

Title: Data Pre-processing and Association rule mining using Python

Aim: Missing Value handling

Data Pre-processing

Association rule mining

Task: Finding the Frequent item set

Association rule mining

Exercises:

1. Import the numpy and pandas libraries
2. load the data set ls-01
3. Create the two copies of the data set
4. Check the null values
5. Missing data handling
6. Remove the missing values
7. Filling the missing values
8. Group the data
9. Label the missing values
10. Group the data according to the gender
11. Import the data set as
12. Dataset=[['milk', 'onion', 'nugmeg', 'kidney beans', 'eggs', 'yogurt'],
 ['oill', 'onion', 'nutmeg', 'kidney beans', 'eggs', 'yogurt'],
 ['milk', 'apple', 'kidney beans', 'eggs'],
 ['milk', 'unicorn', 'corn', 'kidney beans', 'yogurt']
 ['corn', 'onion', 'onion', 'kidney beans', 'ice cream', 'eggs']]
13. Import the pandas and mlxtend libraries
14. Create the dataset as a dataframe
15. Create the possible item set for the min_support=0.6 using apriori algorithm
16. Create the association rule mining
17. Create the sub set of antecedents,consequents,support,confidence,lift
18. Retrieve the itemset which itemset have confidence ≥ 1