

# บทที่ 1

## บทนำ

### 1.1 ความสำคัญและที่มาของปัญหา

ปัจจุบันพฤติกรรมการซื้อสินค้าออนไลน์มีความซับซ้อนและเปลี่ยนแปลงอย่างรวดเร็ว ลูกค้าไม่ได้เพียงแค่เข้าชมสินค้า แต่ยังมีกระบวนการตัดสินใจหลายขั้นตอน เช่น การเปรียบเทียบราคา การเพิ่มสินค้าไว้ในตะกร้า และการซื้อจริง ซึ่งข้อมูลเหล่านี้หากไม่มีการนำมาวิเคราะห์อย่างเป็นระบบ ธุรกิจอาจไม่สามารถเข้าใจความต้องการที่แท้จริงของลูกค้าได้ หนึ่งในปัญหาที่ธุรกิจอีคอมเมิร์ซพบเจอคือ ลูกค้าซื้อเพียงครั้งเดียวแล้วไม่กลับมาซื้อซ้ำ หรือไม่สามารถแยกแยะได้ว่าลูกค้ากลุ่มใดมีแนวโน้มจะเป็นลูกค้าประจำ หากไม่สามารถคาดการณ์หรือทำความเข้าใจพฤติกรรมลูกค้าได้อย่างแม่นยำ ธุรกิจจะสูญเสียโอกาสในการสร้างความสัมพันธ์ระยะยาวและต้องแบกรับต้นทุนในการหาลูกค้าใหม่ที่สูงขึ้น

ดังนั้นจึงมีความจำเป็นที่จะต้องใช้กระบวนการวิเคราะห์ข้อมูลเชิงลึกเพื่อนำไปสู่การจัดกลุ่มลูกค้าและการวิเคราะห์พฤติกรรมการซื้อซ้ำ เครื่องมืออย่าง RFM Analysis และ Machine Learning สามารถช่วยให้เข้าใจความแตกต่างของลูกค้าได้อย่างชัดเจนและมีแบบจำลองที่แม่นยำมากขึ้น

โครงการนี้จึงเลือกใช้ชุดข้อมูลจาก Kaggle เรื่อง E-commerce Behavior Data from Multi-category Store ซึ่งบันทึกพฤติกรรมผู้ใช้ เช่น การดูสินค้า การเพิ่มลงตะกร้า และการสั่งซื้อ เพื่อนำมาวิเคราะห์ด้วยเทคนิค RFM, การจัดกลุ่ม (Clustering) กฎความสัมพันธ์ (Association Rules) และต้นไม้ตัดสินใจ (Decision Tree) เพื่อหาความเข้าใจเชิงลึกเกี่ยวกับลูกค้า และสามารถประยุกต์ใช้ต่อยอดในเชิงกลยุทธ์ทางการตลาดและการสร้างความภักดีของลูกค้าได้อย่างเหมาะสม

จากการวิเคราะห์ข้อมูลพฤติกรรมเหล่านี้ ผู้จัดทำจะนำเทคนิค RFM (Recency, Frequency, Monetary) มาคำนวณค่าสำหรับแต่ละผู้ใช้ แล้วทำการจัดกลุ่มลูกค้าด้วยเทคนิค K-Means Clustering และ Decision Tree เพื่อแยกกลุ่มลูกค้าตามลักษณะพฤติกรรม และใช้กฎความสัมพันธ์ (Association Rules) เพื่อค้นหารูปแบบหรือพฤติกรรมการซื้อพร้อมกันของลูกค้า ซึ่งอาจนำไปสู่การพัฒนาแนวทางในการแนะนำสินค้า (Product Recommendation) ได้อย่างเหมาะสมยิ่งขึ้น เพื่อวิเคราะห์พฤติกรรมการซื้อของลูกค้า และใช้ในการตัดสินใจเชิงกลยุทธ์

หรือการสร้างความภักดีของลูกค้านอกจากนี้ยังสามารถแสดงผลการวิเคราะห์ในรูปแบบกราฟหรือแดชบอร์ดโดยใช้ Visualization เช่น Power BI พร้อมนำเสนอผ่าน Web Browser เพื่อให้ผู้ใช้งานทั่วไปสามารถเข้าถึงได้สะดวก

## 1.2 วัตถุประสงค์

1.2.1 เพื่อดูพฤติกรรมกรรมการซื้อของลูกค้าโดยใช้เทคนิค RFM (Recency, Frequency, Monetary) , K-Means, Association Rules, Decision Tree ในการจัดกลุ่มลูกค้าตามระดับความสัมพันธ์ทางธุรกิจ

1.2.2 เพื่อหาโมเดลในการวิเคราะห์พฤติกรรมกรรมการซื้อของลูกค้าจะมีการมาซื้อในอนาคตด้วยเทคนิค Machine Learning

