(Pages : 3) G – 1657

Reg. No. : .....

Name : .....

Sixth Semester B.Sc./B.C.A. Degree Examination, April 2019
Career Related FDP Under CBCSS
Group 2(B) : Computer Science/Computer Applications
Elective Course CS 1661.3/Core Course CP 1643
DATA MINING AND DATA WAREHOUSING
(2014 Admission Onwards)

Time : 3 Hours Max. Marks : 80

SECTION – A

(Very Short Answer Type)

(One word to maximum of **one** sentence, answer **all** questions) : **(10×1=10 Marks)**

1. Define a distance measure in clustering.
2. What is text mining ?
3. Define database system.
4. Define a cluster.
5. What do you mean by data integration ?
6. Define an attribute.
7. Define data reduction.
8. Define a data cube.
9. Define a classifier.
10. Define outliers.

P.T.O.

SECTION – B

(Short Answer)

(**Not** to exceed **one** paragraph, answer **any eight** questions. **Each** question carries **two** marks). **(8×2=16 Marks)**

11. Define an Online Transaction Processing System.

12. Which are the steps involved in data mining ?

13. Explain the concept of data cleaning.

14. Which are the two steps involved in classification ?

15. Discuss about meta data.

16. Explain the use of star schema.

17. Explain the features of a data warehouse.

18. What do you mean by decision tree induction ?

19. What do you mean by density based clustering ?

20. Discuss about "Lazy Leaners".

21. Define support in association rule mining.

22. Explain different data representation techniques.

SECTION – C

(Short Essay)

(**Not** to exceed **120** words, answer **any six** questions. **Each** question carries **four** marks). **(6×4=24 Marks)**

23. Write short notes on frequent pattern mining.

24. Explain in detail the steps involved in data pre-processing.

25. Explain the concept of a data cubes.

26. Explain the use of web mining.

27. What do you mean by Boolean Association ?

28. Discuss about OLAP.

29. What do you mean by a data mart ?

30. Explain Bayesian classification.

31. Explain k-nearest neighbor method.

SECTION – D

(Long Essay)

(Answer **any two** questions. **Each** question carries **15** marks). **(2×15=30 Marks)**

32. Discuss about decision trees and rule generation using decision trees.

33. Explain about Apriori algorithm in detail.

34. Explain in detail about scope and applications of data mining.

35. Explain any two clustering algorithms in detail.

———————

C – 1615

(Pages : 2)

Reg. No. : 33214B07012

Name : Syamkumar S

Sixth Semester B.Sc./B.C.A. Degree Examination, April 2017

Career Related FDP under CBCSS

Group 2(B) : COMPUTER SCIENCE / COMPUTER APPLICATIONS

Elective Course CS 1661.3/Core Course CP 1643

Data Mining and Data Warehousing

(2014 Admission)

Time : 3 Hours

Max. Marks : 80

SECTION – A

(Very Short Answer Type)

One word to maximum of **one** sentence. Answer **all** questions : **(10×1=10 Marks)**

1. Define information.
2. What is web mining ?
3. Define a decision tree.
4. Name a popular Boolean association rule mining algorithm. Apriori
5. Define data cleaning.
6. Define a continuous attribute.
7. What do you mean by smoothing ?
8. Define multi-dimensional data model.
9. Explain data reduction.
10. Name a hierarchical clustering method.

SECTION – B

(Short Answer)

Not to exceed **one** paragraph, answer **any eight** questions. **Each** question carries **two** marks : **(8×2=16 Marks)**

11. Define a data warehouse.
12. Define knowledge discovery in data mining.
13. Write short notes on types of data in cluster analysis.

P.T.O.

14. Differentiate between classifier and predictor in classification.
15. Explain slice and dice operation.
16. Define supervised learning.
17. Define data mining model.
18. What do you mean by attribute selection in classification ?
19. Explain grid based clustering.
20. Explain outlier points with example.
21. Define confidence in association rule.
22. Explain the steps in data pre-processing.

SECTION – C

(Short Essay)

Not to exceed **120** words, answer **any six** questions. **Each** question carries **four** marks : **(6×4=24 Marks)**

23. Where do we use regression in data mining ?
24. Which are the different operations performed with a data cube ?
25. Explain different data transformation operations.
26. Explain the use of text mining.
27. Define meta data with example.
28. Discuss about OLAP.
29. Write short note on Tree Pruning.
30. Explain Baye's theorem.
31. Explain the methods used for distance calculation in clustering.

SECTION – D

(Long Essay)

Answer **any two** questions. **Each** question carries **15** marks : **(2×15=30 Marks)**

32. Discuss about different clustering methods.
33. Explain in detail about rule-based classification.
34. Discuss about association rule mining technique with an example.
35. Discuss about different schemas used for multi-dimensional data models.

———————

(Pages : 3) **E – 1899**

Reg. No. : .....

Name : .....

Sixth Semester B.Sc./B.C.A. Degree Examination, April 2018
Career Related FDP under CBCSS
Group 2(b) : COMPUTER SCIENCE/COMPUTER APPLICATIONS
Elective Course CS1661.3/Core Course CP 1643
Data Mining and Warehousing
(2014 Admission Onwards)

Time : 3 Hours Max. Marks : 80

PART – A
(Very short answer type)

(**One** word to maximum of **one** sentence, answer **all** questions. **Each** question carries **one** mark.)

1. What is information ?
2. _____ is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management decisions.
3. A/an _____ system is market-oriented and is used for data analysis by knowledge workers, including managers, executives and analysts.
4. KDD stands for _____
5. OLAP is used to explore the _____ knowledge.
6. Define metadata.
7. In K-nearest neighbor the input is translated to _____
8. Define decision tree.
9. The process of grouping a set of physical or abstract objects into classes of similar objects is called _____
10. Data objects, which are grossly different from or inconsistent with the remaining set of data, are called _____

P.T.O.

PART – B
(Short answer)

(**Not** to exceed **one** paragraph, answer **any eight** questions. **Each** question carries **two** marks)

11. What do you mean by data mining ?

12. Briefly explain the term data cleaning.

13. List out various steps in data transformation.

14. Compare and contrast database systems and data warehouses.

15. Write short note on market basket analysis.

16. What is the use of apriori algorithm ?

17. What is classification ?

18. What are lazy learners ?

19. Write a note on decision trees.

20. What is cluster analysis ?

21. Why is outlier mining important ?

22. What is the use of dissimilarity matrix in cluster analysis ?

PART – C
(Short essay)

(**Not** to exceed **120** words, answer **any six** questions. **Each** question carries **four** marks)

23. What do you mean by knowledge discovery ?

24. What are the needs of data integration ?

25. How will you generate association rules from frequent item sets ?

26. Explain the process of mining single dimensional boolean association rule.

27. How will you use IF-THEN rules for classification ?

28. Explain the k-nearest negihbor method.

29. Briefly outline the major ideas of Naïve Bayesian classification.

30. What are the different categories of clustering ?

31. Explain different methods for outlier detection.

PART – D
(Long essay)

(Answer **any two** questions. **Each** question carries **15** marks)

32. What do you understand by pre-processing data ? What are the different forms of data pre-processing ?

33. Explain in detail multidimensional data models.

34. Explain classification and prediction. Describe issues regarding preprocessing the data for classification and prediction.

35. Explain the partitioning methods for classifying clustering methods. Write commonly used partitioning methods.
