

## บทที่ 3

### วิธีดำเนินโครงการ

โครงการนี้มุ่งเน้นการวิเคราะห์ปัญหาที่เกิดขึ้นจากระบบลงทะเบียนเรียนออนไลน์ของนักศึกษาในมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ โดยใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) และการวิเคราะห์ข้อมูลแบบ CRISP-DM เพื่อทำการวิเคราะห์ปัญหาการลงทะเบียน และสาเหตุของการเกิดปัญหา โมเดลที่ใช้ในการวิเคราะห์ประกอบไปด้วย 3 โมเดลหลัก ได้แก่ การจับกลุ่มข้อมูลแบบค่าเฉลี่ย k (K-Means) การวิเคราะห์เชิงพรรณนา (Descriptive Analysis) การหากฎความสัมพันธ์ (Association Rules) เพื่อสะท้อนสภาพการใช้งานระบบในปัจจุบันระบบที่พัฒนาขึ้นจะช่วยสนับสนุนการตัดสินใจของฝ่ายทะเบียนและผู้บริหารในการวางแผนปรับปรุงระบบให้มีประสิทธิภาพและลดโอกาสการล่มของระบบในช่วงเวลาสำคัญ ซึ่งในบทนี้จะแสดงถึงวิธีในการดำเนินงานดังนี้

3.1 การดำเนินงานกระบวนการด้วย CRISP-DM

3.2 ER-Diagram

3.3 Context Diagram

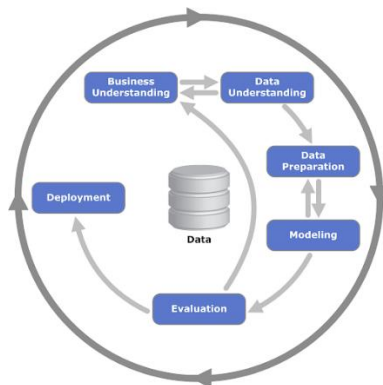
3.4 Data Flow Diagram Level 0

3.5 การออกแบบและพัฒนาเว็บไซต์

3.6 บทสรุป

#### 3.1 การดำเนินงานกระบวนการด้วย CRISP-DM

CRISP-DM เป็นกระบวนการหลักในการจัดทำเหมืองข้อมูลเพื่อการวิเคราะห์และใช้ประโยชน์ในทางธุรกิจ กระบวนการวิเคราะห์ข้อมูล ด้วย CRISP-DM หรือ (Cross Industry Standard Process for Data Mining) ประกอบด้วย 6 ขั้นตอน แต่ละขั้นตอนในรูปแบบจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่อง ตัวอย่างเช่นเมื่อ ได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในขั้น Modeling และหลังจากนั้น อาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าจะโมเดลที่ให้ความถูกต้องมาก ขึ้นก็ได้ เป็นต้น



ภาพที่ 3.1 แสดงกระบวนการ CRISP-DM

ที่มา : <https://kamboonchob.medium.com>

ในกระบวนการนี้ประกอบด้วย 6 ขั้นตอน

### 3.1.1 ความเข้าใจในธุรกิจ (Business Understanding)

ผู้วิเคราะห์มุ่งเน้นไปที่การทำความเข้าใจกระบวนการทางข้อมูลและระบบวิเคราะห์ข้อมูล จากสิ่งรอบตัวการสัมภาษณ์หรือรับฟังปัญหา รวมถึงการมองหาปัญหาจากเรื่องต่าง ๆ เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ให้เกิดประโยชน์ โดยจัดลำดับความสำคัญ และกำหนดวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูล ฐานข้อมูลขนาดใหญ่และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลทาง Data Mining หารูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลของข้อมูลการลงทะเบียนเรียนออนไลน์ เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ ความต้องการทั้งหมดจะนำมาจัดลำดับความสำคัญและกำหนดวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล

### 3.1.2 การทำความเข้าใจข้อมูล (Data Understanding)

ขั้นตอนการจัดเก็บและรวบรวมข้อมูล ผู้วิเคราะห์ต้องพิจารณาตรวจสอบความถูกต้องของข้อมูลที่ได้รับ โดยเลือกที่จะใช้ข้อมูลทั้งหมดหรือบางส่วนในการวิเคราะห์ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ ผู้วิเคราะห์ได้ดำเนินการรวบรวมข้อมูลพฤติกรรมการใช้งานระบบลงทะเบียนรายวิชา ซึ่งเป็นข้อมูล Log File จากระบบฐานข้อมูลของมหาวิทยาลัย เพื่อนำมาศึกษา

และทำความเข้าใจรูปแบบความสัมพันธ์ที่ซ่อนอยู่ของข้อมูลพฤติกรรมการลงทะเบียนและการ  
ลบบรายวิชา โดยมีรายละเอียดขั้นตอนดังนี้

1) การรวบรวมข้อมูล ผู้วิเคราะห์ข้อมูลดำเนินการรวบรวมข้อมูลจาก  
ฐานข้อมูลระบบลงทะเบียน (Registration Log) ซึ่งประกอบด้วยไฟล์ข้อมูลหลักคือ log\_regis.csv  
ที่บันทึกพฤติกรรมการเข้าใช้งานในช่วงเวลาที่กำหนด

2) การเลือกข้อมูลที่จะใช้: ผู้วิเคราะห์เลือกใช้ข้อมูลพฤติกรรมการทำรายการ  
ของผู้ใช้งาน ซึ่งประกอบด้วย 3 แอตทริบิวต์หลักที่สำคัญต่อการวิเคราะห์ ได้แก่

log\_id รหัสอ้างอิงของรายการ เพื่อใช้ในการนับจำนวนภาระงาน (Load)

log\_time วันที่และเวลาที่เกิดรายการ เพื่อใช้ในการวิเคราะห์อนุกรมเวลา  
(Time Series Analysis)

log\_text รายละเอียดของกิจกรรม เช่น "ลงทะเบียนสำเร็จ" หรือ "ลบบรายวิชา  
สำเร็จ" เพื่อใช้ในการจำแนกประเภทกิจกรรม

3) การตรวจสอบข้อมูลที่รวบรวมได้: ผู้วิเคราะห์ได้ทำการตรวจสอบ  
โครงสร้างและปริมาณของข้อมูล Log การลงทะเบียนรายวิชา เพื่อพิจารณาความพร้อมก่อน  
นำไปประมวลผล โดยมีรายละเอียดดังนี้

3.1 จำนวนข้อมูล ข้อมูลที่นำมาวิเคราะห์เป็นไฟล์บันทึกกิจกรรม (Log  
File) ในรูปแบบนามสกุล .csv

3.2 โครงสร้างของแอตทริบิวต์ ในแต่ละรายการของข้อมูล (Row)  
ประกอบด้วยแอตทริบิวต์ (Attribute) ทั้งหมด 3 ส่วนหลักที่สำคัญ ได้แก่

3.3 รหัสรายการ (log\_id) ข้อมูลประเภทตัวเลข (Integer) ใช้เป็นคีย์  
หลักเพื่อระบุตัวตนของแต่ละกิจกรรม

3.4 วันเวลาที่เกิดรายการ (log\_time) ข้อมูลประเภทข้อความ  
(Nominal) ที่ระบุลำดับเวลาในการทำรายการ

3.5 รายละเอียดกิจกรรม (log\_text) ข้อมูลประเภทข้อความ (Nominal)  
ที่ระบุประเภทพฤติกรรม เช่น การลงทะเบียน หรือ การลบบรายวิชา

การเตรียมข้อมูลเป็นขั้นตอนสำคัญก่อนการวิเคราะห์ โดยเป็นกระบวนการแปลงข้อมูลที่รวบรวมมาให้อยู่ในรูปแบบที่ถูกต้อง สมบูรณ์ และพร้อมใช้งาน ขั้นตอนนี้เรียกว่า การทำความสะอาดข้อมูล (Data Cleaning) ซึ่งมักใช้เวลาามากที่สุดในกระบวนการวิเคราะห์ข้อมูลทั้งหมด กระบวนการดังกล่าวประกอบด้วย การตรวจสอบและแก้ไขข้อมูลที่ผิดพลาด การจัดการค่าที่หายไป การปรับรูปแบบข้อมูลให้เป็นมาตรฐาน และการคัดกรองข้อมูลที่ไม่เกี่ยวข้อง เพื่อให้ข้อมูลมีความถูกต้องและสอดคล้องกันมากที่สุด ปัจจุบันระบบจัดเก็บข้อมูลสมัยใหม่ช่วยลดความผิดพลาดจากการป้อนข้อมูลของมนุษย์ด้วยเทคโนโลยีอัตโนมัติ เช่น การสแกนและการดึงข้อมูลจากแหล่งที่เชื่อถือได้ อย่างไรก็ตาม การทำความสะอาดข้อมูลยังคงเป็นขั้นตอนที่ใช้เวลามากกว่า 50% ของระยะเวลาการทำงานทั้งหมด และมีผลโดยตรงต่อคุณภาพและความน่าเชื่อถือของผลการวิเคราะห์ในขั้นตอนต่อไป

3.1.3 การเตรียมข้อมูล (Data Preparation) ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาซึ่งเป็นขั้นตอนที่ใช้เวลามาก ผู้วิเคราะห์ข้อมูลทำการจัดการกับข้อมูลที่ได้โดยรวบรวมข้อมูลและแปลงข้อมูลที่ได้ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้เพื่อลดความผิดพลาดให้น้อยที่สุด โดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลที่มีประสิทธิภาพอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออกดำเนินการตามขั้นตอนดังนี้

1) การคัดเลือกข้อมูล (Data Selection) ขั้นตอนนี้เป็นการคัดเลือกข้อมูลที่เกี่ยวข้องกับการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ โดยผู้จัดทำได้รับข้อมูลจากหน่วยงานทะเบียนในรูปแบบไฟล์ข้อมูลจำนวน 4 ไฟล์ ได้แก่

1.1 ไฟล์ข้อมูลการเข้าใช้งานระบบ (log\_login) ใช้สำหรับวิเคราะห์พฤติกรรมกรการเข้าสู่ระบบของนักศึกษา เช่น วัน เวลา และความถี่ในการเข้าใช้งาน

log_id	std_id_pri	log_text	log_ip_address	log_browser	log_time
2	160635 70341849747	login	49.230.155.222	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
3	160636 70341848044	login	171.100.168.49	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
4	160637 70341872115	login	58.9.12.115	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
5	160638 70341872909	login	203.158.164.34	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
6	160639 70341849747	logout	49.230.155.222	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
7	160640 70341858112	login	118.175.247.121	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
8	160641 70341856727	login	27.55.16.248	Mozilla/5.0 Unknown Windows OS	8/15/2015 22:30
9	160642 70341850922	login	49.49.206.130	Chrome/30.0.0.0 Linux Generic Mobile	8/15/2015 22:30
10	160643 70341875802	login	180.183.144.241	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:30
11	160644 70341875299	login	49.228.153.151	Safari/600.1.4 Mac OS X Apple iPhone	8/15/2015 22:30
12	160645 70341875588	login	49.228.153.151	Chrome/44.0.2403.130 Unknown Windows OS	8/15/2015 22:30
13	160647 70341875544	login	61.7.185.111	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:30
14	160648 70341873782	login	118.172.7.159	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:30
15	160649 70341845153	login	203.150.116.232	Safari/600.1.4 Mac OS X Apple iPhone	8/15/2015 22:30
16	160650 70341850989	login	113.53.84.126	Safari/600.1.4 Mac OS X Apple iPhone	8/15/2015 22:30
17	160651 70341878459	logout	118.172.24.215	Chrome/31.0.1650.57 Unknown Windows OS	8/15/2015 22:30
18	160652 70341852854	login	171.4.241.137	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:31
19	160653 70341850592	login	49.48.173.1	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:31
20	160654 70341850358	login	180.180.156.105	Safari/600.1.4 Mac OS X Apple iPhone	8/15/2015 22:31
21	160655 70341868788	login	171.4.242.251	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:31
22	160656 70341875181	login	180.183.207.11	Mozilla/5.0 Mac OS X Apple iPhone	8/15/2015 22:31
23	160657 70341855042	login	27.55.161.198	Chrome/43.0.2357.93 Linux Generic Mobile	8/15/2015 22:31
24	160658 70341874438	logout	124.121.32.242	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:31
25	160659 70341867876	login	171.5.201.54	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:31
26	160660 70341875299	logout	61.7.185.174	Safari/600.1.4 Mac OS X Apple iPhone	8/15/2015 22:31