1.2.3 เพื่อนำแบบจำลองที่ได้จากการวิเคราะห์มาประยุกต์ใช้วิเคราะห์พฤติกรรมกรรมการซื้อของลูกค้า

## 1.3 ประโยชน์ที่จะได้รับ

1.3.1 ได้พฤติกรรมกรรมการซื้อของลูกค้าโดยใช้เทคนิค RFM (Recency, Frequency, Monetary) K-Means, Association Rules, Decision Tree

1.3.2 ได้โมเดลในการวิเคราะห์พฤติกรรมกรรมการซื้อของลูกค้า

1.3.3 ได้แบบจำลองที่ได้จากการวิเคราะห์มาประยุกต์ใช้วิเคราะห์พฤติกรรมกรรมการซื้อของลูกค้า

## 1.4 ขอบเขต

1.4.1 ขอบเขตวิเคราะห์ข้อมูล

1.4.1.1 วิเคราะห์ข้อมูลได้รวบรวมข้อมูลประวัติการซื้อสินค้าที่ได้จากเว็บไซต์ Kaggle (<https://www.kaggle.com/>) เพื่อนำมาวิเคราะห์ ข้อมูลที่ได้นำมาวิเคราะห์ ได้แก่ หมวดหมู่สินค้า ยี่ห้อสินค้า ราคา ประเภท

1.4.1.2 วิเคราะห์ข้อมูลทำการกลั่นกรองข้อมูล (Data Cleaning) ข้อมูลที่ทำการบวกรวมนี้ อาจจะต้องมีการทำข้อมูลให้ถูกต้อง เช่น ข้อมูลที่ผิดปกติแบบ เต็มข้อมูลที่หายไปให้ครบ ให้ข้อมูลมีความสมบูรณ์ ตัดข้อมูลที่ไม่น่าเป็นออก เป็นต้น

1.4.1.3 วิเคราะห์ทำนายว่าบุคคลแต่ละบุคคลมีโอกาสจะกลับมาซื้อสินค้าอีกหรือไม่ โดยใช้ เทคนิคการทำเหมืองข้อมูล (Data Mining) และนำโปรแกรม Rapid Miner เข้ามาช่วยในการ ทำนายผล ซึ่งมีวิธี ดังนี้

- 1) ใช้กระบวนการ RFM Analysis ในการวิเคราะห์ข้อมูล
- 2) ใช้เทคนิค Machine Learning ประเภทการจำแนกกลุ่ม (Clustering) เช่น RFM Analysis, K-Means, Association Rules, Decision Tree โดยใช้โปรแกรม RapidMiner ในการทดลองเพื่อการวิเคราะห์พฤติกรรมการณ์ซื้อของลูกค้า

#### 1.4.2 ขอบเขตผู้ใช้ทั่วไป

1.4.2.1 ทำระบบกรอกข้อมูลใส่ข้อมูลไปว่าลูกค้าที่อยู่ในที่มีค่า RFM นี้ควรจะอยู่จะ ถูกจัดอยู่ในกลุ่มประเภทลูกค้าอะไร และควรมีวิธีการอย่างไรในการดูแลลูกค้ากลุ่มนี้

