**Data Mining**

Pranaya Bhat

1001408965

Apriori

In my code implementation, I have used frozen set to store the data. The *reason* to use a frozen set rather than a regular set is to reap the benefits of immutability. The function ‘getTransList’ takes the data from file and returns list of items and a list of transactions

and generate two separate sets of items and transactions. The function ‘dataFromFile’ reads from the file and yields a generator of frozen sets. It reads through the blank spaces in between using ‘strip’ and ‘split’. ‘returnMinSupport’ calculates the support for items in the itemSet and returns a subset of the itemSet each of whose elements satisfies the minimum support. The function ‘Apriori’ returns both items and the rules for pretuple, posttuple, confidence.

Association rule generation is usually split up into two separate steps: 1. First, minimum support is applied to find all frequent itemsets in a database. 2. Second, these frequent itemsets and the minimum confidence constraint are used to form rules.