ภาพที่ 3.2 แสดงข้อมูลจากจำนวนผู้เข้าใช้งาน

1.2 ไฟล์ข้อมูลการลงทะเบียนรายวิชา (log\_regis) แสดงกิจกรรมการเพิ่ม-ถอนรายวิชา เพื่อนำมาวิเคราะห์ช่วงเวลาที่มีการใช้งานระบบสูงผิดปกติ

log_id	std_id_pri	course_id_pri	log_text	log_time
2	4 70341860455	24117770341883500	ลงทะเบียน	8/9/2015 23:58
3	5 70341860455	24117770341884700	ลงทะเบียน	8/10/2015 0:11
4	6 70341860455	24117770341884800	ลงทะเบียน	8/10/2015 0:12
5	7 70341860455	24117770341883700	ลงทะเบียน	8/10/2015 0:14
6	8 70341860455	24117770341883500	ลงทะเบียน	8/10/2015 0:15
7	9 70341860455	24117770341884800	ลงทะเบียน	8/10/2015 0:30
8	10 70341860455	24117770341883500	ลงทะเบียน	8/10/2015 0:39
9	11 70341860455	24117770341883500	ลบรายวิชา	8/10/2015 0:40
10	12 70341859067	24117770341884200	ลงทะเบียน	8/10/2015 8:22
11	13 70341876247	24117770341882800	ลงทะเบียน	8/10/2015 8:22
12	14 70341858684	24117770341884000	ลงทะเบียน	8/10/2015 8:22
13	15 70341858708	24117770341884000	ลงทะเบียน	8/10/2015 8:22
14	16 70341847611	24117770341884300	ลงทะเบียน	8/10/2015 8:22
15	17 70341858708	24117770341884000	ลงทะเบียน	8/10/2015 8:22
16	18 70341844288	24117770341883300	ลงทะเบียน	8/10/2015 8:22
17	19 70341856844	24117770341884500	ลงทะเบียน	8/10/2015 8:23
18	20 70341876241	24117770341882800	ลงทะเบียน	8/10/2015 8:23
19	21 70341854406	24117770341884700	ลงทะเบียน	8/10/2015 8:23
20	22 70341848296	24117770341882700	ลงทะเบียน	8/10/2015 8:23
21	23 70341875439	24117770341884500	ลงทะเบียน	8/10/2015 8:23
22	24 70341858777	24117770341884000	ลงทะเบียน	8/10/2015 8:23
23	25 70341858813	24117770341884200	ลงทะเบียน	8/10/2015 8:23
24	26 70341860542	24117770341884700	ลงทะเบียน	8/10/2015 8:23
25	27 70341848296	24117770341882700	ลงทะเบียน	8/10/2015 8:23
26	28 70341876241	24117770341882600	ลงทะเบียน	8/10/2015 8:23

ภาพที่ 3.3 แสดงข้อมูลจากจำนวนผู้ลงทะเบียน

2) การกลั่นกรองข้อมูล (Data Cleaning) คือการทำความสะอาดข้อมูล เป็นกระบวนการตรวจสอบและการแก้ไข (หรือลบ) รายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูลตารางหรือฐานข้อมูล ซึ่งเป็นหลักสำคัญของฐานข้อมูล ทางผู้วิเคราะห์ข้อมูลได้ดำเนินการหลังจากทำความสะอาดข้อมูลแล้ว ผู้จัดทำได้ดำเนินการรวมข้อมูลจากทั้ง 4 ไฟล์เข้าด้วยกันโดยใช้รหัสวิชาและช่วงเวลาเป็นตัวเชื่อมโยง เพื่อให้สามารถวิเคราะห์ความสัมพันธ์ระหว่างพฤติกรรมการใช้งานกับรายวิชาที่ลงทะเบียนได้จากนั้นได้ทำการแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมต่อการวิเคราะห์ ดังนี้

Row No.	ip_log_id	std_id_pri	log_text	log_ip_address	log_browser	log_time
1	160635	70341849747	login	49.230.155.222	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
2	160636	70341848044	login	171.100.168.49	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
3	160637	70341872115	login	58.9.12.115	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
4	160638	70341872909	login	203.158.164.34	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
5	160639	70341849747	logout	49.230.155.222	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
6	160640	70341858112	login	118.175.247.121	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:29
7	160641	70341856727	login	27.55.16.248	Mozilla/5.0 Unknown Windows OS	8/15/2015 22:30
8	160642	70341850922	login	49.49.206.130	Chrome/30.0.0.0 Linux Generic Mobile	8/15/2015 22:30
9	160643	70341875802	login	180.183.144.241	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:30
10	160644	70341875299	login	49.228.153.151	Safari/600.14 Mac OS X Apple iPhone	8/15/2015 22:30
11	160645	70341875588	login	49.228.153.151	Chrome/44.0.2403.130 Unknown Windows OS	8/15/2015 22:30
12	160647	70341875544	login	617.185.111	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:30
13	160648	70341873782	login	118.172.7.159	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:30
14	160649	70341845153	login	203.150.116.232	Safari/600.14 Mac OS X Apple iPhone	8/15/2015 22:30
15	160650	70341850989	login	113.53.84.126	Safari/600.14 Mac OS X Apple iPhone	8/15/2015 22:30
16	160651	70341847849	logout	118.172.24.215	Chrome/310.1650.57 Unknown Windows OS	8/15/2015 22:30
17	160652	70341852854	login	171.14.241.137	Chrome/44.0.2403.155 Unknown Windows OS	8/15/2015 22:31

ภาพที่ 3.4 แสดงการจัดการค่าที่ขาดหายไป

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	course_id_pri	course_name	study_year	study_term	course_type	subject_id	campus_id	department	major_id	p_person_lir	has_lab	course_id	place_stuc	course_sti	grade_sen	grade_ser	study_type	comment	jobtraining	ass_cognit
2	2.3841E+16	S1000003_SEC_1	2564	1	5	2.3841E+16	1	5	208	14	0	0	1	0	2	1	0			
3	2.3841E+16	S1002101_SEC_1	2564	1	5	2.3841E+16	1	9	53	13	0	0	1	0	2	1	0			
4	2.41178E+16	SEC2 2558 การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
5	2.41178E+16	SEC2 2558 การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
6	2.41178E+16	SEC2 2558 การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
7	2.41178E+16	SEC2 2558 การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
8	2.41178E+16	SEC2 2558 การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
9	2.41178E+16	SEC2 2558 การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
10	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
11	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	50	0	0	1	0	2	0	0			
12	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
13	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
14	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
15	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
16	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
17	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	80	0	0	1	0	2	0	0			
18	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
19	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
20	2.41178E+16	SEC1 น.1-การจัดการ	2558	1	5	2.3841E+16	6	23	202	50	0	0	1	0	2	0	0			
21	2.41178E+16	SEC1 น.1-การตลาด	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
22	2.41178E+16	SEC1 น.1-การตลาด	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
23	2.41178E+16	SEC1 น.1-การตลาด	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
24	2.41178E+16	SEC1 น.1-การตลาด	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
25	2.41178E+16	SEC1 น.1-การบัญชี	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			
26	2.41178E+16	SEC1 น.1-การบัญชี	2558	1	5	2.3841E+16	6	23	202	40	0	0	1	0	2	0	0			

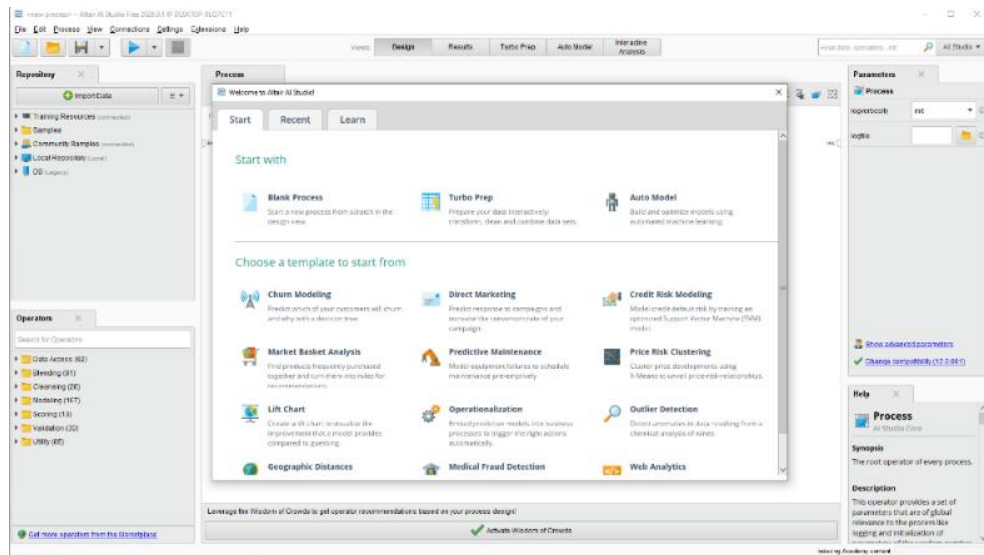
ภาพที่ 3.5 แสดงการแยกคอลัมน์ที่มีออกมาเป็นคอลัมน์ย่อยๆ

3.1.4 การสร้างแบบจำลอง (Modeling) ขั้นตอนการสร้างแบบจำลองเป็นกระบวนการนำข้อมูลที่ผ่านการเตรียมและทำความสะอาดแล้ว มาวิเคราะห์ด้วยเทคนิคทางสถิติและเหมืองข้อมูล เพื่อค้นหารูปแบบ แนวโน้ม และปัจจัยที่เกี่ยวข้องกับปัญหาการลงทะเบียนเรียนออนไลน์

ของนักศึกษา โดยการศึกษาครั้งนี้ใช้โปรแกรม RapidMiner เป็นเครื่องมือหลักในการพัฒนาแบบจำลอง

