FACULTY OF ENGINEERING

B.E. 4/4 (CSE) II-Semester (Main) Examination, April / May 2013

Subject : Data Mining

Time : 3 Hours

Max. Marks: 75

(10)

Note: Answer all questions of Part - A and answer any five questions from Part-B.

PART – A (25 Marks)

	What is Data Mining? Give techniques for Data Transformation	n.	(2) (3)	
3.	3. Define support and confidence. (2			
4.	Define Linear Regression		(2)	
5.	What is bit map indexing and Join index	king methods to speed up query processing?	(3)	
6.	6. Distinguish Data mart and Data Warehouse. (2)			
1.	Differentiate snow flake and Fact col	instellations schemas for multi dimensional	(2)	
0	Udiabases.	in a Data Warahayaa?	(3)	
о. О	9 What is binning? Smooth the following data using by bin means and by bin			
5.	boundaries 4 8 15 21 24 25 28 34	ing data daing by bin means and by bin	(3)	
10	What is Web Mining?		(3)	
	PART – B (5	5x10=50 Marks)	(0)	
11.	11.(a) Explain various Data Mining functionalities with examples. (5)			
	(b) Discuss various issues in Data Mining.			
			. ,	
12.	12.(a) Draw three-tier data Warehouse Architecture and explain various levels in the			
	architecture.		(6)	
	(b) Differentiate OLAP and OLTP.		(4)	
40	12 (a) Evalain variaus Data Mining Drimitivas			
13.	(7) (b) Define (i) Lease equaling (ii) Semi tight equaling (iii) tight equaling arehitecture			
	of data mining		(3)	
	or data mining		(3)	
14	Muite and explain Appierial eleventities to fi			
	. Write and explain Apriori aldorithm to th	ind all frequent item sets and strong		
	association rules for the following data	ind all frequent item sets and strong base, where min_sup=60% and		
	association rules for the following data min conf=80%.	ind all frequent item sets and strong base, where min_sup=60% and (10)	
	association rules for the following datal min_conf=80%.	ind all frequent item sets and strong base, where min_sup=60% and (Items	10)	
	association rules for the following datal min_conf=80%.	ind all frequent item sets and strong base, where min_sup=60% and (<u>Items</u> {K, A, D, B}	10)	
	 write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. Tid T100 T200 { 	ind all frequent item sets and strong base, where min_sup=60% and (7 Items {K, A, D, B} D, A, C, E, B}	10)	
	. Write and explain Aprior algorithm to it association rules for the following datal min_conf=80%.	ind all frequent item sets and strong base, where min_sup=60% and (7 <u>Items</u> {K, A, D, B} (D, A, C, E, B} {C, A, B, E}	10)	
	Tid T100 T200 { T300 T400	ind all frequent item sets and strong base, where min_sup=60% and $(\begin{tabular}{lllllllllllllllllllllllllllllllllll$	10)	
	Tid T100 T200 { T300 T400	ind all frequent item sets and strong base, where min_sup=60% and (* <u>Items</u> {K, A, D, B} D, A, C, E, B} {C, A, B, E} {B, A, D}	10)	
15.	 Write and explain Aprior algorithm to it association rules for the following datal min_conf=80%. Tid T100 T200 { T300 T400 (a) Explain in detail Bayesian Classifier. 	and all frequent item sets and strong base, where min_sup=60% and $(\frac{Items}{\{K, A, D, B\}})$ (A, C, E, B) (C, A, B, E) $\{B, A, D\}$	10) (7)	
15.	 Write and explain Aprior algorithm to it association rules for the following datal min_conf=80%. Tid T100 T200 { T300 T400 (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be seen to be see	base, where min_sup=60% and $\overline{\text{Items}}$ $\overline{\{K, A, D, B\}}$ $\overline{\{D, A, C, E, B\}}$ $\overline{\{C, A, B, E\}}$ $\overline{\{B, A, D\}}$ be estimated.	10) (7) (3)	
15.	 Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. Tid T100 T200 { T300 T400 (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be for the following datal min_conf=80%. 	base, where min_sup=60% and \overline{Items} $\overline{\{K, A, D, B\}}$ $\overline{\{C, A, B, E\}}$ $\overline{\{B, A, D\}}$ be estimated.	10) (7) (3)	
15. 16.	 Write and explain Aprior algorithm to it association rules for the following datal min_conf=80%. <u>Tid</u> <u>T100</u> <u>T200</u> { <u>T300</u> <u>T400</u> (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be shown for the following datales. 	base, where min_sup=60% and $\frac{\text{Items}}{\{K, A, D, B\}}$ $D, A, C, E, B\}$ $\{C, A, B, E\}$ $\{B, A, D\}$ be estimated. owing example to find the cluster where k =3	10) (7) (3)	
