classification) และรูปแบบของแผนภูมิต้นไม้ (Decision tree) ข้อมูลอาจารย์และนักศึกษา สามารถนำรูปแบบ (Model) จากโปรแกรม Weka 3.8.4 พบว่าการทดสอบประสิทธิภาพโมเดล Decision Tree (J48) พิจารณาได้ว่าโมเดลที่ถูกสร้างขึ้น มีค่าความถูกต้องเฉลี่ยในทุกโมเดล เท่ากับ 80.44% ซึ่งมีความน่าเชื่อถือ ที่สามารถนำมาวางแผนการเพิ่มอัตราบุคลากรสายวิชาการ ของคณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ นำไปสู่การวางแผนต่อไปในอนาคต เป็นขั้นตอนแรกของกระบวนการนำข้อมูลอาจารย์และนักศึกษา ปีการศึกษา 2558 ถึง 2563 จำนวนรายการทั้งหมด 46 รายการ รวบรวมข้อมูลและพิจารณาตรวจสอบความถูกต้องและความครบถ้วนของข้อมูล พิจารณาข้อมูลทั้งหมดว่าส่วนใดของข้อมูลที่เป็นต่อการวิเคราะห์ข้อมูลให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ นำข้อมูลที่ได้มาทำการคัดเลือกข้อมูล (Data Selection) การกลั่นกรองข้อมูล (Data Cleaning) การแปลงรูปแบบของข้อมูล (Data Transformation) ผู้จัดทำวิเคราะห์ข้อมูล ด้วยการทำเหมืองข้อมูลแบบการจำแนกข้อมูล (Data classification) ด้วยการสร้างโมเดล Decision Tree ที่สร้างขึ้นด้วยโปรแกรม Weka และโปรแกรม KNIME เพื่อเปรียบเทียบความแม่นยำของโปรแกรม