**Business Data Mining**

**Assignment 6**

**Name: Fahad Khalid Alofani ID: 437104121**

**Problem 1 [50pts]**. A man recorded a list of transactions for his candy truck that took place at a specified time. The man wants to understand his customer purchase behavior to make specific marketing actions. Download and use the dataset in the file named “CandyTruck.csv”, and answer the following questions (do not use RapidMiner):

1. [5pts] How many transactions are there, and how many items (products) does the man sell?

There are 100 transactions, and the man sells 3 products only (Apples, Cookies, Popcorn)

1. [15pts] List the possible itemsets that can be generated from the dataset, and provide their frequencies as counts and percentages.

|  |  |  |
| --- | --- | --- |
| Itemset | Frequency | Percentage |
| Apples | 4 | 4% |
| Cookies | 8 | 8% |
| Popcorn | 28 | 28% |
| Cookies, Popcorn | 25 | 25% |
| Apples, Cookies | 25 | 25% |
| Apples, Cookies, Popcorn | 10 | 10% |

1. [20pts] Generate all possible rules and calculate their support, confidence, and lift.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Apples | Cookies | Popcorn |
| Apples | 39 | 35 | 10 |
| Cookies | 35 | 68 | 35 |
| Popcorn | 10 | 35 | 63 |

|  |  |  |  |
| --- | --- | --- | --- |
| Rule | Support | Confidence | Lift |
| If Apples, then Cookies | 35% | 89.74% | 1.3 |
| If Apples, then Popcorn | 10% | 25.64% | 0.41 |
| If Apples, then {Cookies, Popcorn} | 10% | 25.64% | 0.73 |
| If Cookies, then Apples | 35% | 51.47% | 1.3 |
| If Cookies, then Popcorn | 35% | 51.47% | 0.81 |
| If Cookies, then {Apples, Popcorn} | 10% | 14.7% | 1.47 |
| If Popcorn, then Apples | 10% | 15.87% | 0.41 |
| If Popcorn, then Cookies | 35% | 55.55% | 0.81 |
| If Popcorn, then {Apples, Cookies} | 10% | 15.87% | 0.45 |
| If {Apples, Cookies}, then Popcorn | 10% | 28.57% | 0.45 |
| If {Apples, Popcorn}, then Cookies | 10% | 100% | 1.47 |
| If {Cookies, Popcorn}, then Apples | 10% | 28.57% | 0.73 |

1. [10pts] Based on the analysis above, provide one or two marketing actions, and state why the man should follow your advice?
2. Since Apples sales are low compared to the others, and the rule “If Apples, then Cookies” has high support, confidence and lift. I suggest that you should make a bundle for apples and cookies. For example: buy 5 apples and get 1 cookie for free
3. Same as the above but if you buy 2 apples and 2 popcorn you get 1 cookie for free

**Bonus Problem [50pts]**. A supermarket recorded set of transactions for a specific week (GroceryData.csv). The supermarket manager is interested to make a specific marketing change to his store (promotion, bundling, product placement, etc.). However, he needs your help to analyze the data, understand consumer purchase behavior, and then come up with a specific recommendation, or set of recommendations. In order to do this, you may follow these steps:

* Use the RapidMiner sample process to import and examine the data (“Association Rules Starter Process.rmp”)
* How many 1-temset, 2-itemset, 3-itemset, and 4-itemset are there?

1-itemset=157 , 2-itemset=2981 , 3-itemset=6831 , and 4-itemset=0

* What is the most frequent item (top-selling item)?

Top selling item is whole milk

* What are the two most frequent items that sell together? Three? Four?

2- Whole milk, other vegetables

1. Whole milk, other vegetables, and root vegetables
2. none

* Generate association rules and specify minimum thresholds. What are some prominent rules? List few of them (5 to 10).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number | Premises | Conclusion | support | confidence | LaPlace | Gain | p-s | lift |
| 1 | yogurt, root vegetables | whole milk | 0.015 | 0.6 | 0.99 | -0.04 | 0.008 | 2.203 |
| 2 | other vegetables, domestic eggs | whole milk | 0.012 | 0.55 | 0.99 | -0.032 | 0.007 | 2.162 |
| 3 | root vegetables, tropical fruit | other vegetables | 0.012 | 0.59 | 0.99 | -0.03 | 0.008 | 3.021 |
| 4 | root vegetables, tropical fruit | whole milk | 0.012 | 0.57 | 0.99 | -0.03 | 0.007 | 2.231 |
| 5 | other vegetables, butter | whole milk | 0.011 | 0.57 | 0.99 | -0.03 | 0.006 | 2.25 |
| 6 | root vegetables, citrus fruit | other vegetables | 0.01 | 0.59 | 0.99 | -0.03 | 0.007 | 3.03 |
| 7 | yogurt, curd | whole milk | 0.01 | 0.58 | 0.99 | -0.03 | 0.006 | 2.28 |
| 8 | other vegetables, curd | whole milk | 0.01 | 0.57 | 0.99 | -0.03 | 0.005 | 2.25 |
| 9 | root vegetables, whipped/sour cream | whole milk | 0.009 | 0.55 | 0.99 | -0.03 | 0.005 | 2.17 |
| 10 | yogurt, butter | whole milk | 0.009 | 0.64 | 0.99 | -0.02 | 0.006 | 2.5 |

* Provide up to 3 marketing actions (recommendations) based on the analysis above

1. Make a yogurt and butter next to each other to try to increase their sales.
2. Make whole milk far from yogurt and butter to make the costumer see all our product.
3. Make promotion for root vegetables and tropical fruit and other vegetables