

Association rules

FP-Growth

Dataset - Chapter05Exercise.csv (market basket)

The screenshot displays the RapidMiner Studio interface with the following components:

- Repository:** Shows a tree structure with 'Community Samples (connected)', 'DB (Legacy)', 'Local Repository (MC-Y56-A) (Sasipa)', 'Connections (MC-Y56-A) (Sasipa)', 'data (MC-Y56-A) (Sasipa)', and 'processes (MC-Y56-A) (Sasipa)'. Under 'data', there are files like 'Chapter05Exercise_marketbasket', 'glass', 'glass_dataset', and 'OrderSource1'.
- Process:** A workflow diagram showing the sequence of operators: 'Retrieve Chapter05...', 'Remove Attribute R...', 'Numerical to Binomi...', 'FP-Growth', and 'Create Association...'. The 'FP-Growth' operator is highlighted.
- Parameters:** A panel for the 'FP-Growth' operator with settings: 'input format' (items in dummy c...), 'positive value' (empty), 'min requirement' (support), 'min support' (0.5), 'min items per itemset' (1), 'max items per itemset' (0), 'max number of itemsets' (1000000), and a checked 'find min number of itemsets' option.
- Help:** A panel for 'FP-Growth' with a 'Concurrency' tag, a list of tags (Associations, Market Basket, Upselling, Up-selling, Cross-selling, Itemset, Item-set, Item set, Mining, Frequent Patterns), a 'Synopsis' section, and a 'Jump to Tutorial Process' link.

แบบฝึกหัด

Dataset - Chapter05DataSet.csv

ข้อมูลนี้ได้จากการสร้างแบบสอบถามออนไลน์เพื่อเก็บข้อมูลของสมาชิกในชุมชน

- **Elapsed_Time:** เวลาที่ผู้ตอบแบบสอบถามใช้ในการตอบแบบสอบถาม โดยมีการเก็บเป็นหน่วยนาที่ เช่น 4.5 หมายถึง ผู้ตอบแบบสอบถามใช้เวลาในการตอบทั้งหมด 4 นาที 30 วินาที
- **Time_in_Community:** ระยะเวลาที่ผู้ตอบแบบสอบถามอาศัยอยู่ในชุมชนโดยแบ่งเป็น 3 ช่วงเวลาคือ Short (0-2 ปี), Medium (3-9 ปี) และ Long (10 ปีขึ้นไป)
- **Gender:** เพศของผู้ตอบแบบสอบถาม
- **Working:** ผู้ตอบแบบสอบถามมีงานทำหรือไม่ โดยการตอบเป็นแบบ yes/no
- **Age:** อายุของผู้ตอบแบบสอบถาม
- **Family:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกขององค์กร/สมาคม ประเภท ครอบครัว เช่น สมาคม Big Brothers/Big Sisters

- **Hobbies:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกองค์กร/สมาคมที่เกี่ยวข้องกับงานอดิเรก เช่น สมาคมวิทยุสมัครเล่น ชมรมเดินป่า และชมรมจักรยาน
- **Social_Club:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกองค์กร/สมาคมด้านสังคม เช่น สมาคมโรตารี และสมาคมไลอ้อน
- **Political:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกองค์กร/สมาคมทางการเมืองต่างๆ เช่น ผู้สนับสนุนพรรคการเมือง
- **Professional:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกองค์กร/สมาคมทางวิชาชีพ เช่น สมาคมทางการแพทย์ สมาคมนักกฎหมาย
- **Religious:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกองค์กร/สมาคมทางศาสนา เช่น ศรัทธาวัด และ สมาชิกโบสถ์คริสต์จักร
- **Support_Group:** คำตอบแบบ yes/no เพื่อแสดงว่าผู้ตอบแบบสอบถามเป็นสมาชิกองค์กร/สมาคมที่ให้ความช่วยเหลือผู้ประสบปัญหา เช่น ผู้ติดยาเรื้อรัง และ แม่เลี้ยงเดี่ยว (single mom)

ข้อมูลที่เลือกมาวิเคราะห์ (ใช้ select attributes operator) คือ

Family, Hobbies, Social_Club, Political, Professional, Religious, Support_Group

ทำแล้ววิเคราะห์ผลลัพธ์ที่ได้