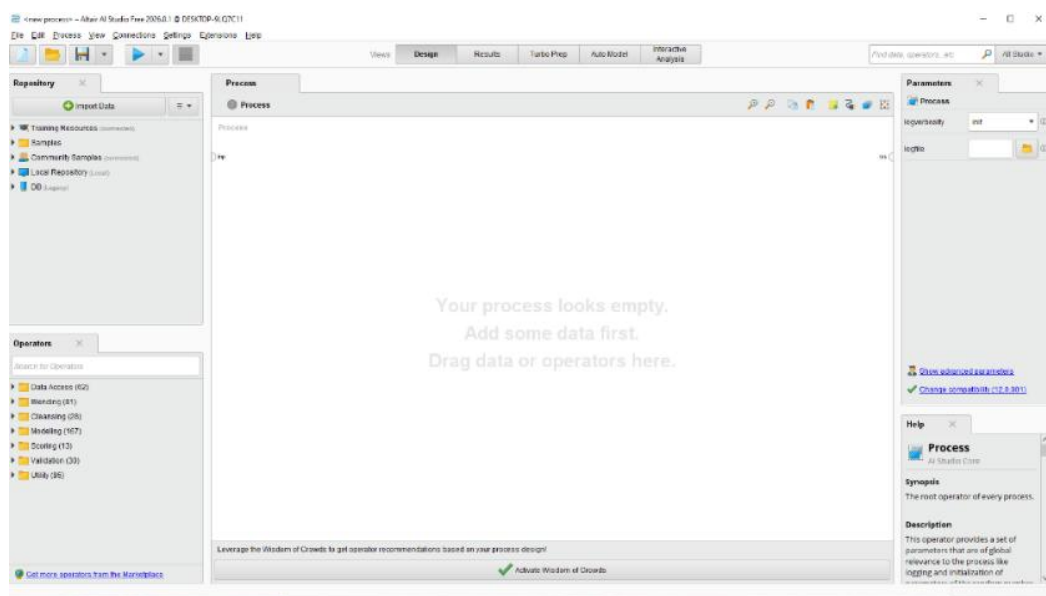
### 1) การนำข้อมูลเข้าโปรแกรม RapidMiner

1.1) เปิดโปรแกรม RapidMiner และเลือก Blank Process เพื่อเริ่มต้นกระบวนการทำงานใหม่ แสดงดังภาพ



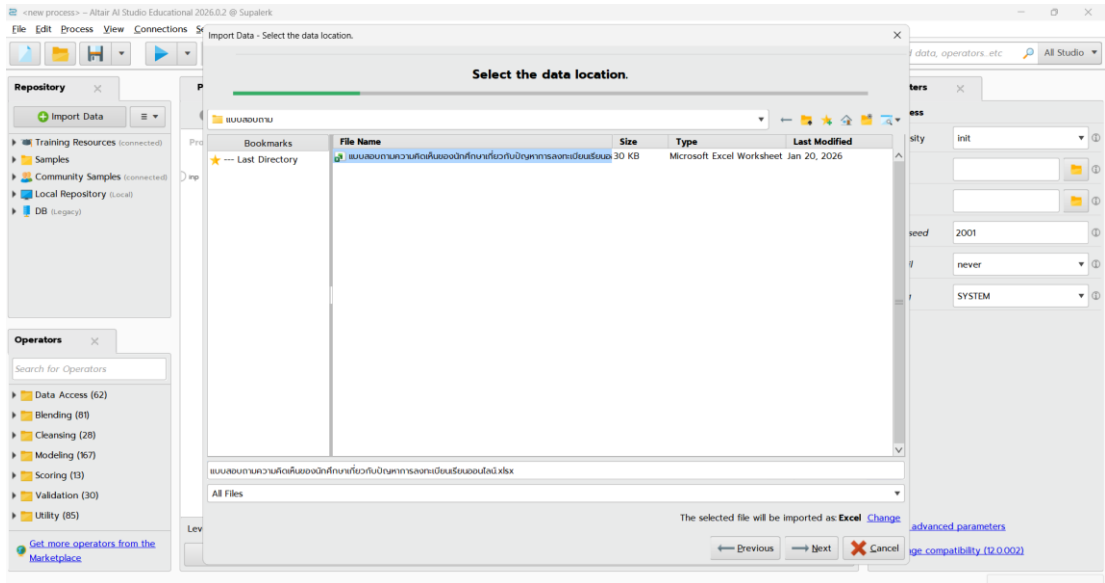
ภาพที่ 3.6 แสดงหน้าหลักของโปรแกรม RapidMiner

1.2) นำเข้าไฟล์ข้อมูลแบบสอบถามที่ผ่านการทำความสะอาดข้อมูลแล้ว โดยเลือกคำสั่ง Import Data



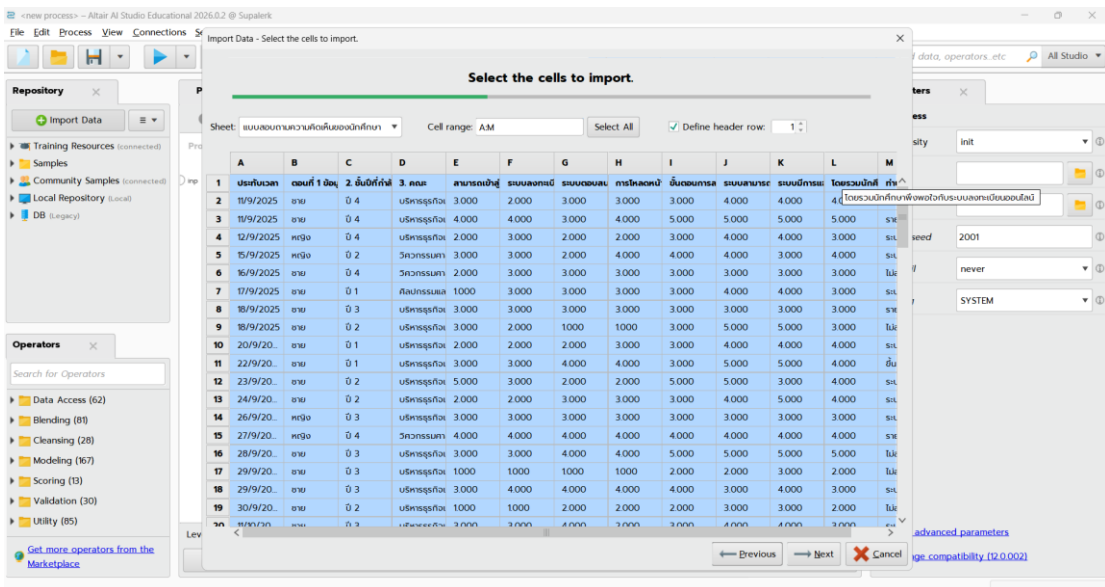
ภาพที่ 3.7 แสดงการนำเข้าข้อมูล

3) เลือกไฟล์ข้อมูลแบบสอบถามในรูปแบบ CSV เพื่อใช้ในการวิเคราะห์



ภาพที่ 3.8 แสดงข้อมูลที่จัดเก็บไว้ใน Folder

4) จัดเก็บชุดข้อมูลไว้ในไฟล์เตอร์ Data ภายในโปรแกรม เพื่อความสะดวกในการเรียกใช้งาน



ภาพที่ 3.9 แสดงข้อมูลที่ให้นำเข้า

5) ข้อมูลที่ได้จากการนำเข้าข้อมูลจะมีทั้งหมด จำนวน 1,048,575

แถว

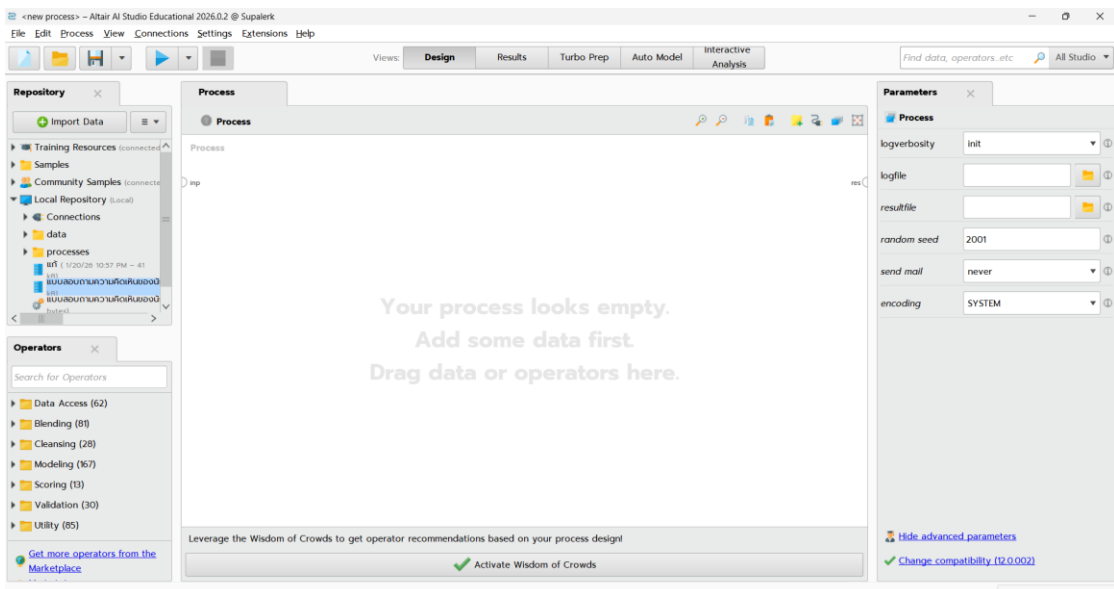
6 แอททริบิวต์

Row No.	ระยะเวลา	คนที่ 1 ชื่อเล่น	2. อันดับกำลังศึกษา	3. คณะ	สามารถเข้าสู่ระบบ	ระบบลงทะเบียน	ระบบขอเสนอแนะ	การไหลหน้าเว็บ	ขั้นตอนการลงทะเบียน	ระดับ
1	11/9/2025	ชาย	0 4	บริหารธุรกิจและ...	3	2	3	3	3	4
2	11/9/2025	ชาย	0 4	บริหารธุรกิจและ...	4	4	3	4	5	5
3	12/9/2025	หญิง	0 4	บริหารธุรกิจและ...	2	3	2	2	3	4
4	15/9/2025	หญิง	0 2	วิศวกรรมศาส...	3	3	2	4	4	4
5	16/9/2025	ชาย	0 4	วิศวกรรมศาส...	2	3	3	3	3	3
6	17/9/2025	ชาย	0 1	ศิลปกรรมและสถา...	1	3	3	3	3	4
7	18/9/2025	ชาย	0 3	บริหารธุรกิจและ...	3	3	3	3	3	3
8	18/9/2025	ชาย	0 2	บริหารธุรกิจและ...	3	2	1	1	3	5
9	20/9/2025	ชาย	0 1	บริหารธุรกิจและ...	2	2	2	3	3	4
10	22/9/2025	ชาย	0 1	บริหารธุรกิจและ...	3	3	4	4	3	5
11	23/9/2025	ชาย	0 2	บริหารธุรกิจและ...	5	3	2	2	5	5
12	24/9/2025	ชาย	0 2	บริหารธุรกิจและ...	2	2	3	3	3	4
13	26/9/2025	หญิง	0 3	บริหารธุรกิจและ...	3	3	3	3	3	3
14	27/9/2025	หญิง	0 4	วิศวกรรมศาส...	4	4	4	4	4	4
15	28/9/2025	ชาย	0 3	บริหารธุรกิจและ...	3	3	4	4	5	5
16	29/9/2025	ชาย	0 3	บริหารธุรกิจและ...	1	1	1	1	2	2
17	29/9/2025	ชาย	0 3	บริหารธุรกิจและ...	3	4	4	4	4	3

