FUTURE VISION BIE

One Stop for All Study Materials

& Lab Programs

Future Vision

By K B Hemanth Raj

Scan the QR Code to Visit the Web Page

Or

Visit : https://hemanthrajhemu.github.io

Gain Access to All Study Materials according to VTU,
CSE – Computer Science Engineering,
ISE – Information Science Engineering,
ECE - Electronics and Communication Engineering
& MORE...

Join Telegram to get Instant Updates: https://bit.ly/VTU_TELEGRAM

Contact: MAIL: futurevisionbie@gmail.com

INSTAGRAM: www.instagram.com/hemanthraj_hemu/

INSTAGRAM: www.instagram.com/futurevisionbie/

WHATSAPP SHARE: https://bit.ly/FVBIESHARE

USN | | | | | | | | | | |

10IS74

Seventh Semester B.E. Degree Examination, Dec.2017/Jan.2018
Data Warehousing and Data Mining

Time: 3 hrs. Max. Marks:100

Note: ***Answer any FIVE full questions, selecting at least TWO questions from each part.***

PART – A

1 a. Give the definition of data warehousing. Discuss the need for data warehousing. **(06 Marks)**

b. Give the difference between OLTP and data warehouse systems. **(04 Marks)**

c. Discuss the characteristics of operational data store with its design and implementation issues. **(10 Marks)**

2 a. Describe the operations of data cube. **(10 Marks)**

b. Present five major characteristics from Codd's rule. **(05 Marks)**

c. Explain the difference between MOLAP and ROLAP. **(05 Marks)**

3 a. Explain various tasks of data mining with example for each. **(10 Marks)**

b. Explain: (i) Data mining applications, (ii) Issues in proximity calculation. **(10 Marks)**

4 a. What is Frequent Itemset Generation? Explain Frequent Itemset Generation using Apriori principle. **(10 Marks)**

b. Given the following set of transactions in market basket model. Build a frequency pattern (FP tree) show each transaction separately.

Transaction ID	Items bought
01	Milk, bread, cookies, juice
02	Milk, juice
03	Milk, eggs
04	Bread, cookies
05	Juice, eggs
06	Bread, eggs

(10 Marks)

PART – B

5 a. Explain Hunts algorithm. Using Hunts algorithm write decision tree for the following data:

Tid	Home owner	Annual Income	Marital Status	Default borrower
1	Yeas	125 K	Single	No
2	No	100 K	Married	No
3	No	70 K	Single	No
4	Yes	120 K	Married	No
5	No	95 K	Divorced	Yes
6	No	60 K	Married	No
7	Yes	220 K	Divorced	No
8	No	85 K	Single	Yes
9	No	75 K	Married	No
10	No	90 K	Single	Yes

(10 Marks)

b. Explain the various measures for selecting the best splits. **(05 Marks)**

c. Explain the rule evaluation criteria for classification. **(05 Marks)**

6 a. What are Bayesian classifiers? Explain Baye's theorem for classification. **(10 Marks)**
b. Explain how the predictive accuracy of classification methods be estimated. **(10 Marks)**

7 a. Give the definition of cluster analysis. Explain desired features of cluster analysis. **(10 Marks)**
b. Explain the following clustering technique with algorithm.
i) K-means method
ii) Divisive hierarchical method. **(10 Marks)**

8 a. What is Web data mining? Explain Web document clustering. **(06 Marks)**
b. Explain different text mining approach. **(08 Marks)**
c. Describe sequential mining technique, with an example. **(06 Marks)**

* * * * *