

Experiment 1-04

- Objective:- Implementation of F-P growth algo using mka explorer on any dataset of market basket analysis.
- Theory:- F-P growth method for determining frequent itemsets and F-P growth rule.
- Steps:-
 - ① Create itemsets starting from elements in each, till all the possible itemsets are considered.
 - ② Remove those itemsets having s.c & confidence less than the threshold.
 - ③ In the last step, itemsets with max possible element will be the frequent itemsets.
 - ④ from the frequent itemsets, determine the association rule by computing confidence for each possible FP growth rule and consider only those with confidence value more than the threshold.
- Result:- In this experiment have seen that total rules generated theoretically & by the mka are matching, hence the rules generated are 5.

EXPERIMENT 4

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Choose

FPGrowth - P 2 4 -1 -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1

Result list (right-click for ...)

210536 - FPGrowth

Start

Stop

Associator output

==== Run information ====

Scheme: weka.associations.FPGrowth -P 2 -T -1 -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1

Relation: supermarket

Instances: 4627

Attributes: 217

[list of attributes omitted]

==== Associator model (full training set) ====

FPGrowth found 16 rules (displaying top 10)

1. [fruit<=, frozen food<=, biscuit<=, total=high]: 788 ==> [bread and cake<=]: 723 <conf:(0.92)> lift:(1.27) lev:(0.03) conv:(3.35)

2. [fruit<=, baking needs<=, biscuit<=, total=high]: 760 ==> [bread and cake<=]: 696 <conf:(0.92)> lift:(1.27) lev:(0.03) conv:(3.28)

3. [fruit<=, baking needs<=, frozen food<=, total=high]: 770 ==> [bread and cake<=]: 705 <conf:(0.92)> lift:(1.27) lev:(0.03) conv:(3.27)

4. [fruit<=, vegetable<=, biscuit<=, total=high]: 815 ==> [bread and cake<=]: 746 <conf:(0.92)> lift:(1.27) lev:(0.03) conv:(3.26)

5. [fruit<=, party snack food<=, total=high]: 854 ==> [bread and cake<=]: 779 <conf:(0.91)> lift:(1.27) lev:(0.04) conv:(3.15)

6. [vegetable<=, frozen food<=, biscuit<=, total=high]: 797 ==> [bread and cake<=]: 725 <conf:(0.91)> lift:(1.26) lev:(0.03) conv:(3.06)

7. [vegetable<=, baking needs<=, biscuit<=, total=high]: 772 ==> [bread and cake<=]: 701 <conf:(0.91)> lift:(1.26) lev:(0.04) conv:(3.01)

8. [fruit<=, biscuit<=, total=high]: 954 ==> [bread and cake<=]: 866 <conf:(0.91)> lift:(1.26) lev:(0.04) conv:(3)

9. [fruit<=, vegetable<=, frozen food<=, total=high]: 834 ==> [bread and cake<=]: 757 <conf:(0.91)> lift:(1.26) lev:(0.03) conv:(3.01)

10. [fruit<=, frozen food<=, total=high]: 969 ==> [bread and cake<=]: 877 <conf:(0.91)> lift:(1.26) lev:(0.04) conv:(2.92)

Status

OK

Log

x0

● Viva Questions:-

Q1 Briefly discuss the strategy of FP-growth algo for determining frequent itemset in any dataset.

- FP growth is a tree based algo for mining freq. itemset without generating candidate sets.
- It constructs an FP tree by compressing the dataset based on freq. items.
 - Recursive tree traversal extracts freq. itemsets.
 - It is more efficient than apriori for large dataset.

Q2 Enlist / State any 5 real time applications each for FP growth algo w.r.t. data mining domain.

-
- ① Market Basket Analysis
 - ② fraud detection in financial transaction
 - ③ Healthcare analysis
 - ④ Social n/w analysis
 - ⑤ E-commerce upselling strategies.