

Online Assignment 2

CAP446: Data Warehousing and Data Mining

Max Marks: 30

Instructions:

1. Attempt all the questions
 2. Solutions must be handwritten and individual.
 3. Copied cases/ Wrong uploads will be marked zero.
 4. Upload assignments on UMS navigation → Learning management system → Assignment upload → choose CAP446 from dropdown
1. For the following given transaction dataset, generate rules using Apriori algorithm. Consider values as SUPPORT=30% and CONFIDENCE=60% [15]

TID	PRODUCTS			
1	Milk	Egg	Bread	Butter
2	Milk	Butter	Egg	Ketchup
3	Bread	Butter	Ketchup	
4	Milk	Bread	Butter	
5	Bread	Butter	Cookies	
6	Milk	Bread	Butter	Cookies
7	Milk	Cookies		
8	Milk	Bread	Butter	
9	Bread	Butter	Egg	Cookies
10	Milk	Butter	Bread	
11	Milk	Bread	Butter	
12	Milk	Bread	Cookies	Ketchup

2. Explain the various steps involved in Data Pre-processing by taking help of a dataset gone through the process in Rapid Miner. Attach relevant screenshots. [10]
3. Prepare a FP growth tree for the following transactions. Given Minimum support=7 [5]

Transaction No	Products
1	beer, wine, cheese
2	beer, potato chips
3	eggs, flour, butter, cheese
4	eggs, flour, butter, beer, potato chips
5	wine, cheese
6	potato chips
7	eggs, flour, butter, wine, cheese
8	eggs, flour, butter, beer, potato chips
9	wine, beer
10	beer, potato chips
11	butter, eggs
12	beer, potato chips
13	flour, eggs
14	beer, potato chips
15	eggs, flour, butter, wine, cheese
16	beer, wine, potato chips, cheese
17	wine, cheese
18	beer, potato chips
19	wine, cheese
20	beer, potato chips