

Total No. of Printed Pages:2

SUBJECT CODE NO: E-202
FACULTY OF ENGINEERING AND TECHNOLOGY
B.E.(CSE/ IT) Examination Nov/Dec 2017
Data Warehousing & Data Mining (CSE-IT)
(REVISED)

[Time: Three Hours]

[Max.Marks:80]

Please check whether you have got the right question paper.

N.B

- i) Q.1 and Q.6 are compulsory.
- ii) Solve any two from question 2,3,4,5 and any two from question 7,8,9,10.

Section A

- | | | |
|-----|---|----|
| Q.1 | a) What is data mining & Explain its application in web search engine. | 04 |
| | b) Explain snowflake schema using diagram. | 03 |
| | c) What is histogram? Explain its use in data preprocessing. | 03 |
| Q.2 | a) Draw and explain multi-tier architecture for data warehousing. | 08 |
| | b) Explain various proximity measure for binary and ordinal attributes. | 07 |
| Q.3 | a) What is data discretization and data normalization? Explain various techniques for data discretization. | 08 |
| | b) Define range, Quartiles, Inter quartile range, boxplot & outlier give one example for each. | 07 |
| Q.4 | a) Consider the following data values 10000, 30000, 40000, 60000, 80000 and apply min-max normalization, z-score normalization and decimal scaling normalization. | 08 |
| | b) Explain various steps in data preprocessing. | 07 |
| Q.5 | a) Calculate correlation coefficient and covariance of numeric data given in following table. | 08 |

Time Point	All Electronics	Hitech
+1	6	20
+2	5	10
+3	4	14
+4	3	5
+5	2	5

- | | |
|--|----|
| b) Explain various trends in data mining | 07 |
|--|----|

2017

Section-B

- Q.6 a) Explain the technique of classification with suitable example. 04
 b) Distinguish between supervised and unsupervised classification. 03
 c) Define support, confidence and minimum support count. 03

- Q.7 a) What is Decision tree? Explain various attribute selection measure for decision tree. 08
 b) Explain business –pressure – response support model with diagram. 07

- Q.8 a) Consider the following dataset with 7 Transaction. Find frequent itemset using apriori algorithm and also generate association rules (MSC=2, min-sup = 70%) 10

transaction	item bought
To1	A, D, B
To2	A, C, D
To3	D, C,
To4	A, D
To5	C, D, B
To6	A, C, D

- b) Explain linear & non linear regression with example. 05

- Q.9 a) Apply the decision tree (ID3) classification algorithm on following dataset & find decision tree. 10

Name	Experience	Qualification	Class allotted
P	06	PhD	PG
Q	10	ME	PG
R	8	ME	UG
S	3	PhD	PG
T	6	ME	UG
U	7	ME	UG
V	15	ME	PG

- b) Explain the architecture of business intelligent with diagram. 05

- Q.10 a) Explain K-mean clustering with suitable example. 08
 b) Write a note on “Intelligent creation and BI Governance”. 07