Model Questions

SKR/KW/24/2065

Faculty of Science & Technology
Seventh Semester B.E. (Information Technology) (C.B.S.) Examination
DATA WAREHOUSING & MINING

Time : Three Hours]   [Maximum Marks : 80

INSTRUCTIONS TO CANDIDATES

(1) All questions carry marks as indicated.

(2) Solve Question No. **1 OR** Question No. **2**.

(3) Solve Question No. **3 OR** Question No. **4**.

(4) Solve Question No. **5 OR** Question No. **6**.

(5) Solve Question No. **7 OR** Question No. **8**.

(6) Solve Question No. **9 OR** Question No. **10**.

(7) Solve Question No. **11 OR** Question No. **12**.

1. (a) Explain three tier architecture of Data warehouse with neat diagram.   7

    (b) Describe the failure of past decision support system.   6

OR

2. (a) What is data mart ? How it differs from data warehouse ?   7

    (b) Explain Life Cycle of Data warehouse with neat sketch.   6

3. (a) Summarize the data pre-processing steps in brief.   7

    (b) What is data reduction ? Explain different methods of data reduction.   6

OR

4. (a) How data cleaning can be handled in Pre-processing ?   6

    (b) Explain MIN-MAX normalization and Z-score normalization.   7

5. (a) What is OLAP ? Discuss basic operations of OLAP with example.   8

    (b) Differentiate between OLTP and OLAP.   6

OR

6. (a) Describe STAR and SNOWFLAKE scheme with example.   8

    (b) Define the following :

         (i)   ROLAP

         (ii)   MOLAP

         (iii) HOLAP.   6

Model Questions

7. (a) Draw and explain Knowledge Discovery process. 8

    (b) What are the various issues in Data mining ? 6

OR

8. (a) With the schematic diagram, describe the architecture of data mining system. 8

    (b) Explain any three applications of data mining in detail. 6

9. (a) Explain market basket analysis in detail with example. 7

    (b) Define the following terms with examples :

        (i) Frequent item set

        (ii) Closed item set. 6

OR

10. A database has five transactions. Let min-sup = 60% and min-conf = 80%.

TID	Items-bough
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I, E}

    Find all frequent item sets using Apriori and FP-growth respectively. Compare the efficiency of two mining processes. 13

11. (a) Write in detail about BI framework 2020 with neat diagram. 7

    (b) Discuss Business intelligence application used in Logistics and Production domain. 6

OR

12. (a) Define Big data. State its challenges. 5

    (b) Describe the architectural component of Hadoop ecosystem. 8

2

Model Questions

PRS/KS/24/2392

Faculty of Science & Technology

Seventh Semester B.E. (Information Technology) (C.B.S.) Examination

DATA WAREHOUSING AND MINING

Time : Three Hours] [Maximum Marks : 80

INSTRUCTIONS TO CANDIDATES

(1) All questions carry marks as indicated.

(2) Solve Question No. **1 OR** Question No. **2**.

(3) Solve Question No. **3 OR** Question No. **4**.

(4) Solve Question No. **5 OR** Question No. **6**.

(5) Solve Question No. **7 OR** Question No. **8**.

(6) Solve Question No. **9 OR** Question No. **10**.

(7) Solve Question No. **11 OR** Question No. **12**.

(8) Assume suitable data wherever necessary.

1. (a) Explain three tier architecture of data warehouse with neat diagram.   7

    (b) Elaborate the terms mean, mode and median. How to determine the mean, mode and median in the given data set. {3, 4, 5, 5, 6, 6, 6, 7, 8, 8, 9, 9} ?   6

OR

2. (a) What is data mart ? How it differs from data warehouse ?   7

    (b) Find the mean deviation about the mean for the following data,

    {6, 7, 10, 12, 13, 4, 8, 12}.   6

3. (a) What is the need of data pre-processing ? Explain in brief.   5

    (b) What is data transformation ? Explain different methods of data transformation by normalization.   8

OR

4. (a) What is data reduction ? Explain different methods of data reduction.   9

    (b) What are different measures of data dispersion ?   4

5. (a) Describe the STAR and SNOWFLAKE scheme with neat sketch. 7

   (b) Explain various types of OLAP operations with example. 7

<center>**OR**</center>

6. (a) Differentiate between OLAP and OLTP. 6

   (b) Write a short note on web based OLAP. 4

   (c) Discuss the architecture of HOLAP in detail with the help of suitable diagram. 4

7. (a) Explain the architecture of data mining with neat diagram. 9

   (b) List and explain the major issues of data mining. 4

<center>**OR**</center>

8. (a) What are several stages involved in knowledge data discovery process ? 5

   (b) Explain the application of data mining in (any **two**) :

     (i) Direct marketing

     (ii) Fraud Detection

     (iii) Insurance and health care. 8

9. (a) Explain market basket analysis in detail. 7

   (b) A database has six transactions, find the frequent itemsets and generate association rule on this. Assume that minimum support threshold (s = 33.33%) and minimum confidence threshold (c = 60%) :

Transaction ID	Item
T1	Hot Dogs, Buns, Ketchup
T2	Hot Dogs, Buns
T3	Hot Dogs, Coke, Chips
T4	Chips, Coke
T5	Chips, Ketchup
T6	Hot Dogs

6

<center>**OR**</center>

Model Questions

10. (a) Write a short note on association rule mining. 4

(b) Explain FP-Growth algorithm, for the following data, with Min_Support = 3, and Draw A Tree :

Transaction ID	Item
T100	{K, E, M, N, O, Y}
T200	{D, E, K, N, O, Y}
T300	{A, E, K, M}
T400	{C, K, M, U, Y}
T500	{C, E, I, K, O}

9

11. (a) Draw and explain map reduce technique with example. 10

(b) Write a note on characteristics of big data. 4

OR

12. (a) Describe architecture of Hadoop. Also give its challenges and application. 10

(b) Write a note on important factors of Business Intelligence. 4