Data Mining (CSE 4052)

Frequent Item-Set Mining for Supermarket

Frequent item-set mining is a widely used data-mining technique used for discovering sets of frequently occurring items in large databases. A typical example of where this algorithm is used is in large supermarket stores where the sellers intend to observe the frequently bought together products in the store. A supermarket owner needs to take a decision regarding renovation and modification of the products to be ordered and maintained the the store. A miniature form of the database transactions made in the store are given in the following table. We need to write PYTHON code to mine the following transactional data for finding the frequently bought together items, such that the decision making process would be easy. Each alphabet in the table represents an item starting with the alphabet, such as: M- Milk, O- Oat, N- Nescafe, K- Kellogs, E- Egg, Y- Yoghurt. The database has five transactions. Let minimum support be 60% and minimum confidence be 80%.

TID	Items bought
T100	M, O, N, K, E, Y
T200	D, O, N, K, E, Y
T300	M, A, K, E
T400	M, U, C, K, Y
T500	C, O, O, K, I, E

Find all frequent item sets which are to be maintained in sufficient amount in the supermarket using Apriori and FP-growth, respectively. List all the strongly associated items to be kept close by in the store.