Experiment 07 - Association Rule Mining using Data Mining Tool (RapidMiner)

Roll No.	19
Name	Manav Jawrani
Class	D15A
Subject	Business Intelligence Lab
LO Mapped	LO2: Organize and prepare the data needed for data mining algorithms in terms of attributes and class inputs, training, validating, and testing files. LO3: Implement the appropriate data mining methods like classification, clustering or association mining on large data sets using open source tools like WEKA

Experiment 7. 10 DATE:

Aim - To perform Association Rule mining using Rapid miner

Theory -

Introduction to Association mining -

Association mining is a data mining technique that identifies patterns or relationships between items in a dataset, often used for market basket analysis.

Apriori algorithm -

It is a classic data mining algorithm used for association rule learning. The algorithm uses a bottom-up approach by identifying frequent individual items and gradually combining them into larger itemsets until no more frequent itemsets can be found.

Example -

TID	List of items
T1	B, J, P
T2	B, P
T3	B, M, P
T4	E, B
T5	E, M

min support count = 40%

Find frequent closed, maximal itemset.

DATE:

→ Soln:

All items are frequent except J.

Item	count
B	4
E	2
J	1
M	2
P	3

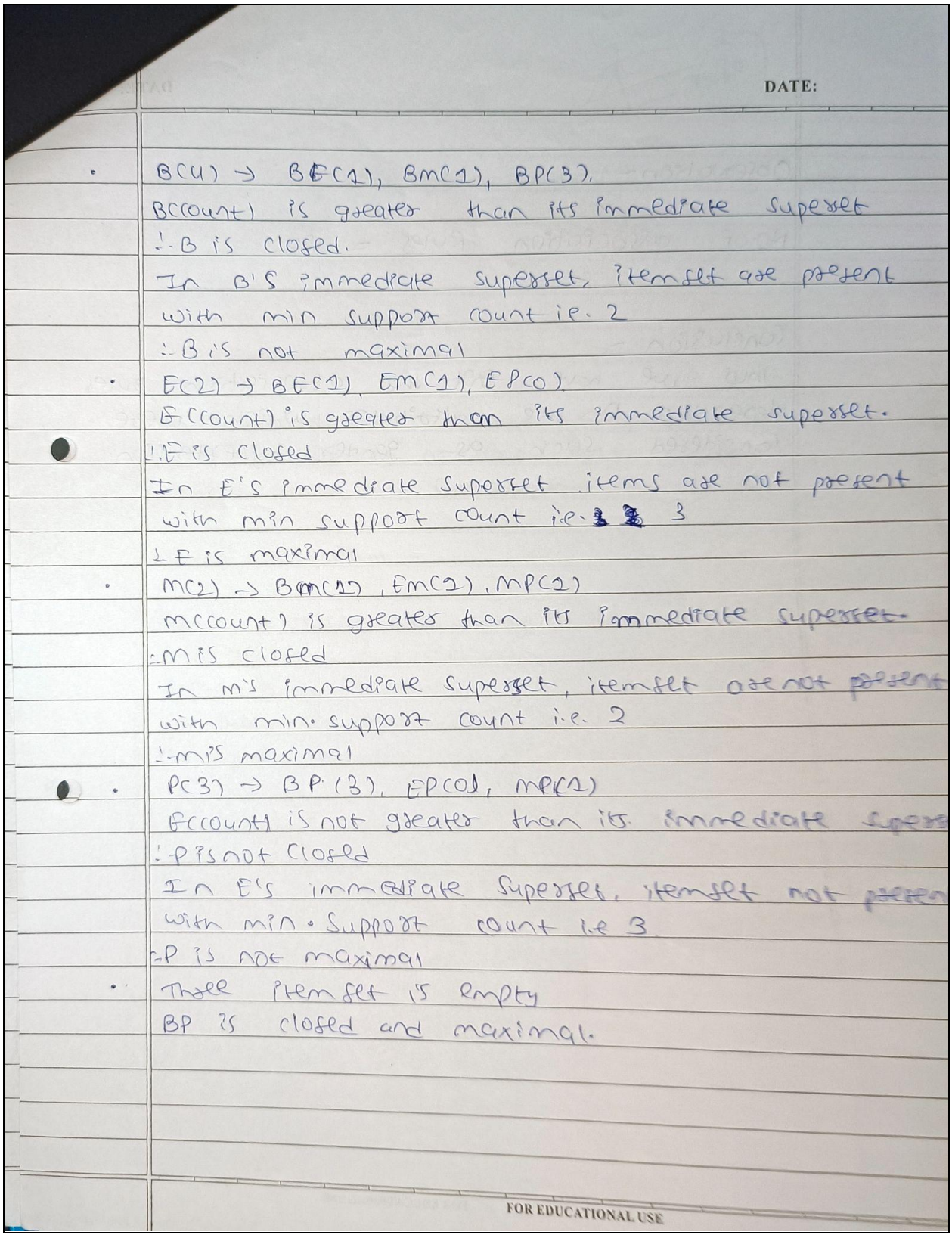
Frequent 1-itemsets

Item	count
B	4
E	2
M	2
P	3

Frequent 2- Itemsets

Item	count
B,E	1
B,M	1
B,P	3
E,M	1
E,P	--
M,P	1

All 2 items are not frequent. except BP.



DATE:

B(4) → BE(1), Bm(1), BP(3).
B(count) is greater than its immediate superset
∴ B is closed.
In B's immediate superset, itemset are present with min support count ie. 2
∴ B is not maximal

E(2) → BE(1), Em(1), EP(0).
E(count) is greater than its immediate superset.
∴ E is closed
In E's immediate superset items are not present with min support count ie. 3
∴ E is maximal

M(2) → Bm(1), Em(1), MP(1)
M(count) is greater than its immediate superset.
∴ M is closed
In M's immediate superset, itemset are not present with min. support count i.e. 2
∴ M is maximal

P(3) → BP(3), EP(0), MP(1)
P(count) is not greater than its immediate superset
∴ P is not closed
In E's immediate superset, itemset not present with min. Support count i.e 3
∴ P is not maximal

Third itemset is empty
BP is closed and maximal.

DATE:

Observation -

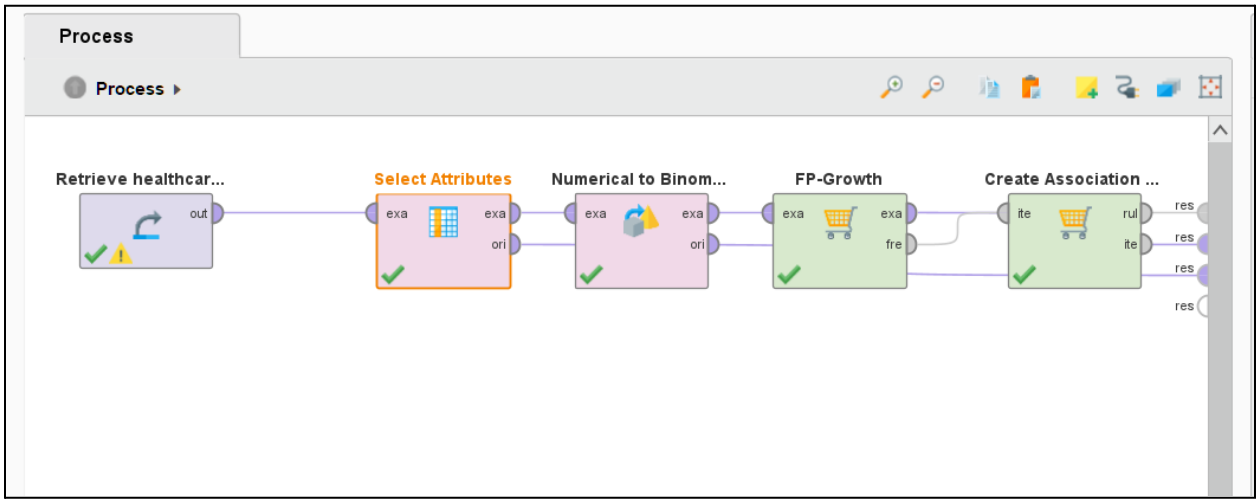
No of association Rules = 23

Conclusion -

Thus we have got the association rules depending on the attributes which were considered such as gender, hypertension etc.

