

## ปฏิบัติการที่ 5

### การใช้เครื่องมือการวิเคราะห์ตะกร้าตลาด

#### วัตถุประสงค์

1. เพื่อให้สามารถใช้เครื่องมือวิเคราะห์ตะกร้าตลาด (Market Basket Analysis) ในการหา frequent item set ได้
2. เพื่อให้สามารถใช้เครื่องมือวิเคราะห์ตะกร้าตลาด (Market Basket Analysis) ในการหาความสัมพันธ์ (Association Rule) ได้

#### 1. ชุดข้อมูลปฏิบัติการ

- ชุดข้อมูล Mini Market Basket (สำหรับการสาธิต)
- ชุดข้อมูล Simulated Online Shopping Carts (สำหรับการฝึกปฏิบัติการ)

#### 2. ขั้นตอนปฏิบัติการ

ขั้นตอนปฏิบัติการ มีดังนี้

1. นำเข้าชุดข้อมูล Mini Market Basket จากแฟ้มข้อมูล mini\_market\_basket.csv ตั้งชื่อชุดข้อมูลเป็น mini market basket
2. ทำการสร้างการทดลอง โดยกำหนดชื่อการทดลองเป็น “Practice 5”
3. นำชุดข้อมูล mini market basket เข้าสู่การทดลองโดยลากโมดูลชุดข้อมูลซึ่งที่อยู่ภายใต้ Saved Datasets → My Datasets ในหน้าต่างย่อย Modules มาวางบน Workspace
4. หา Frequent Item Sets ในชุดข้อมูล mini market basket โดยลากโมดูล Discover Association Rules (ภายใต้ Custom หากไม่พบโมดูลดังกล่าวให้ทำการนำเข้ามาจาก Azure AI Gallery ดูเนื้อหาหัวข้อ การนำเข้าโมดูลจาก Azure AI Gallery) มาวางยัง workspace และเพิ่มข้อความอธิบาย (Comment) โดยการคลิกขวาที่โมดูล เลือก Edit Comment พิมพ์ข้อความอธิบายเป็น “Find Frequent Item Sets”
5. นำข้อมูลส่งออกจากโมดูลข้อมูล mini market basket เป็นข้อมูลนำเข้า Dataset ของโมดูล Discover Association Rules
6. คลิกที่กล่องโมดูล Discover Association Rules ที่หน้าต่างย่อย Properties จะปรากฏฟอร์มสำหรับกำหนดพารามิเตอร์ของโมดูล (สามารถศึกษาเพิ่มเติมได้ที่เว็บไซต์ <https://gallery.azure.ai/CustomModule/Discover-Association-Rules-1>)

24. กำหนด Input Dataset Type เป็น Items list แล้วคลิก Launch column selector เพื่อเลือกตัวแปรที่จัดเก็บรายการสินค้าที่ต้องการประมวลผล ในที่นี้เลือกตัวแปร items แล้วคลิก ✓
7. กำหนดค่าพารามิเตอร์ ดังนี้
- Minimal Support: 0.1
  - Minimal Confidence: 0.1
  - Minimal Number of Items in a Rule: 1
  - Maximal Number of Items in a Rule: 5
  - Return type: Frequent Itemsets
8. คลิก **RUN** เพื่อทำการประมวลผล แล้วดูผลลัพธ์จากข้อมูลออกของโมดูล Discover Association Rules โดยคลิกที่โหนดส่วนต่อประสานข้อมูลออก แล้วเลือก Visualize จะปรากฏผลลัพธ์ดังรูป

Practice 5 > Discover Association Rules > Rules

rows	columns	Frequent Item Sets		
20	3	id	items	support
view as				
		1	{mineral water}	0.39604
		2	{spaghetti}	0.331683
		3	{eggs}	0.257426
		4	{chocolate}	0.247525
		5	{french fries}	0.237624

9. หากหาความสัมพันธ์ (Association Rules) ในชุดข้อมูล mini market basket โดยลากโมดูล Discover Association Rules มาวางยัง workspace และเพิ่มข้อความอธิบาย (Comment) เป็น “Find Association Rules”
10. นำข้อมูลส่งออกจากโมดูลข้อมูล mini market basket เป็นข้อมูลนำเข้า Dataset ของโมดูล Discover Association Rules (สำหรับหาความสัมพันธ์)
25. คลิกที่กล่องโมดูล Discover Association Rules (สำหรับหาความสัมพันธ์) กำหนด Input Dataset Type เป็น Items list แล้วคลิก Launch column selector เพื่อเลือกตัวแปรที่จัดเก็บรายการสินค้าที่ต้องการประมวลผล ในที่นี้เลือกตัวแปร items แล้วคลิก ✓
11. กำหนดค่าพารามิเตอร์ ดังนี้
- Minimal Support: 0.1
  - Minimal Confidence: 0.1
  - Minimal Number of Items in a Rule: 1