ภาพที่ 3.10 แสดงข้อมูลทั้งหมดที่นำเข้า

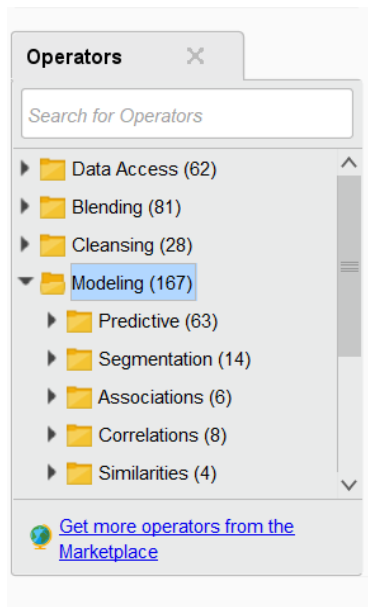
6) ใช้พื้นที่ Blank Process เป็นพื้นที่สำหรับลากและเชื่อมต่อ

โอเปอเรเตอร์เพื่อสร้างกระบวนการวิเคราะห์ข้อมูล



ภาพที่ 3.11 แสดงหน้า Blank Process

7) ดึงโอเปอเรเตอร์จาก Operators Panel ลงมาและเชื่อมต่อ เพื่อสร้าง Process



ภาพที่ 3.12 แสดง Operator ต่าง ๆ

## 2) ขั้นตอนการดำเนินงานของโมเดล

2.1) Data Transformation รวมข้อมูลจากหลายแหล่ง ทำความสะอาด แปลงรูปแบบข้อมูลและสร้างตัวแปรใหม่เพื่อให้ข้อมูลพร้อมสำหรับการวิเคราะห์

2.1.1 Retrieve log\_login นำเข้าข้อมูล Log การเข้าใช้งานระบบ

2.1.2 Retrieve log\_login นำเข้าข้อมูล Log การเข้าใช้งาน

ระบบ

2.1.3 Retrieve log\_regis นำเข้าข้อมูล Log การลงทะเบียน

2.1.4 Rename (2) ปรับชื่อคอลัมน์ของข้อมูลชุดที่สอง

2.1.5 Join รวมข้อมูลทั้งสองชุดเข้าด้วยกันโดยใช้ Key เดียวกัน

(เช่น user\_id)

2.1.6 Select Attributes เลือกเฉพาะตัวแปรที่จำเป็น ลดข้อมูล

ที่ไม่เกี่ยวข้อง

2.1.7 Remove Duplicates ลบข้อมูลที่ซ้ำซ้อน เพื่อลดความคลาดเคลื่อน

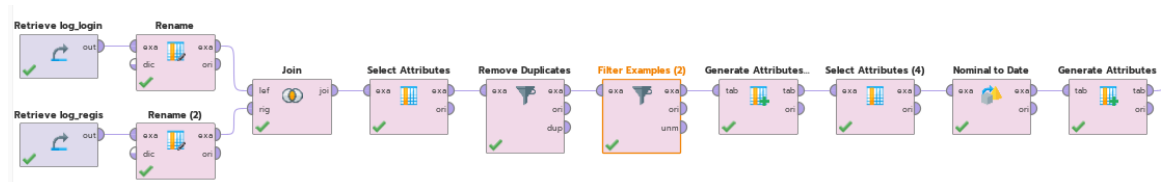
2.1.8 Filter Examples กรองข้อมูลตามเงื่อนไขที่ต้องการ เช่น ปีการศึกษา หรือค่าที่ไม่สมบูรณ์

2.1.9 Generate Attributes สร้างตัวแปรใหม่จากข้อมูลเดิม เช่น ดึงชั่วโมงจาก timestamp

2.1.10 Select Attributes (อีกครั้ง) เลือกเฉพาะตัวแปรที่ผ่านการสร้างและจะใช้จริง

2.1.11 Nominal to Date แปลงข้อมูลวันที่จากข้อความให้เป็นชนิด Date เพื่อให้สามารถคำนวณเชิงเวลาได้

2.1.12 Generate Attributes (รอบสุดท้าย) คำนวณค่าที่ต้องใช้ในการวิเคราะห์ เช่น จำนวนครั้ง login จำนวนครั้งลงทะเบียน ค่าเฉลี่ยเวลาเข้าใช้



ภาพที่ 3.13 แสดงเครื่องมือกระบวนการเปลี่ยนรูปแบบ

Open in Turbo Prep Auto Model Interactive Analysis

Row No.	std_id_pri	log_id	log_text_re...	login_flag	log_time	hour
1	70341849747	107078	ลงทะเบียน	1	Mar 8, 2016	22
2	70341848044	83835	ลงทะเบียน	1	Mar 8, 2016	22
3	70341848044	94785	สลายวิชา	1	Mar 8, 2016	22
4	70341872115	152179	ลงทะเบียน	1	Mar 8, 2016	22
5	70341872115	369264	สลายวิชา	1	Mar 8, 2016	22
6	70341872909	290150	ลงทะเบียน	1	Mar 8, 2016	22
7	70341872909	290198	สลายวิชา	1	Mar 8, 2016	22
8	70341849747	107078	ลงทะเบียน	0	Mar 8, 2016	22
9	70341858112	42501	ลงทะเบียน	1	Mar 8, 2016	22
10	70341858112	409207	สลายวิชา	1	Mar 8, 2016	22
11	70341856727	31750	ลงทะเบียน	1	Mar 8, 2016	22
12	70341856727	40468	สลายวิชา	1	Mar 8, 2016	22
13	70341850922	408090	ลงทะเบียน	1	Mar 8, 2016	22
14	70341850922	417851	สลายวิชา	1	Mar 8, 2016	22
15	70341875802	21099	ลงทะเบียน	1	Mar 8, 2016	22
16	70341875802	25529	สลายวิชา	1	Mar 8, 2016	22
17	70341875299	5634	ลงทะเบียน	1	Mar 8, 2016	22

ภาพที่ 3.14 แสดงผลกระบวนการเปลี่ยนรูปแบบ

2.2) การวิเคราะห์เชิงพรรณนา (Descriptive Analytics) เป็นเทคนิคที่ใช้สรุปลักษณะการใช้งานระบบลงทะเบียนเรียนออนไลน์ โดยอ้างอิงจากข้อมูล Log file เพื่อแสดงความถี่และแนวโน้มช่วงเวลาที่มีการใช้งานระบบสูง

2.2.1 การดึงข้อมูล (Retrieve Dataset) นำเข้าข้อมูล Log การใช้งานระบบ (Login Log) ที่ผ่านการทำความสะอาดแล้วเข้าสู่โปรแกรม RapidMiner เพื่อใช้เป็นฐานข้อมูลหลักในการวิเคราะห์พฤติกรรมกรเข้าใช้ระบบของนักศึกษา

2.2.2 การเปลี่ยนชื่อตัวแปร (Rename) ปรับเปลี่ยนชื่อตัวแปรบางรายการเพื่อป้องกันความซ้ำซ้อนของชื่อคอลัมน์ ก่อนทำการรวมชุดข้อมูลกับตารางอื่น โดยเฉพาะตัวแปรที่มีชื่อเหมือนกันแต่มีความหมายต่างกัน

2.2.3 การดึงข้อมูล Log การลงทะเบียน (Retrieve log\_regis) นำเข้าข้อมูล Log การลงทะเบียนเรียน ซึ่งประกอบด้วยข้อมูลการทำรายการลงทะเบียนของนักศึกษา เช่น รหัสนักศึกษา รหัสวิชา และเวลาที่ทำการ

2.2.4 การเปลี่ยนชื่อตัวแปรชุดที่สอง (Rename) ปรับชื่อคอลัมน์ของข้อมูลการลงทะเบียน เพื่อป้องกันปัญหาการชนกันของชื่อฟิลด์เมื่อทำการเชื่อมตาราง (Join)

2.2.5 การรวมชุดข้อมูล (Join) เชื่อมโยงข้อมูล Log การใช้งาน และ Log การลงทะเบียนเข้าด้วยกัน โดยใช้รหัสนักศึกษา (student\_id หรือ user\_id) เป็นกุญแจหลัก (Key) ทำให้ได้ข้อมูลที่แสดงพฤติกรรมการใช้งานและการลงทะเบียนของนักศึกษาอยู่ในชุดเดียวกัน

2.2.6 การเลือกตัวแปรที่เกี่ยวข้อง (Select Attributes) คัดเลือกเฉพาะตัวแปรที่จำเป็นต่อการวิเคราะห์ เช่น รหัสนักศึกษา, เวลาเข้าใช้ระบบ, เวลาลงทะเบียนเพื่อลดข้อมูลส่วนเกินและเพิ่มประสิทธิภาพในการประมวลผล

2.2.7 การลบข้อมูลซ้ำซ้อน (Remove Duplicates) ตรวจสอบและลบข้อมูลที่มีความซ้ำซ้อนออก เพื่อป้องกันการคำนวณค่าทางสถิติที่ผิดพลาด ซึ่งอาจเกิดจากการบันทึก Log ซ้ำ

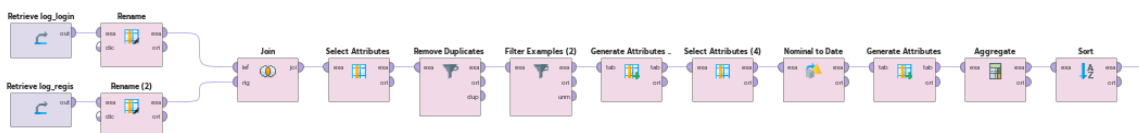
2.2.8 การลบข้อมูลซ้ำซ้อน (Remove Duplicates) ตรวจสอบและลบข้อมูลที่มีความซ้ำซ้อนออก เพื่อป้องกันการคำนวณค่าทางสถิติที่ผิดพลาด ซึ่งอาจเกิดจากการบันทึก Log ซ้ำ

2.2.9 การแปลงชนิดข้อมูล (Nominal to Date) แปลงข้อมูลวันและเวลาที่อยู่ในรูปแบบข้อความ (Nominal) ให้เป็นรูปแบบวันที่ (Date Type) เพื่อให้สามารถคำนวณหรือตั้งค่าเวลา เช่น ชั่วโมง หรือช่วงเวลา ได้อย่างถูกต้อง

