Data Mining Assignment 2

Name: Rigved Satish Patki
E-mail: patki@kth.se
Group: Assignment-group 8

Topic: Discovery of Frequent Itemsets and Association Rules

Solution:
The code for the solution of this assignment is written in typescript which then transpiiles down to ES6 javascript before execution.I am using the same dataset T10I4D100K.dat as the one that was mentioned in this assignment. The dataset contains 100000 transactions containing items that are indicated by numbers.

The flow of the code is as follows:

- The code starts in index.ts in the src file as mentioned in the package.json file. First the dataset file T10I4D100K.dat is consumed and converted into a string.
- Next the getSalesTransactions function converts the string into array of transactions containing the transactionId and array of items that are present in that transaction. The interface for transaction looks like :

```
export interface Transaction {
 transactionId: number;
 items: number[];
}
```

- Then we have a function for each step, function getSingleItemCount goes through all the transactions and returns the count for each item in all the transactions. Function getItemCountWithSupport is used to delete all the item counts that are less than the threshold for support. This threshold is present in config.ts file along with the threshold for confidence.
- getItemCountWithSupport is applied to all the item counts namely, single, double and triple.
- Further we get double item count and triple item count from getDoubleItemCount and getTripleItemCount functions respectively. And then getItemCountWithSupport removes all the item counts that are less than the threshold of support.
- Next we find confidence for double count and triple count of items using getDoubleConfidence and getTripleConfidence.
- The results for all the steps are written to files in the output folder.
- Sample of the results in the output folder can be found here.
- Makefile is present for this assignment which is similar to the one present in the last assignment .