Parshvanath Charitable Trust's
A. P. SHAH INSTITUTE OF TECHNOLOGY
(Approved by AICTE New Delhi & Govt. of Maharashtra, Affiliated to University of Mumbai)
(Religious Jain Minority)

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING (DATA SCIENCE)

RETEST UNIT TEST-I

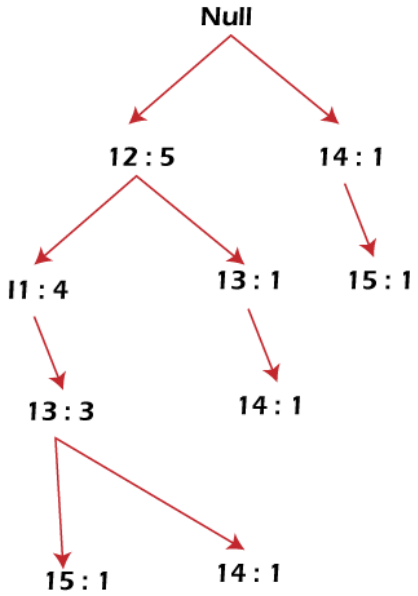
Class:TEDS **Semester:V** **Subject:DWM**

Date: 06-Nov-2023 **Time:02:30pm to 04:00pm** **Max marks: 40**

Note the following instructions

1. **Attempt all questions.**
2. **Draw neat diagrams wherever necessary.**
3. **Write everything in Black ink (no pencil) only.**
4. **Assume data, if missing, with justification.**

Q.N	Questions	MARKS	CO	Bloom Level	PO
Q.1.	**Attempt any two.**				
A.	What is web structure mining? List the approaches used to structure the web pages to improve effectiveness of search engine and crawlers.	[5]	CO6	L2	
B.	Explain web usage mining in detail.	[5]	CO6	L2	
C.	What is web content mining, and how does it differ from traditional data mining techniques?	[5]	CO6	L2	
D.	How is PageRank used by search engines like Google to rank web pages?	[5]	CO6	L2	
Q.2.	**Attempt any two**				
A.	Find the frequent item set in the following database of given transactions with min support 50% and confidence 70% TID / ITEMS 100 / 1 3 4 200 / 2 3 5 300 / 1 2 3 5 400 / 2 5	[10]	CO5	L3	PO2
B.	For the following transactional database construct FP tree.	[10]	CO3	L3	PO2

		TID | Items 1 | B D C A 2 | E D C 3 | A B 4 | A C D 5 | F G D B				
C.	For the following FP tree mine frequent patterns using the FP-growth algorithm.  	[10]	CO3	L3	PO2
Q.3.	**Attempt any one.**					
A.	Find clusters using k-means clustering algorithm if we have several objects (4 types of medicines) and each object have two attributes or features as shown in the table below. The goal is to group these objects into k=2 group of medicine based on the two features (pH and weight index). **Object** | **Attribute 1 (X) Weight Index** | **Attribute 2 (Y) pH** **Medicine A** | **1** | **1** **Medicine B** | **2** | **1** **Medicine C** | **4** | **3** **Medicine D** | **5** | **4**	[10]	CO4	L3	PO1
B.	For given distance matrix draw single and complete linkage dendogram | | 1 | 2 | 3 | 4 | 5 | 1 | 0 2 | 9 | 0 3 | 3 | 7 | 0 4 | 6 | 5 | 9 | 0 5 | 11 | 10 | 2 | 8 | 0	[10]	CO4	L3	PO1