- Maximal Number of Items in a Rule: 5
- Return type: Rules

12. คลิก **RUN** เพื่อทำการประมวลผล แล้วดูผลลัพธ์จากข้อมูลออกของโมดูล Discover Association Rules (สำหรับหาความสัมพันธ์) โดยคลิกที่โหนดส่วนต่อประสานข้อมูลออก แล้วเลือก Visualize จะปรากฏผลลัพธ์ดังรูป

Practice 5 > Discover Association Rules > Rules

rows 23 columns 6

กฎความสัมพันธ์ถูกอธิบายในรูปแบบ lhs → rhs

id	lhs	rhs	support	confidence	lift
1	{chocolate}	{mineral water}	0.123762	0.5	1.2625
2	{eggs}	{mineral water}	0.118812	0.461538	1.165385
3	{spaghetti}	{mineral water}	0.143564	0.432836	1.09291
4	{}	{mineral water}	0.39604	0.39604	1
5	{mineral water}	{spaghetti}	0.143564	0.3625	1.09291
6	{}	{spaghetti}	0.331683	0.331683	1
7	{mineral water}	{chocolate}	0.123762	0.3125	1.2625
8	{mineral water}	{eggs}	0.118812	0.3	1.165385

### 3. แบบฝึกปฏิบัติการ

ให้นักศึกษาทำแบบฝึกปฏิบัติการ ตามลำดับขั้นตอนต่อไปนี้

1. ให้นักศึกษานำชุดข้อมูล Simulated Online Shopping Carts จากแฟ้มข้อมูล simulated\_online\_shopping\_carts.csv เข้าสู่โปรแกรม ML Studio กำหนดชื่อชุดข้อมูล เป็น “simulated online shopping carts”
2. สร้างการทดลอง กำหนดชื่อเป็น “Lab 5” โดยให้นำชุดข้อมูล simulated online shopping carts เข้าสู่การทดลอง
3. หา Frequent Item Sets ในชุดข้อมูล simulated online shopping carts จากตัวแปร products โดยกำหนดค่าพารามิเตอร์ ดังนี้ (เพิ่มข้อความอธิบาย (Comment) โมดูล เป็น “Find Frequent Item Sets”)
  - Minimal Support: 0.01
  - Minimal Confidence: 0.2
  - Minimal Number of Items in a Rule: 1
  - Maximal Number of Items in a Rule: 5

4. หากกฎความสัมพันธ์ (Association Rules) ในชุดข้อมูล simulated online shopping carts จากตัวแปร products โดยกำหนดค่าพารามิเตอร์ ดังนี้ (เพิ่มข้อความอธิบาย (Comment) โมดูล เป็น “Find Association Rules”
- Minimal Support: 0.1
  - Minimal Confidence: 0.2
  - Minimal Number of Items in a Rule: 1
  - Maximal Number of Items in a Rule: 5
5. ทดลองเปลี่ยนค่าพารามิเตอร์ต่างๆ ของโมดูลการหา Frequent Item Sets และกฎความสัมพันธ์เป็นค่าอื่นๆ และศึกษาผลลัพธ์ที่เกิดขึ้น ลองหาคำตอบว่าพารามิเตอร์แต่ละตัว ส่งผลต่อผลลัพธ์ที่เกิดขึ้นอย่างไร

สิ่งที่ต้องส่งเป็นการบ้าน ภาพหน้าจอ Workspace ของนักศึกษาที่ใช้ทำแบบฝึกปฏิบัติการ โดยให้เห็นกล่องโมดูลทั้งหมดและชื่อ Workspace ซึ่งเป็นชื่อของนักศึกษา ตั้งชื่อไฟล์ในรูปแบบ Lab\_05\_id.jpg โดยแทน id ด้วยรหัสนักศึกษา ส่งผ่านเว็บไซต์ <http://hw.cs.science.cmu.ac.th>