1.4.2.2 สามารถแสดง Dashboard ให้ผู้ใช้ได้ดูได้ เช่น กราฟ Pie Chart

1.4.2.3 สามารถดาวน์โหลดข้อมูลเพื่อนำไปศึกษาต่อได้

#### 1.4.3 ขอบเขตผู้ดูแลระบบ

1.4.3.1 สามารถเพิ่ม ลบ แก้ไขข้อมูลบนเว็บไซต์ได้

1.4.3.2 เข้าสู่ระบบได้โดยใช้ไอดีและรหัสผ่าน

1.4.3.3 สามารถอัปเดตข้อมูลบนเว็บไซต์ได้

#### 1.4.4 ชุดข้อมูล

	A	B	C	D	E	F	G	H	I	J	K	L
1	event_time	event_type	product_id	category_id	category_code	brand	price	user_id	user_session			
2	2019-11-01 00:00	view	1003461	20530135556318826	electronics.smartphone	xiaomi	489.07	520088904	4d3b30da-a5e4-49df-b1a8-be5943f1dd33			
3	2019-11-01 00:00	view	5000088	2053013566100866c	appliances.sewing_machine	janome	293.65	530496790	8e5ff4f83-366c-4f70-860e-ca7417414283			
4	2019-11-01 00:00	view	17302664	2053013553853497655		creed	28.31	561587266	755422e7-9040-477b-9bd2-6a6e8fd97387			
5	2019-11-01 00:00	view	3601530	20530135638107759	appliances.kitchen.washer	lg	712.87	518085591	31bf58dc-7892-48cc-8020-2f17e6de6e7f			
6	2019-11-01 00:00	view	1004775	20530135556318826	electronics.smartphone	xiaomi	183.27	558856683	313628f1-6808-460d-84fe-cec7a8796ef2			
7	2019-11-01 00:00	view	1306894	20530135589202171	computers.notebook	hp	360.09	520772685	816a59f3-f5ae-4ccd-9b23-82aa8c23d33c			
8	2019-11-01 00:00	view	1306421	20530135589202171	computers.notebook	hp	514.56	514028527	df8184cc-3694-4549-8c8c-6b5171877376			
9	2019-11-01 00:00	view	15900065	2053013558190408249		rondell	30.86	518574284	5e6ef132-4d7c-4730-8c7f-85aa4082588f			
10	2019-11-01 00:00	view	12708937	205301355359896355		michelin	72.72	532364121	0a899268-31eb-46de-898d-09b2da950b24			
11	2019-11-01 00:00	view	1004258	20530135556318826	electronics.smartphone	apple	732.07	532647354	d2d3d2c6-631d-489e-9f5d-06f340b85be0			
12	2019-11-01 00:00	view	17200570	20530135597926324	furniture.living_room.sofa		437.33	518780843	aa806835-b14c-45af-9538-cd4d1849be3e			
13	2019-11-01 00:00	view	2701517	20530135639114392	appliances.kitchen.refrigerators		155.11	518427361	c89b0d96-247f-4044-9c91-bb5f38c6af9b			
14	2019-11-01 00:00	view	16700260	20530135599016843	furniture.kitchen.chair		31.64	566255262	173d7b72-1db7-4638-8207-be8148bf3c9f			
15	2019-11-01 00:00	view	34600011	2060981320581906480			20.54	512416379	4dfe2c67-e537-4dc2-ae69-0b65263db091			
16	2019-11-01 00:00	view	4600658	20530135639449936	appliances.kitchen.dishwasher	samsung	411.83	526595547	aab33a9a-29c3-4d50-84c1-8a2bc9256104			
17	2019-11-01 00:00	view	24900193	2053013562183385881			1.09	512651494	f603c815-f51a-46fe-9404-cb58e35edaaf			
18	2019-11-01 00:00	view	27400066	2053013563391345499			8.55	551061950	3f6112f1-5695-4e88-bb0a-49f9e36658ff			
19	2019-11-01 00:00	view	5100503	2053013553375346967		xiaomi	22.68	520037415	f54f896a-f3f2-43ac-99a4-fcb2449dd36			
20	2019-11-01 00:00	view	1004566	20530135556318826	electronics.smartphone	huawei	164.84	566265908	52c2c76c-b79e-4794-88ff-badc76d35f5a			
21	2019-11-01 00:00	view	1307115	20530135589202171	computers.notebook	hp	411.59	514028527	df8184cc-3694-4549-8c8c-6b5171877376			
22	2019-11-01 00:00	view	12719154	205301355359896355		yokohama	88.81	512376444	f7e8ffaa-5103-47f3-90ab-c30f313770c0			
23	2019-11-01 00:00	view	1004708	20530135556318826	electronics.smartphone	huawei	151.99	566143627	aa610ab3-5c60-4551-8a2b-8c999ddd8f43			
24	2019-11-01 00:00	view	26022534	2053013562837697343			7.07	566280860	341a87d8-8cf4-4b4a-8b7d-749380697af1			
25	2019-11-01 00:00	view	17200558	20530135597926324	furniture.living_room.sofa		591.78	566280536	34e7a822-0b93-4a9a-8e0a-bae4cadb21d8			
26	2019-11-01 00:00	view	17300016	2053013553853497655		trussardi	55.33	563769246	95ab5d5d-be19-4118-84e9-534d65b7a687			
27	2019-11-01 00:00	view	26019863	2053013562837697343			11.79	566280860	341a87d8-8cf4-4b4a-8b7d-749380697af1			

ภาพที่ 1.1 แสดงตัวอย่างข้อมูลข้อมูลพฤติกรรมอีคอมเมิร์ซการซื้อสินค้าอิเล็กทรอนิกส์จากร้านค้าในช่วงเดือนพฤศจิกายน ปี 2562

จาก Dataset ข้อมูลพฤติกรรมอีคอมเมิร์ซจากร้านค้าหลายหมวดหมู่นี้ ก่อนทำการ  
กลั่นกรองข้อมูล จำนวนข้อมูล 1,048,576 แถว จำนวนคอลัมน์ 9 คอลัมน์ สามารถอธิบายชุด  
ข้อมูลได้ดังนี้

