Discovering Frequent Patterns in Big Data Using ECLATbitset Algorithm

In this tutorial, we will discuss two approaches to find frequent patterns in big data using ECLATbitset algorithm. 1. [__Basic approach:__](#basicApproach) Here, we present the steps to discover frequent patterns using a single minimum support value 2. [__Advanced approach:__](#advApproach) Here, we generalize the basic approach by presenting the steps to discover frequent patterns using multiple minimum support values.

Basic approach: Executing ECLATbitset on a single dataset at a particular minimum support value

Step 1: Import the ECLATbitset algorithm

In [1]:
```
from PAMI.frequentPattern.basic import ECLATbitset  as alg
```

Step 2: Specify the following input parameters

In [2]:
```
inputFile = 'transactional_T10I4D100K.csv'

minimumSupportCount=100  #Users can also specify this constraint between 0 to 1.

seperator='¥t'
```

Step 3: Execute the ECLATbitset algorithm

In [3]:
```
obj = alg.ECLATbitset(iFile=inputFile, minSup=minimumSupportCount, sep=seperator)
obj.startMine()             #Start the mining process
```
```
Frequent patterns were generated successfully using Eclat_bitset algorithm
```

Step 4: Storing the generated patterns

Step 4.1: Storing the generated patterns in a file

In [4]:
```
obj.savePatterns(outFile='frequentPatternsMinSupCount1000.txt')
```

Step 4.2. Storing the generated patterns in a data frame

In [5]:
```
frequentPatternsDF= obj.getPatternsAsDataFrame()
```

Step 5: Getting the statistics

Step 5.1: Total number of discovered patterns

In [6]:
```
print('Total No of patterns: ' + str(len(frequentPatternsDF)))
```
```
Total No of patterns: 27532
```

Step 5.2: Runtime consumed by the mining algorithm

In [7]:
```
print('Runtime: ' + str(obj.getRuntime()))
```

```
Runtime: 1398.162481546402
```

In [8]:
```
##### Step 5.3: Total Memory consumed by the mining algorithm
```

In [9]:
```
print('Memory (RSS): ' + str(obj.getMemoryRSS()))
print('Memory (USS): ' + str(obj.getMemoryUSS()))
```

```
Memory (RSS): 229834752
Memory (USS): 191168512
```