Implementation -

Process -

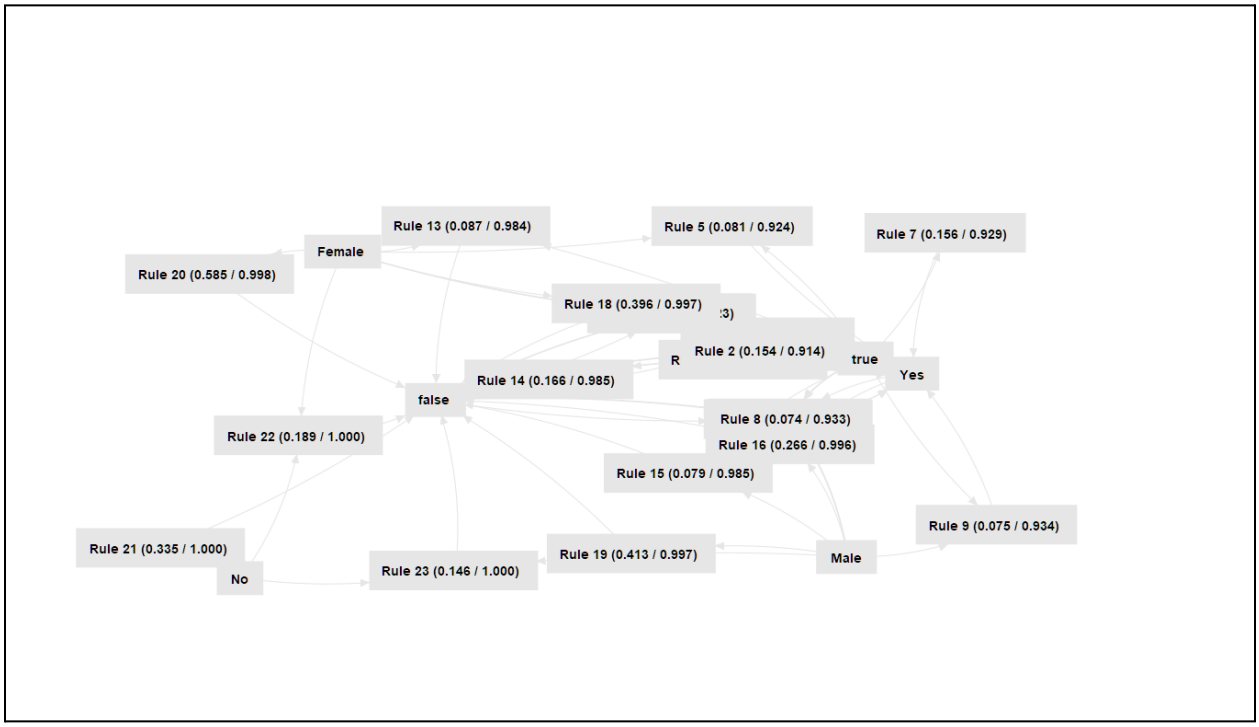


Association Rules -

AssociationRules

```
Association Rules
[Female, true] --> [false, Yes] (confidence: 0.909)
[true] --> [false, Yes] (confidence: 0.914)
[Male, true] --> [false, Yes] (confidence: 0.919)
[false, Female, true] --> [Yes] (confidence: 0.923)
[Female, true] --> [Yes] (confidence: 0.924)
[false, true] --> [Yes] (confidence: 0.928)
[true] --> [Yes] (confidence: 0.929)
[false, Male, true] --> [Yes] (confidence: 0.933)
[Male, true] --> [Yes] (confidence: 0.934)
[Yes, Female, true] --> [false] (confidence: 0.983)
[Yes, true] --> [false] (confidence: 0.984)
[Yes, Male, true] --> [false] (confidence: 0.984)
[Female, true] --> [false] (confidence: 0.984)
[true] --> [false] (confidence: 0.985)
[Male, true] --> [false] (confidence: 0.985)
[Yes, Male] --> [false] (confidence: 0.996)
[Yes] --> [false] (confidence: 0.996)
[Yes, Female] --> [false] (confidence: 0.997)
[Male] --> [false] (confidence: 0.997)
[Female] --> [false] (confidence: 0.998)
[No] --> [false] (confidence: 1.000)
[Female, No] --> [false] (confidence: 1.000)
[Male, No] --> [false] (confidence: 1.000)
```

Graph -


Rule 13 (0.087 / 0.984)
Rule 5 (0.081 / 0.924)
Rule 7 (0.156 / 0.929)
Female
Rule 20 (0.585 / 0.998)
Rule 18 (0.396 / 0.997)
:3)
R
Rule 2 (0.154 / 0.914)
true
Yes
Rule 14 (0.166 / 0.985)
false
Rule 8 (0.074 / 0.933)
Rule 16 (0.266 / 0.996)
Rule 22 (0.189 / 1.000)
Rule 15 (0.079 / 0.985)
Rule 9 (0.075 / 0.934)
Rule 21 (0.335 / 1.000)
Rule 19 (0.413 / 0.997)
Male
No
Rule 23 (0.146 / 1.000)