**ตารางที่ 1.1** ตารางแสดงคอลัมน์และความหมายของชุดข้อมูล

คอลัมน์	ความหมาย
date	วันที่
time	เวลาที่เหตุการณ์เกิดขึ้นที่ (ใน UTC)
event_type	กิจกรรมประเภท
product_id	รหัสของผลิตภัณฑ์
category_id	รหัสหมวดหมู่ของผลิตภัณฑ์
category_code	หมวดหมู่ของผลิตภัณฑ์ (ชื่อรหัส)
category_code_other1	หมวดหมู่ของผลิตภัณฑ์ย่อย 1 (ชื่อรหัส)
category_code_other2	หมวดหมู่ของผลิตภัณฑ์ย่อย 2 (ชื่อรหัส)
category_code_other3	หมวดหมู่ของผลิตภัณฑ์ย่อย 3 (ชื่อรหัส)
brand	ชื่อแบรนด์
price	ราคาของสินค้า ปัจจุบัน
user_id	รหัสผู้ใช้ถาวร
user_session	รหัสเซสชันของผู้ใช้ชั่วคราว เหมือนกัน สำหรับเซสชันของผู้ใช้แต่ละคน มีการ เปลี่ยนแปลงทุกครั้งที่ใช้กลับมาที่ร้านค้า ออนไลน์จากการหยุดชั่วคราวเป็นเวลานาน

หลังจากทำการกลั่นกรองข้อมูลแล้วเหลือจำนวนข้อมูล 653,283 แถว จำนวนคอลัมน์  
9 คอลัมน์ โดยคอลัมน์ที่ข้อมูลขาดหายและทำการกลั่นกรองข้อมูลคือ category\_code, brand

## 1.5 เครื่องมือที่ใช้พัฒนาระบบ

### 1.5.1 Hardware

1.5.1.1 โน้ตบุ๊ก Asus CPU Intel(R) Core i5 3.10 GHz Ram 16 GB SSD 512 GB

### 1.5.2 Software

1.5.2.1 โปรแกรม Rapid Miner ใช้ในการสร้างโมเดลการวิเคราะห์ข้อมูลต่างๆ

1.5.2.2 โปรแกรม Apache เพื่อจำลองเซิร์ฟเวอร์

1.5.2.3 โปรแกรม Visual Studio Code ใช้ในการสร้างเว็บไซต์

1.5.2.4 Microsoft Office Word 2019 ใช้ในการทำเอกสาร

1.5.2.5 Microsoft Office Excel 2019 ใช้ในการจัดการชุดข้อมูล

1.5.2.6 ระบบปฏิบัติการ Window 11 ใช้ในการเปิดซอฟต์แวร์ต่างๆ

## 1.6 สถานที่ใช้ในการดำเนินการศึกษา และรวบรวมข้อมูล

1.6.1 มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาเชียงใหม่ 128 ถนนห้วยแก้ว ตำบล  
ช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

## 1.7 ระยะเวลาในการดำเนินการ

ตารางที่ 1.2 ตารางแสดงระยะเวลาดำเนินงาน

แผนการดำเนินการ	2568							2569		
	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.
1. ศึกษาและกำหนดความต้องการ	←→									
2. วิเคราะห์ออกแบบระบบและสร้างฐานข้อมูล		←→								
3. เขียนและทดสอบโปรแกรม			←→							
4. ติดตั้ง ทดสอบ และปรับปรุงระบบ					←→					
5. ตรวจสอบระบบโดยรวม							←→			
6. ประเมินการใช้งานระบบ								←→		
7. จัดทำคู่มือการใช้งาน			←→							
8. จัดทำเอกสารประกอบโครงการ	←→									

## 1.8 นิยามศัพท์เฉพาะ

1.8.1 RFM Analysis หมายถึง เทคนิคการวิเคราะห์ข้อมูลลูกค้า โดยพิจารณาจาก 3 ปัจจัยหลัก ได้แก่ Recency (R) Frequency (F) และ Monetary (M) เพื่อแบ่งกลุ่มลูกค้าและทำความเข้าใจพฤติกรรมการซื้อของลูกค้าได้ดียิ่งขึ้น

1.8.2 Machine Learning หมายถึง สาขาหนึ่งของปัญญาประดิษฐ์ (AI) ที่ทำให้คอมพิวเตอร์สามารถเรียนรู้และปรับปรุงตัวเองจากข้อมูลได้โดยไม่ต้องตั้งโปรแกรมไว้ล่วงหน้า

1.8.3 Recency (R) หมายถึง ความใหม่หรือช่วงเวลาของลูกค้าซื้อสินค้าหรือบริการครั้งล่าสุด

1.8.4 Frequency (F) หมายถึง ความถี่ในการซื้อสินค้าหรือบริการของลูกค้าในช่วงเวลาหนึ่ง

1.8.5 Monetary (M) หมายถึง มูลค่าการใช้จ่ายทั้งหมดของลูกค้าในช่วงเวลาหนึ่ง