

Northeastern University, Khoury College of Computer Science

CS 6220 Data Mining | Assignment 2

Due: February 25, 2023(100 points)

Yichen Sun

https://github.com/LAnselet/cs6220-datamining

Frequent Itemsets

Question 1

Fk – 1 x F1 = F3 x F1 = {1, 2, 3, 4}, {1, 2, 3, 5}, {1, 2, 4, 5}, {1, 3, 4, 5}, {2, 3, 4, 5}

Question 2

Fk – 1 x Fk - 1 = F3 x F3 = {1, 2, 3, 4}, {1, 2, 3, 5}, {1, 2, 4, 5}, {2, 3, 4, 5}

Question 3

C4 after prune = {1, 2, 3, 4}

Association Rules

Question 4

1. R = 3n – 2n + 1 + 1 = 37 – 28+ 1 = 1932

Question 5

True. As Apriori principle illustrated. If {a, b, c} if a frequent itemset, then all subsets of this item set are frequent.

Question 6

False. {a, b, c} may not even exist.

Question 7

False. The support of {b} could be 50 without overlapping.

Question 8

False. The maximum can potentially generate up to 2k – 1 frequent itemsets which will be 25 – 1 itemsets.

Question 9

A drawing of a hexagon

Description automatically generated