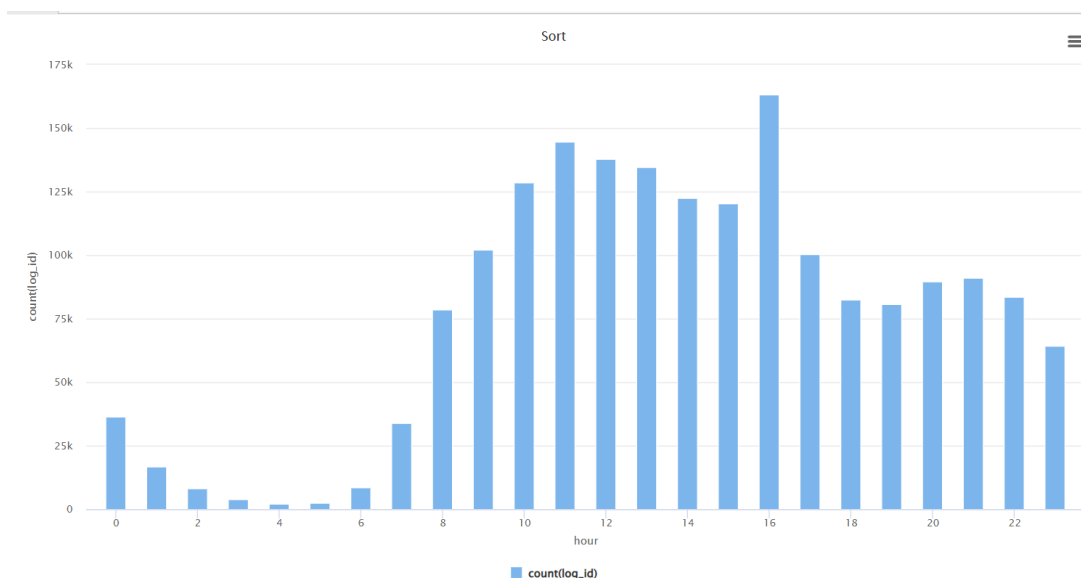
2.2.10 การสร้างตัวแปรใหม่ (Generate Attributes) สร้างตัวแปรเพิ่มเติมจากข้อมูลเดิม เช่น ชั่วโมงที่เข้าใช้งาน (login\_hour) จำนวนครั้งการเข้าใช้ระบบ (total\_login) จำนวนครั้งการลงทะเบียน (total\_regis) ขั้นตอนนี้เป็นเตรียมข้อมูลสำหรับการสรุปผลเชิงสถิติ

2.2.11 การสรุปผลข้อมูล (Aggregate) ใช้เครื่องมือ Aggregate เพื่อจัดกลุ่มข้อมูลตามรหัสนักศึกษา และคำนวณค่าทางสถิติ เช่น นับจำนวนครั้งการเข้าใช้ระบบ (Count login) นับจำนวนครั้งการลงทะเบียน (Count regis) คำนวณค่าเฉลี่ยช่วงเวลาเข้าใช้ระบบ  
ผลลัพธ์ที่ได้จะแสดงพฤติกรรมการใช้งานในระดับรายบุคคล

2.2.12 การเรียงลำดับข้อมูล (Sort) จัดเรียงข้อมูลตามค่าที่คำนวณได้ เช่น เรียงตามจำนวนครั้งการเข้าใช้ระบบจากมากไปน้อย เพื่อให้เห็นกลุ่มผู้ใช้งานหนักและผู้ใช้งานน้อยอย่างชัดเจน



ภาพที่ 3.15 แสดงเครื่องมือของโมเดลการวิเคราะห์เชิงพรรณนา



ภาพที่ 3.16 แสดงผลลัพธ์ของโมเดลการวิเคราะห์เชิงพรรณนา

2.3) K-Means Clustering เป็นเทคนิคการทำเหมืองข้อมูลแบบไม่ต้องสอน (Unsupervised Learning) ที่ใช้สำหรับจัดกลุ่มข้อมูลตามลักษณะความคล้ายคลึงกันของตัวแปร โดยไม่มีการกำหนดกลุ่มล่วงหน้า เหมาะสำหรับการค้นหารูปแบบพฤติกรรมการใช้งานระบบของผู้ใช้จากข้อมูลเชิงพฤติกรรม

2.3.1 การดึงข้อมูล (Retrieve Dataset) ขั้นตอนการนำเข้าข้อมูลดิบ 2 ชุด ได้แก่ บันทึกการเข้าสู่ระบบ (Log\_Login) และบันทึกการลงทะเบียน (Log\_Regis) เพื่อนำข้อมูลวัน-เวลา และรหัสผู้ใช้งานเข้าสู่กระบวนการวิเคราะห์พฤติกรรม

2.3.2 Nominal to Date ทำการแปลงตัวแปรวันที่และเวลาจากรูปแบบข้อความ (Nominal) ให้เป็นชนิดข้อมูลวันที่ (Date Type) เพื่อให้ระบบสามารถคำนวณและสกัดองค์ประกอบของเวลา เช่น ชั่วโมง ออกมาใช้งานได้อย่างแม่นยำ

2.3.3 Generate Attributes สร้างตัวแปรใหม่ (Features) จากข้อมูลเวลา เช่น การสกัด "ชั่วโมงที่เข้าใช้งาน" (login\_hour) เพื่อใช้เป็นตัวแปรหลักในการวิเคราะห์พฤติกรรมช่วงเวลาการใช้งานของนักศึกษา

2.3.4 Rename เปลี่ยนชื่อคอลัมน์ให้สื่อความหมายชัดเจน เช่น เปลี่ยนชื่อคอลัมน์ผลรวมและค่าเฉลี่ยต่าง ๆ ให้เป็นชื่อที่เข้าใจง่ายและพร้อมใช้ในขั้นตอนการวิเคราะห์ถัดไป

2.3.5 Aggregate ทำการสรุปกลุ่มข้อมูลตามรหัสผู้ใช้งาน (User ID) เพื่อคำนวณตัวชี้วัดพฤติกรรมรายบุคคล เช่น จำนวนครั้งที่เข้าใช้งานทั้งหมด (total\_login) และช่วงเวลาเฉลี่ยที่ใช้งาน (avg\_login\_hour) เป็นต้น

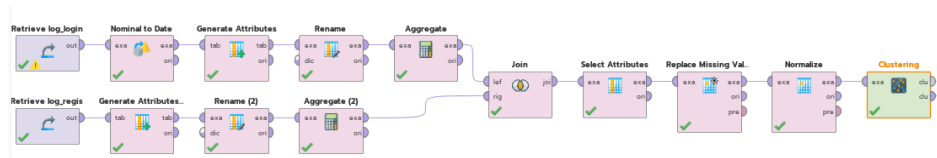
2.3.6 Join เป็นขั้นตอนการรวมข้อมูลจาก 2 แหล่ง (Login และ Regis) เข้าด้วยกันโดยใช้รหัสผู้ใช้งานเป็นตัวเชื่อม (Key) เพื่อให้ได้ชุดข้อมูลที่สมบูรณ์ซึ่งประกอบด้วยทั้งพฤติกรรมการเข้าใช้งานและพฤติกรรมการลงทะเบียนในตารางเดียว

2.3.7 Select Attributes ทำการคัดเลือกเฉพาะตัวแปรที่มีนัยสำคัญต่อการจัดกลุ่ม (Clustering) เช่น total\_login, avg\_login\_hour และตัวแปรพฤติกรรมการลงทะเบียน เพื่อลดความซับซ้อนและเพิ่มประสิทธิภาพให้แก่โมเดล

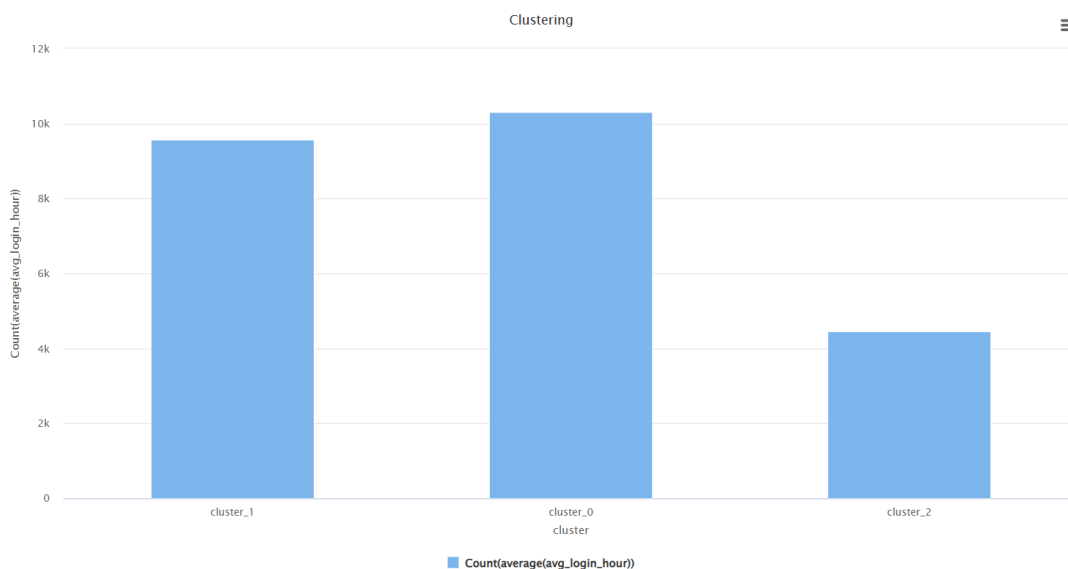
2.3.8 Replace Missing Values จัดการกับข้อมูลที่ขาดหายไป (เช่น นักศึกษาที่ login แต่ไม่ได้กดลงทะเบียน) โดยการแทนที่ด้วยค่าที่เหมาะสม (เช่น 0) เพื่อไม่ให้เกิดข้อผิดพลาดในการคำนวณของอัลกอริทึม

2.3.9 Normalize ปรับมาตรฐานค่าของตัวแปรทั้งหมดให้อยู่ในช่วงเดียวกัน (เช่น 0 ถึง 1) เพื่อป้องกันไม่ให้ตัวแปรที่มีค่าตัวเลขสูง (เช่น จำนวนครั้งการใช้งาน) มีอิทธิพลเหนือตัวแปรที่มีค่าตัวเลขน้อย (เช่น ชั่วโมง) ในขณะที่คำนวณระยะห่างระหว่างจุดข้อมูล

2.3.10 Clustering (K-Means) ประมวลผลด้วยอัลกอริทึม K-Means เพื่อจัดกลุ่มนักศึกษาออกเป็น 3 กลุ่ม (k=3) ตามความคล้ายคลึงของพฤติกรรม โดยระบบจะคำนวณจุดศูนย์กลางกลุ่ม (Centroid) และระบุหมายเลขกลุ่ม (Cluster ID) ให้กับนักศึกษาแต่ละราย



ภาพที่ 3.17 แสดงเครื่องมือของโมเดล K-Means Clustering



ภาพที่ 3.18 แสดงการแบ่งกลุ่ม Cluster

2.4 Association Rules เทคนิคการทำเหมืองข้อมูล (Data Mining) เพื่อค้นหาความสัมพันธ์หรือรูปแบบที่ซ่อนอยู่ในชุดข้อมูลขนาดใหญ่ โดยมักจะแสดงความสัมพันธ์ในรูปแบบ "ถ้าเกิดเหตุการณ์ A แล้ว มีโอกาสจะเกิดเหตุการณ์ B" (If A, then B)

2.4.1 Retrieve log\_login / Retrieve log\_regis นำเข้าข้อมูล Log ทั้งสองชุด ข้อมูลเข้าใช้ระบบ ข้อมูลการลงทะเบียน

2.4.2 Rename ปรับชื่อคอลัมน์เพื่อป้องกันการชนกันของชื่อ ก่อน Join

2.4.3 Join รวมข้อมูลสองชุดเข้าด้วยกันโดยใช้ Key เช่น user\_id ทำให้ได้ข้อมูลพฤติกรรมทั้งหมดอยู่ในตารางเดียว

2.4.4 Select Attributes เลือกเฉพาะตัวแปรที่เกี่ยวข้องกับพฤติกรรม เช่น is\_register is\_delete is\_error timestamp ตัดข้อมูลที่ไม่จำเป็นออก

