

DWM PARTICAL NO 2

Assignment No 2: Perform data preprocessing tasks and demonstrate performing association rule mining on data sets

market_data.csv file

Note: house_price.csv file is made in Mysql Database and exported to weka in csv format

Algorithm: In Associate "Choose" → associations → Apriori

Code:

```
CREATE DATABASE IF NOT EXISTS market_data;
```

```
USE market_data;
```

```
CREATE TABLE transactions (  
    id INT AUTO_INCREMENT PRIMARY KEY,  
    bread ENUM('yes', 'no'),  
    milk ENUM('yes', 'no'),  
    butter ENUM('yes', 'no'),  
    jam ENUM('yes', 'no'),  
    eggs ENUM('yes', 'no')  
);
```

```
INSERT INTO transactions (bread, milk, butter, jam, eggs) VALUES  
( 'yes', 'yes', 'no', 'no', 'yes'),  
( 'yes', 'no', 'yes', 'no', 'yes'),  
( 'no', 'yes', 'yes', 'no', 'no'),  
( 'yes', 'yes', 'yes', 'yes', 'yes'),  
( 'no', 'no', 'no', 'yes', 'no'),  
( 'yes', 'yes', 'yes', 'no', 'yes'),  
( 'no', 'yes', 'no', 'yes', 'yes'),  
( 'yes', 'yes', 'no', 'yes', 'no'),
```

('yes', 'no', 'no', 'yes', 'yes'),
('no', 'no', 'yes', 'yes', 'yes'),
('yes', 'yes', 'yes', 'yes', 'yes');

SELECT * FROM transactions;

Preprocessing Step:

Note: Go To Weka and covert the numeric values to nominal

Filter → unsupervised → attribute → NumericToNominal

Output:

```
Apriori
=====

Minimum support: 0.2 (2 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 16

Generated sets of large itemsets:

Size of set of large itemsets L(1): 10
Size of set of large itemsets L(2): 33
Size of set of large itemsets L(3): 30
Size of set of large itemsets L(4): 7
Size of set of large itemsets L(5): 1

Best rules found:

1. bread=yes butter=yes 4 ==> eggs=yes 4    <conf:(1)> lift:(1.38) lev:(0.1) [1] conv:(1.09)
2. jam=no eggs=yes 3 ==> bread=yes 3    <conf:(1)> lift:(1.57) lev:(0.1) [1] conv:(1.09)
3. bread=yes jam=no 3 ==> eggs=yes 3    <conf:(1)> lift:(1.38) lev:(0.07) [0] conv:(0.82)
4. butter=yes jam=yes 3 ==> eggs=yes 3    <conf:(1)> lift:(1.38) lev:(0.07) [0] conv:(0.82)
5. milk=yes butter=yes eggs=yes 3 ==> bread=yes 3    <conf:(1)> lift:(1.57) lev:(0.1) [1] conv:(1.09)
6. bread=yes milk=yes butter=yes 3 ==> eggs=yes 3    <conf:(1)> lift:(1.38) lev:(0.07) [0] conv:(0.82)
7. bread=yes milk=no 2 ==> eggs=yes 2    <conf:(1)> lift:(1.38) lev:(0.05) [0] conv:(0.55)
8. bread=no milk=no 2 ==> jam=yes 2    <conf:(1)> lift:(1.57) lev:(0.07) [0] conv:(0.73)
9. bread=no butter=no 2 ==> jam=yes 2    <conf:(1)> lift:(1.57) lev:(0.07) [0] conv:(0.73)
10. bread=no eggs=yes 2 ==> jam=yes 2    <conf:(1)> lift:(1.57) lev:(0.07) [0] conv:(0.73)
```