15. 16.	 Write and explain Aprior algorithm to it association rules for the following datal min_conf=80%. <u>Tid</u> <u>T100</u> <u>T200</u> { <u>T300</u> <u>T400</u> (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be and initial points are assigned to cluster and initial points are assigned to cluster and initial points are assigned to cluster. 	ind all frequent item sets and strong base, where min_sup=60% and $(\frac{Items}{\{K, A, D, B\}})$ D, A, C, E, B} $\{C, A, B, E\}$ $\{B, A, D\}$ be estimated. be estimated. owing example to find the cluster where k =3 ers as C ₁ = {A ₁ , A ₂ , A ₃ }, C ₂ ={A ₄ , A ₅ , A ₆ },	10) (7) (3)	
15.	 Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. <u>Tid</u> <u>T100</u> <u>T200</u> { <u>T300</u> <u>T400</u> (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be classifier accuracy can be classifier and initial points are assigned to cluster C₃={A₇, A₈} 	the all frequent item sets and strong (7) (7) (7) (7) (7) (7) (7) (7)	10) (7) (3) 10)	
15.	 Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. <u>Tid</u> <u>T100</u> <u>T200</u> { <u>T300</u> <u>T400</u> (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be and initial points are assigned to cluster C₃={A₇, A₈} 	the all frequent item sets and strong base, where min_sup=60% and (7) \overline{Items} $\overline{\{K, A, D, B\}}$ $\overline{\{D, A, C, E, B\}}$ $\overline{\{C, A, B, E\}}$ $\overline{\{B, A, D\}}$ be estimated. owing example to find the cluster where k =3 ers as C ₁ = {A ₁ , A ₂ , A ₃ }, C ₂ ={A ₄ , A ₅ , A ₆ }, $\overline{X Y}$ $\overline{2 10}$	10) (7) (3) 10)	
15. 16.	 Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. Tid T100 T200 { T300 T400 (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be considered and initial points are assigned to cluster C₃={A₇, A₈} 	the all frequent item sets and strong base, where min_sup=60% and (7) Items $\{K, A, D, B\}$ $D, A, C, E, B\}$ $\{C, A, B, E\}$ $\{B, A, D\}$ be estimated. owing example to find the cluster where k =3 ers as C ₁ = {A ₁ , A ₂ , A ₃ }, C ₂ ={A ₄ , A ₅ , A ₆ }, X Y 2 10 2 5	10) (7) (3) 10)	
15.	 Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. Tid T100 T200 { T300 T400 (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be and initial points are assigned to cluster C₃={A₇, A₈} 	the all frequent item sets and strong base, where min_sup=60% and (A = A + A + A + A + A + A + A + A + A +	10) (7) (3) 10)	
15.	5. Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. Iid I100 $I200$ { I300 I400 5. (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be 5. Explain K-means algorithm for the follow and initial points are assigned to cluster $C_3=\{A_7, A_8\}$ A_1 A_2 A_3 A_3	ind all frequent item sets and strong base, where min_sup=60% and $\begin{array}{r} \hline Items \\ \hline \{K, A, D, B\} \\ \hline D, A, C, E, B\} \\ \hline \{C, A, B, E\} \\ \hline \{B, A, D\} \end{array}$ be estimated. be estimated. owing example to find the cluster where k =3 ers as C ₁ = {A ₁ , A ₂ , A ₃ }, C ₂ ={A ₄ , A ₅ , A ₆ }, $\begin{array}{r} \hline X & Y \\ \hline 2 & 10 \\ \hline 2 & 5 \\ \hline 8 & 4 \\ \hline 5 & 8 \end{array}$ (7)	10) (7) (3) 10)	
15.	2. Write and explain Apriori algorithm to it association rules for the following datal min_conf=80%. $\frac{\text{Tid}}{1100}$ $\frac{\text{Tid}}{1200}$ $\frac{1}{300}$ $\frac{1}{7400}$ 3. (a) Explain in detail Bayesian Classifier. (b) Explain how classifier accuracy can be 5. Explain K-means algorithm for the follow and initial points are assigned to cluster $C_3=\{A_7, A_8\}$ $\frac{A_1}{A_2}$ A_3 A_4	ind all frequent item sets and strong base, where min_sup=60% and (7) $Items$ $\{K, A, D, B\}$ $D, A, C, E, B\}$ $\{C, A, B, E\}$ $\{B, A, D\}$ be estimated. owing example to find the cluster where k =3 ers as C ₁ = {A ₁ , A ₂ , A ₃ }, C ₂ ={A ₄ , A ₅ , A ₆ }, $\boxed{X Y}$ $2 10$ $2 5$ $8 4$ $5 8$ $7 5$	10) (7) (3) 10)	

17. Write short notes on any two of the following:

- (a) Sequential pattern mining
- (b) Mining multimedia databases
- (c) Mining text databases

4

1 2

9

 A_7

 A_8