2.4.5 Filter Examples กรองข้อมูลตามเงื่อนไขที่ต้องการศึกษา เช่น ภาคเรียน หรือปีการศึกษา

2.4.6 Remove Duplicates ลบข้อมูลซ้ำซ้อน ป้องกันการคำนวณผิดพลาด

2.4.7 Generate Attributes สร้างตัวแปรสรุป เช่น จำนวนครั้งที่ register จำนวนครั้งที่ delete จำนวน error

2.4.8 Nominal to Date แปลงข้อมูลเวลาให้อยู่ในรูปแบบ Date เพื่อให้สามารถคำนวณเชิงเวลาได้

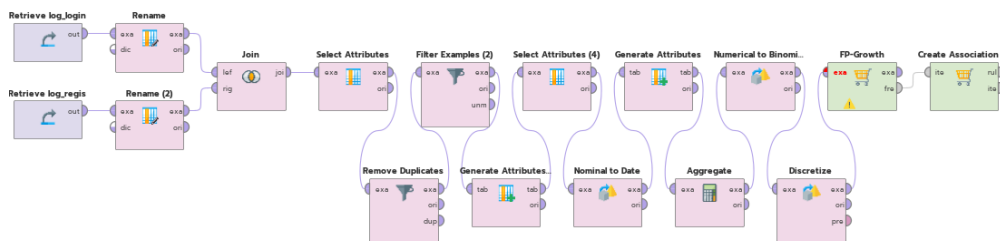
2.4.9 Aggregate รวมข้อมูลในระดับนักศึกษา เช่น sum(is\_register) sum(is\_delete) sum(is\_error) ตอนนี้อยู่ในรูปแบบ “หนึ่งแถวต่อหนึ่งนักศึกษา”

2.4.10 Discretize แปลงค่าตัวเลขให้เป็นช่วง (เช่น สูง/ต่ำ หรือ 0/1) Association Rules ใช้ข้อมูลแบบเชิงหมวดหมู่จึงต้องแปลงค่าตัวเลขก่อน

2.4.11 Numerical to Binominal แปลงค่าตัวแปรให้เป็น 0/1 (True/False) เพื่อให้ FP-Growth อ่านข้อมูลเป็นรายการสินค้า (Itemset)

2.4.12 FP-Growth ขั้นตอนนี้จะค้นหา Frequent Itemsets หาชุดพฤติกรรมที่เกิดร่วมกันบ่อยกว่าค่า Minimum Support ที่กำหนด ตัวอย่างเช่น [is\_delete, is\_error] เกิดร่วมกันบ่อย

2.4.13 Create Association Rules สร้างกฎความสัมพันธ์จาก Frequent Itemsets รูปแบบของกฎคือ Premise → Conclusion เช่น sum(is\_delete) → sum(is\_register)



ภาพที่ 3.19 แสดงเครื่องมือของโมเดล Association Rules

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Convict...
4	sum(is_register)	sum(is_delete)	0.814	0.814	0.907	-1.186	-0.000	1.000	1.000
5	sum(is_delete)	sum(is_register)	0.814	1.000	1.000	-0.814	-0.000	1.000	0.814
6	sum(is_delete), sum(is_error)	sum(is_register)	0.220	1.000	1.000	-0.220	-0.000	1.000	0.823
7	sum(is_error)	sum(is_register)	0.279	1.000	1.000	-0.279	0.000	1.000	1.045

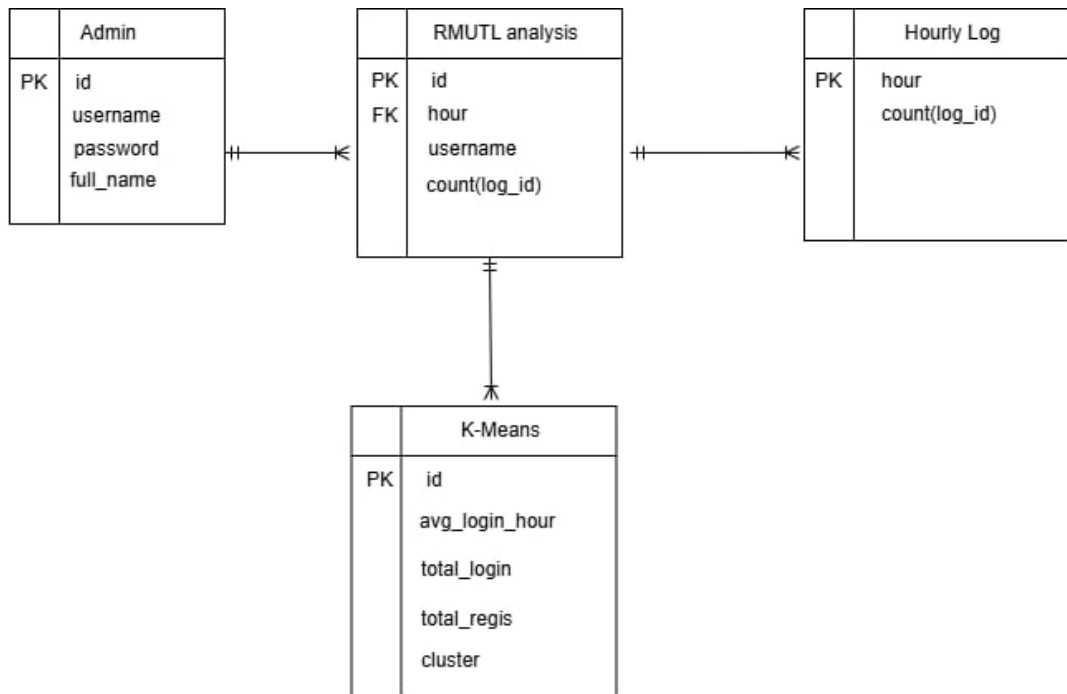
ภาพที่ 3.20 แสดงผลลัพธ์ของกฎความสัมพันธ์ Association Rules

3.1.5 การวัดประสิทธิภาพของโมเดล (Evaluation) ในขั้นตอนการประเมินประสิทธิภาพของโมเดล เป็นการตรวจสอบความถูกต้องและคุณภาพของผลลัพธ์ที่ได้จากเทคนิคการทำเหมืองข้อมูลทั้ง 3 รูปแบบ เพื่อให้มั่นใจว่าสารสนเทศที่ค้นพบสามารถนำไปตอบโจทยปัญหาการลงทะเบียนเรียนออนไลน์ได้อย่างแม่นยำ โดยการวิเคราะห์เชิงพรรณนา (Descriptive Analysis) จะถูกประเมินจากความสามารถในการสะท้อนภาพรวมของกิจกรรมในระบบได้อย่างชัดเจน โดยเฉพาะการระบุช่วงเวลาวิกฤต (Peak Time) ที่สอดคล้องกับพฤติกรรมการใช้งานจริงของนักศึกษา สำหรับการจัดกลุ่มข้อมูลด้วยเทคนิค K-Means Clustering จะประเมินจากความแตกต่างของลักษณะเด่นในแต่ละกลุ่ม (Cluster Profiles) ผ่านค่าเฉลี่ยของตัวแปรพฤติกรรม เช่น จำนวนครั้งและช่วงเวลาในการเข้าใช้งาน เพื่อยืนยันว่าการแบ่งกลุ่มนักศึกษาออกเป็น 3 กลุ่มนั้น มีความแตกต่างกันอย่างมีนัยสำคัญและสามารถแยกแยะกลุ่มพฤติกรรมได้อย่างเป็นรูปธรรม และในส่วนของ การหากฎความสัมพันธ์ (Association Rules) จะประเมินผ่านค่าตัวชี้วัดทางสถิติ ได้แก่ ค่าสนับสนุน (Support) ค่าความเชื่อมั่น (Confidence) และค่าดึงดูด (Lift) เพื่อพิสูจน์ความแข็งแรงของกฎความสัมพันธ์ที่ค้นพบว่ามีเหตุการณืที่เกิดขึ้นโดยบังเอิญ ซึ่งผลการประเมินทั้งหมดนี้ช่วยยืนยันว่าโมเดลที่เลือกใช้มีความเหมาะสมต่อการนำไปวิเคราะห์ปัญหาและสาเหตุการล่มของระบบได้อย่างมีประสิทธิภาพ

3.1.6 การนำโมเดลไปใช้งานจริง (Deployment) การนำแบบจำลองไปใช้งานเป็นขั้นตอนสุดท้ายที่เปลี่ยนผลลัพธ์เชิงข้อมูลให้กลายเป็นเครื่องมือสนับสนุนการตัดสินใจสำหรับฝ่ายทะเบียนและผู้บริหาร โดยมุ่งเน้นการนำเสนอผลที่ได้จาก Descriptive Analysis, K-Means Clustering และ Association Rules ไปประยุกต์ใช้ในรูปแบบของการออกแบบส่วนแสดงผลบนเว็บไซต์ (Web Visualization) เพื่อนำเสนอข้อมูลในรูปแบบ Dashboard ที่เข้าใจง่ายผ่านกราฟและสถิติต่าง ๆ ซึ่งจะช่วยให้ผู้บริหารสามารถคาดการณ์ช่วงเวลาที่มีความเสี่ยงต่อการล่มของระบบ และวางแผนสำรองทรัพยากรเครื่องแม่ข่าย (Server) ได้อย่างตรงจุดตามกลุ่มพฤติกรรมของนักศึกษา นอกจากนี้ กฎความสัมพันธ์ที่ค้นพบยังช่วยในการวางแผนจัดลำดับการเปิดลงทะเบียนรายวิชาเพื่อกระจายความหนาแน่นของการเข้าใช้งานพร้อมกัน โดยกระบวนการทั้งหมดนี้จะถูกรวบรวมและแสดงผลผ่านโครงสร้างเว็บไซต์ที่ได้ออกแบบไว้ เพื่อให้เกิดการนำ

ข้อมูลไปใช้งานจริงในการปรับปรุงประสิทธิภาพระบบลงทะเบียนเรียนออนไลน์ของมหาวิทยาลัยให้ดียิ่งขึ้น

