

Data Science 2019
HW4 - Frequent Itemset Mining
Due: 2019/11/20 09:00 a.m.

1. Please use the **Apriori algorithm** to mine frequent itemsets and then generate association rules for the database D.

minimum support count: 3 transactions (50%)

minimum confidence: 66%

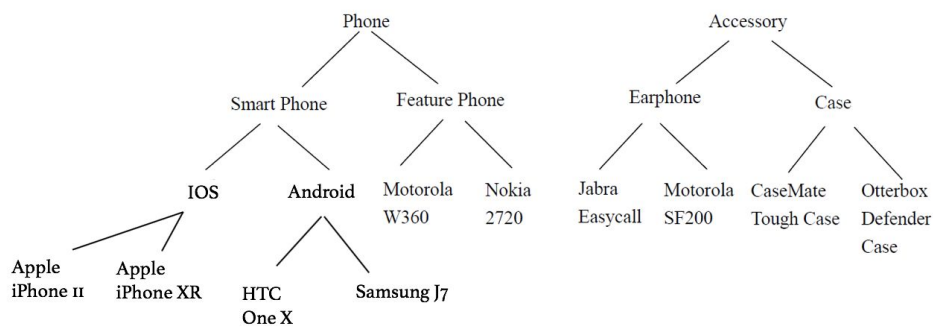
- a. Do the large itemset generation [20%]
- b. Derive any 5 strong association rules [15%]
- c. List all closed itemsets [10%]
- d. List all maximal itemsets [10%]
- e. Show one strong association rule whose lift is less than 1 [5%]

D

TID	Items
100	I1 I3
200	I3 I4 I5
300	I1 I3 I4 I5
400	I4 I5
500	I1 I2 I5
600	I1 I2 I3 I4 I5

2. Given the database below,

- a. What is the confidence and support of rule: “Phone → Earphone”? [20%]
- b. What is the confidence and support of rule: “Smart Phone → Case”? [20%]



Tx	Items bought
100	Apple iPhone 11, Jabra Easycall, CaseMate Tough Case
200	Apple iPhone XR, CaseMate Tough Case
300	Nokia2720, Jabra Easycall
400	HTC One X, Jabra Easycall
500	Motorola W360
600	Samsung J7, Otterbox Defender Case
700	HTC One X, Otterbox Defender Case
800	MotorolaSF200
900	Apple iPhone 11, CaseMate Tough Case
1000	Motorola W360, MotorolaSF200

Submission

- Write down the process you get the answer as clear as possible.
- Please submit your homework in **PDF** format.
- You may submit handwriting by scanning or capturing, be sure that it's clear enough.
- Wrong submission format will get 10 points penalty.
- No plagiarism.
- Submission deadline: **2019/11/20 09:00 a.m.**
- Accept late submission for 2 days after deadline.
- Late submission penalty is 20% per day (start from 9:00a.m.)