

## การหาความสัมพันธ์ของประเภท application แต่ละประเภท

### โดยใช้เทคนิคการหาความสัมพันธ์ด้วยวิธี Apriori

จากข้อมูลแอปพลิเคชันทั้งหมดจำนวน 459,943 แอป ทำการสุ่มข้อมูลตัวอย่างมาจำนวน 2,500 ข้อมูล เพื่อตรวจสอบว่า แต่ละแอปพลิเคชันนั้นเป็นประเภทแอปพลิเคชันอะไรบ้าง ซึ่งก็จะพบว่าแอปพลิเคชันหนึ่งแอปนั้นสามารถเป็นหลากหลายประเภท จึงทำให้เราสามารถค้นหาความสัมพันธ์ของประเภทแอปพลิเคชันได้จากข้อมูลประเภทแอปพลิเคชันเหล่านั้น

ตารางแสดงตัวอย่างข้อมูลประเภทแอปพลิเคชัน

แอปพลิเคชัน	ประเภทของแอปพลิเคชัน
A	Industry tag, 1 free
B	Game, Tencent, Custom label
C	Game, 1 free

นำข้อมูลแอปพลิเคชันไปหาความสัมพันธ์ของแอปพลิเคชันแต่ละประเภทโดยวิธี Apriori โดยกำหนดค่าสนับสนุนขั้นต่ำเป็น 0.2 และคำนวณค่าสนับสนุนของแต่ละรูปแบบประเภทแอปพลิเคชัน ดังตาราง

ตารางแสดงการคำนวณค่าสนับสนุนของ application แต่ละประเภท

No.	Application	Support	Size
1	Game	0.50	1
2	Tencent	0.51	1
3	Custom label	0.59	1
4	Industry tag	0.55	1
5	Property Industry 2.0	0.43	1
6	1 free	0.23	1
7	game,Tencent	0.50	2
8	game,Custom label	0.50	2
9	Tencent,Custom label	0.51	2
10	Industry tag,Property Industry 2.0	0.43	2
11	Industry tag,1 free	0.23	2
12	Property Industry 2.0,1 free	0.23	2
13	game,Tencent,Custom label	0.50	3
14	Industry tag,Property Industry 2.0,1 free	0.23	3

จากค่าสนับสนุนของแต่ละรูปแบบประเภทแอปพลิเคชันที่คำนวณได้ดังกล่าว สามารถให้ความหมายข้อมูลได้ถึงจำนวนประเภทแอปพลิเคชันที่พบและประเภทของแอปพลิเคชันที่พบร่วมกันในข้อมูลตัวอย่างที่สุ่มมา ยกตัวอย่างเช่น ข้อมูลตารางแถวหมายเลข 1 แอปพลิเคชันประเภท Game ที่มีค่าสนับสนุน 0.50 นั้นหมายความว่า แอปพลิเคชันที่พบจะเป็นแอปพลิเคชันประเภท Game เป็น 50% ของจำนวนข้อมูลตัวอย่างแอปพลิเคชันทั้งหมด หรือ ข้อมูลตารางแถวหมายเลข 12 ประเภท Industry tag 2.0, 1 free ที่มีค่าสนับสนุน 0.23 หมายความว่า แอปพลิเคชันที่พบจะเป็นแอปพลิเคชันประเภท Industry tag 2.0 และ 1 free ร่วมกันเป็น 23% ของจำนวนข้อมูลตัวอย่างแอปพลิเคชันทั้งหมด

ดังนั้นจึงทำให้เราสามารถนำค่าสนับสนุนไปคำนวณหาความเชื่อมั่นและค่าลิฟท์ เพื่อแสดงถึงความสัมพันธ์ของแต่ละประเภทแอปพลิเคชันต่างๆ ดังแสดงในตาราง

ตารางแสดงกฎความสัมพันธ์ด้วยความเชื่อมั่นและค่าลิฟท์ของแต่ละประเภทแอปพลิเคชัน

No.	Association Rules	Confidence	Lift
1	game => Tencent	1.00	0.51
2	Tencent => game	0.98	0.49
3	game => Custom label	1.00	0.59
4	Custom label => game	0.85	0.43
5	Tencent => Custom label	1.00	0.59
6	Custom label => Tencent	0.87	0.44
7	Industry tag => Property Industry 2.0	0.79	0.34
8	Property Industry 2.0 => Industry tag	1.00	0.55
9	Industry tag => 1 free	0.41	0.09
10	1 free => Industry tag	1.00	0.55
11	Property Industry 2.0 => 1 free	0.53	0.12
12	1 free => Property Industry 2.0	1.00	0.43
13	game => Tencent,Custom label	1.00	0.51
14	Tencent => game,Custom label	0.98	0.49
15	Custom label => game,Tencent	0.85	0.43
16	game,Tencent => Custom label	1.00	0.59
17	game,Custom label => Tencent	1.00	0.51
18	Tencent,Custom label => game	0.98	0.49
19	Industry tag => Property Industry 2.0,1 free	0.41	0.09
20	Property Industry 2.0 => Industry tag,1 free	0.53	0.12
21	1 free => Industry tag,Property Industry 2.0	1.00	0.43
22	Industry tag,Property Industry 2.0 => 1 free	0.53	0.12
23	Industry tag,1 free => Property Industry 2.0	1.00	0.43
24	Property Industry 2.0,1 free => Industry tag	1.00	0.55

จากข้อมูลตารางแสดงกฎความสัมพันธ์ด้วยค่าความเชื่อมั่นและค่าลิฟท์ของแต่ละประเภทแอปพลิเคชันดังกล่าว สามารถให้ความหมายได้ถึง ค่าความสัมพันธ์ของประเภทแอปพลิเคชันหนึ่งๆมีความสัมพันธ์ต่อการเป็นอีกประเภทของ แอปพลิเคชันหนึ่งๆมาน้อยเพียงใด ยกตัวอย่างเช่น ข้อมูลในตารางแถวหมายเลข 3 ความสัมพันธ์ของประเภทแอปพลิเคชัน Game => Custom label ให้ความหมายถึงการที่แอปพลิเคชันเป็นประเภท Game แล้วจะเป็นแอปพลิเคชันประเภท Custom label ด้วย หรืออีกตัวอย่างข้อมูลในตารางแถวหมายเลข 24 ความสัมพันธ์ของประเภทแอปพลิเคชัน Property Industry 2.0, 1 free => Industry tag ให้ความหมายถึงแอปพลิเคชันที่เป็นประเภท Industry 2.0 และ 1 free แล้วจะเป็นแอปพลิเคชันประเภท Industry tag ด้วย ส่วน confidence ที่มีค่า 1.00 นั้นให้ความหมายถึง แอปพลิเคชันที่เป็นประเภท Industry 2.0 และ 1 free แล้วจะเป็นแอปพลิเคชันประเภท Industry tag ด้วยนั้นมีค่าความเชื่อมั่นของกฎความสัมพันธ์ที่จะเกิดขึ้นเป็นจำนวน  $100\%(1.00 \times 100)$  ส่วน lift ที่มีค่า 0.55 เป็นการบ่งบอกถึงรูปแบบแอปพลิเคชัน Property Industry 2.0, 1 free และ Industry tag มีความสัมพันธ์ในเชิงขึ้นต่อกันเป็นค่า 0.55