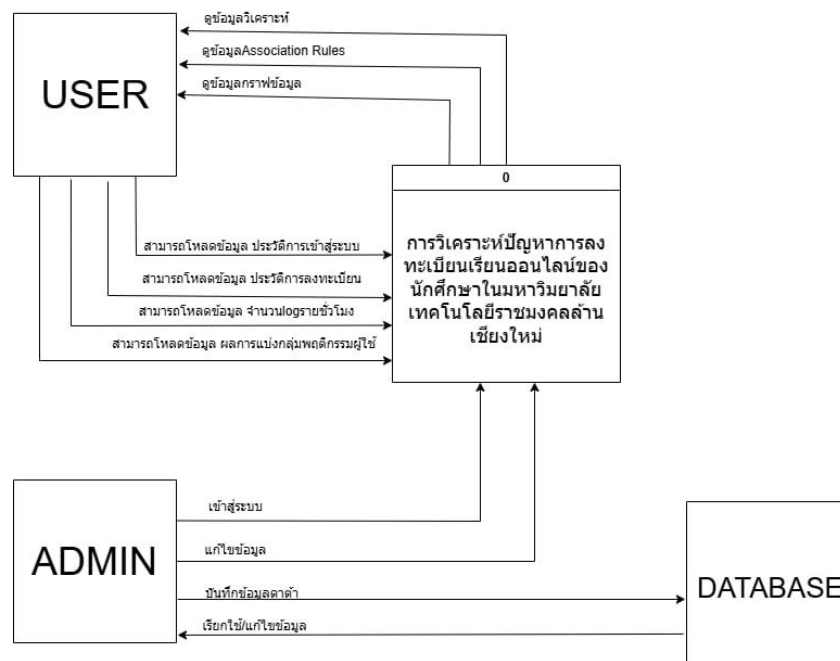
### 3.2 ER-Diagram



ภาพที่ 3.21 ภาพของ ER-Diagram

จากแผนภาพ โครงสร้างฐานข้อมูลของระบบประกอบด้วย 4 ตารางหลัก ได้แก่ ตาราง Admin สำหรับจัดเก็บข้อมูลบัญชีผู้ดูแลระบบ ตาราง RMUTL analysis สำหรับเก็บข้อมูลการวิเคราะห์การเข้าใช้งานระบบ ตาราง Hourly Log สำหรับสรุปจำนวนการเข้าใช้งานรายชั่วโมง และตาราง K-Means สำหรับจัดเก็บผลลัพธ์การจัดกลุ่มข้อมูล ข้อมูลการใช้งานรายชั่วโมงจากตาราง Hourly Log จะถูกเชื่อมโยงมายังตาราง RMUTL analysis เพื่อใช้ในการประมวลผลเชิงพรรณนา และนำผลลัพธ์ไปวิเคราะห์ต่อยังเทคนิค K-Means โดยผลการจัดกลุ่มจะถูกจัดเก็บไว้ในตาราง K-Means พร้อมระบุคุณลักษณะของข้อมูลแต่ละรายการ โครงสร้างดังกล่าวทำให้ระบบสามารถนำข้อมูลดิบมาวิเคราะห์ สรุปผล และจัดกลุ่มระดับความหนาแน่นของการใช้งานได้อย่างถูกต้อง พร้อมรองรับการแสดงผลผ่านหน้า Dashboard ได้อย่างเป็นระบบ

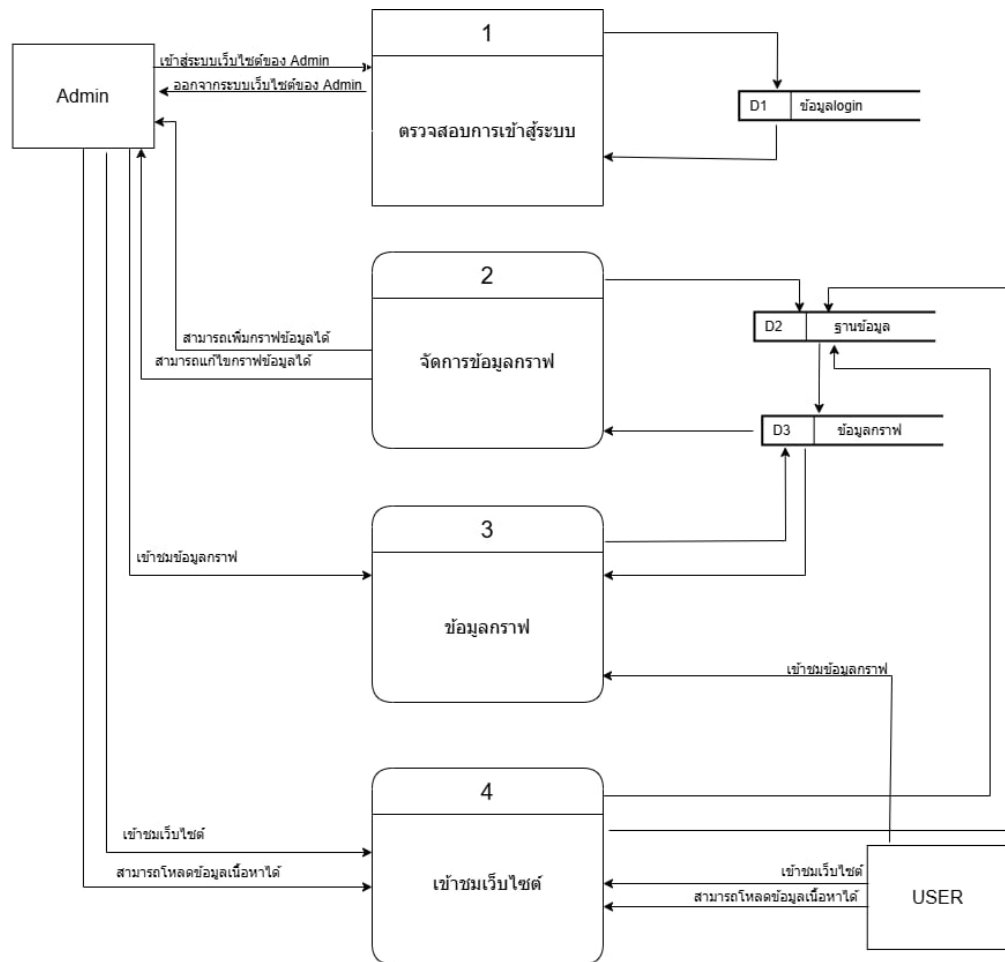
### 3.3 Context Diagram



ภาพที่ 3.22 ภาพของ Context Diagram

จากแผนภาพกระแสข้อมูล (Context Diagram) แสดงภาพรวมการทำงานของระบบวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ โดยมีผู้เกี่ยวข้อง 2 ส่วนหลัก ได้แก่ USER และ ADMIN ซึ่งมีสิทธิ์การใช้งานแตกต่างกัน USER สามารถเข้าสู่ระบบเพื่อดูผลการวิเคราะห์ ข้อมูลสถิติ กราฟแนวโน้ม และดาวน์โหลดรายงานได้ ส่วน ADMIN มีหน้าที่จัดการข้อมูล เช่น เพิ่มแก้ไข หรือลบข้อมูล และควบคุมการทำงานของระบบ ระบบทำหน้าที่รับคำสั่งจากผู้ใช้งานประมวลผลข้อมูล และเรียกใช้หรือบันทึกข้อมูลทั้งหมดผ่านฐานข้อมูล (Database) ซึ่งเป็นแหล่งจัดเก็บข้อมูลกลางของระบบ

### 3.4 Data Flow Diagram Level 0



ภาพที่ 3.23 ภาพของ Data Flow Diagram Level 0

แผนภาพ DFD level 0 ของระบบการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ของนักศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่

**ตารางที่ 3.1** แสดงคำอธิบายการประมวลผลของโปรเซส 1.0 : ตรวจสอบการเข้าสู่ระบบ

Process Description	
System	ระบบวิเคราะห์ข้อมูลการใช้งานระบบ ลงทะเบียนเรียนออนไลน์
DFD Number	1.0
Process Name	ตรวจสอบการเข้าสู่ระบบ
Input Data Flows	ข้อมูลการเข้าสู่ระบบของ Admin
Output Data Flows	สถานะการเข้าสู่ระบบของ Admin
Data Stores	D1 ข้อมูล Login
Description	โปรเซสตรวจสอบชื่อผู้ใช้และรหัสผ่านของผู้ดูแลระบบกับฐานข้อมูล เพื่ออนุญาตให้เข้าสู่ระบบและจัดการข้อมูลในเว็บไซต์

**ตารางที่ 3.2** แสดงคำอธิบายการประมวลผลของโปรเซส 2.0 : จัดการข้อมูลกราฟ

Process Description	
System	ระบบวิเคราะห์ข้อมูลการใช้งานระบบ ลงทะเบียนเรียนออนไลน์
DFD Number	2.0
Process Name	จัดการข้อมูลกราฟ
Input Data Flows	คำสั่งเพิ่ม แก้ไข และลบข้อมูลกราฟจาก Admin
Output Data Flows	ข้อมูลกราฟที่ถูกปรับปรุงแล้ว
Data Stores	D2 ฐานข้อมูล, D3 ข้อมูลกราฟ
Description	โปรเซสผู้ดูแลระบบสามารถเพิ่ม แก้ไข และลบข้อมูลกราฟ โดยระบบจะบันทึกและปรับปรุงข้อมูลในฐานข้อมูล

**ตารางที่ 3.3** แสดงคำอธิบายการประมวลผลของโปรเซส 3.0 : ข้อมูลกราฟ

Process Description	
System	ระบบวิเคราะห์ข้อมูลการใช้งานระบบ ลงทะเบียนเรียนออนไลน์
DFD Number	3.0
Process Name	ข้อมูลกราฟ
Input Data Flows	คำสั่งเรียกดูข้อมูลกราฟ
Output Data Flows	ข้อมูลกราฟสำหรับการแสดงผล
Data Stores	D3 ข้อมูลกราฟ
Description	โปรเซสการดึงข้อมูลกราฟจากฐานข้อมูล เพื่อนำมาแสดงผลบนเว็บไซต์ เพื่อให้ผู้ใช้งาน สามารถดูและวิเคราะห์ข้อมูลได้สะดวก

**ตารางที่ 3.4** แสดงคำอธิบายการประมวลผลของโปรเซส 4.0 : เข้าชมเว็บไซต์

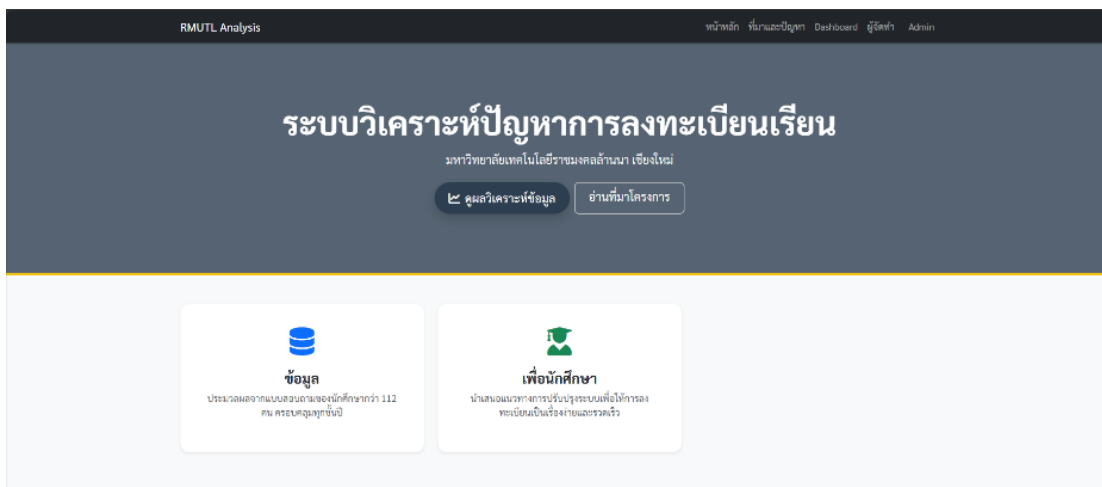
Process Description	
System	ระบบวิเคราะห์ข้อมูลการใช้งานระบบ ลงทะเบียนเรียนออนไลน์
DFD Number	4.0
Process Name	เข้าชมเว็บไซต์
Input Data Flows	คำสั่งเข้าชมเว็บไซต์จากผู้ใช้งาน
Output Data Flows	ข้อมูลสำหรับการแสดงผลหน้าเว็บไซต์
Data Stores	D3 ข้อมูลกราฟ
Description	โปรเซสในการให้ผู้ใช้งานเข้าชมเว็บไซต์ โดย ระบบจะรับคำสั่งการเข้าชมเว็บไซต์และดึง ข้อมูลที่เกี่ยวข้องจากฐานข้อมูลมาแสดงผล เช่น ข้อมูลกราฟหรือข้อมูลสรุปต่าง ๆ ที่ใช้ ในการวิเคราะห์ข้อมูล

### 3.5 การออกแบบและพัฒนาเว็บไซต์

การวางแผนการจัดลำดับเนื้อหาสาระของเว็บไซต์ ออกเป็นหมวดหมู่ เพื่อจัดทำเป็นโครงสร้างในการจัดวางหน้าเว็บเพจทั้งหมด ทำให้เห็นโครงสร้างทั้งหมดของเว็บไซต์และการออกแบบโครงสร้างหรือจัดระเบียบของข้อมูลที่ชัดเจน แยกย่อยเนื้อหาออกเป็นส่วนต่าง ๆ ที่สัมพันธ์กัน และให้อยู่ในมาตรฐานเดียวกัน จะช่วยให้นำมาใช้งานและง่ายต่อการเข้าอ่านเนื้อหาของผู้ใช้เว็บไซต์

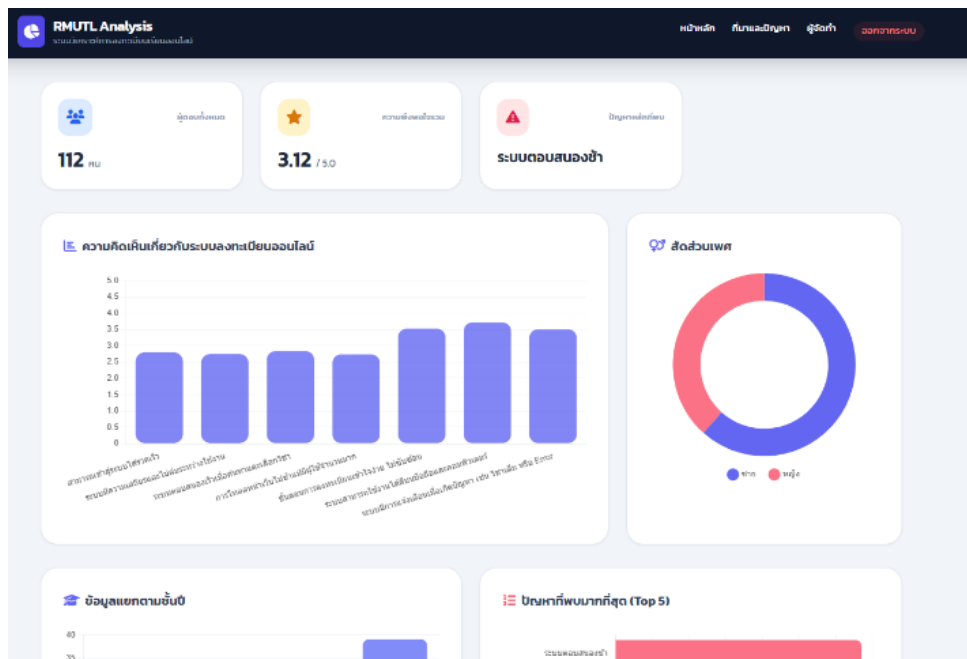
#### 3.4.1 การออกแบบ Wireframe หน้าจอเว็บไซต์

1) หน้าแรก of เว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บและเนื้อหาที่เกี่ยวข้องกับข้อมูลการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ เช่น กราฟ เนื้อหาต่าง ๆ



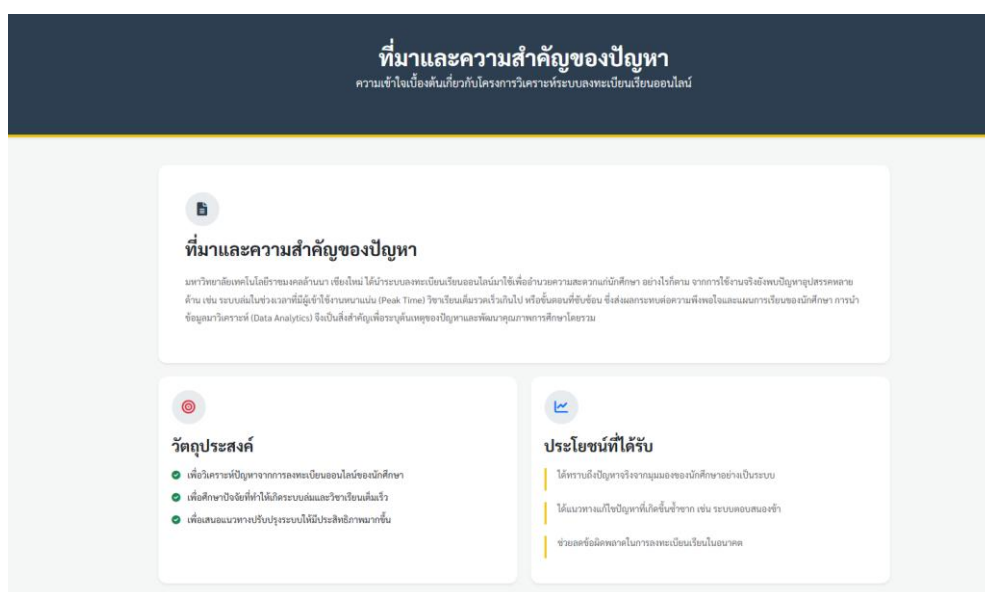
ภาพที่ 3.24 แสดงโครงร่างการออกแบบหน้าแรกของเว็บไซต์

2) หน้าแสดง Dashboard ของข้อมูลมูลค่าการส่งออกของการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์



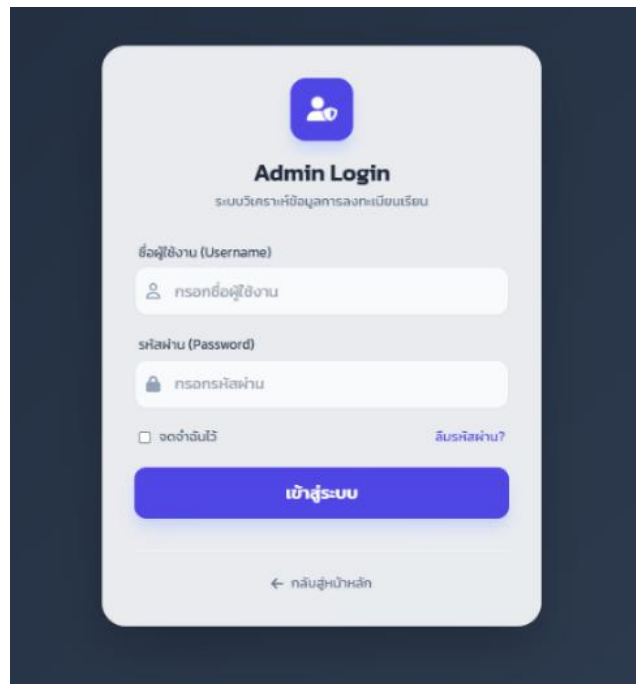
ภาพที่ 3.25 แสดงโครงร่างการออกแบบหน้าแสดงผลข้อมูล

3) หน้าแสดงผลข้อมูลที่มาของการวิเคราะห์ข้อมูลการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์



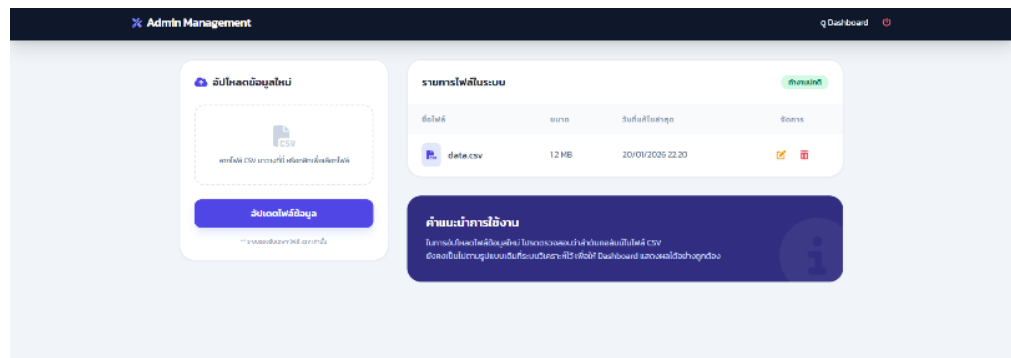
ภาพที่ 3.26 แสดงโครงร่างการออกแบบหน้าแสดงผลที่มาของการวิเคราะห์ข้อมูล

4) หน้าแสดงผลในส่วนของการเข้าสู่ระบบเพื่อไปจัดการกับข้อมูลภายในเว็บไซต์



ภาพที่ 3.27 แสดงโครงร่างการออกแบบในส่วนของหน้าการเข้าสู่ระบบ

5) หน้าแสดงผลในส่วนของ Data Set เป็นหน้าแสดงชุดข้อมูลที่ใช้ ในการวิเคราะห์ข้อมูล



ภาพที่ 3.28 แสดงโครงร่างการออกแบบในส่วนของ Data Set

### 3.6 บทสรุป

การศึกษาในบทนี้ได้อธิบายถึงขั้นตอนและวิธีการดำเนินโครงการอย่างเป็นระบบ เริ่มตั้งแต่การกำหนดประชากรและกลุ่มตัวอย่างที่เป็นนักศึกษาผู้ใช้งานระบบลงทะเบียนเรียนออนไลน์ เครื่องมือที่ใช้ในการวิจัยทั้งด้านฮาร์ดแวร์และซอฟต์แวร์สำหรับการวิเคราะห์ข้อมูล ไปจนถึงวิธีการเก็บรวบรวมข้อมูลบันทึกการใช้งาน (Log Files) เพื่อนำมาประมวลผล โดยมีหัวใจสำคัญคือการประยุกต์ใช้แนวคิดกระบวนการวิเคราะห์ข้อมูลตามกรอบมาตรฐาน CRISP-DM เป็นแนวทางหลักในการวางลำดับขั้นตอนการดำเนินงาน การวิเคราะห์ข้อมูลมุ่งเน้นการบูรณาการเทคนิคเหมืองข้อมูล (Data Mining) 3 รูปแบบหลัก ได้แก่ การวิเคราะห์เชิงพรรณนา (Descriptive Analysis) เพื่อสะท้อนสภาพการณ์และปัญหาการใช้งานระบบในปัจจุบัน การจัดกลุ่มข้อมูล (K-Means Clustering) เพื่อจำแนกพฤติกรรมนักศึกษา และการหากฎความสัมพันธ์ (Association Rules) เพื่อค้นหารูปแบบการลงทะเบียนที่ซ่อนอยู่ ซึ่งผลลัพธ์ที่ได้จากโมเดลเหล่านี้ จะถูกนำไปต่อยอดสู่การออกแบบและพัฒนาส่วนแสดงผลบนเว็บไซต์ (Web Visualization) เพื่อใช้เป็นเครื่องมือสำคัญในการสนับสนุนการตัดสินใจของฝ่ายทะเบียนและผู้บริหาร ในการวางแผนปรับปรุงระบบลงทะเบียนเรียนออนไลน์ให้มีประสิทธิภาพและลดความเสี่ยงในการล่มของระบบได้อย่างยั่งยืน