DIP-C-186 **DDS–02**

DIPLOMA EXAMINATION – DECEMBER, 2020.

DATA MINING TECHNIQUES

Time: 3 hours Maximum marks: 75

PART A — (5 × 5 = 25 marks)

Answer any FIVE questions.

1. Discuss Shared memory Architecture.

2. What is Metadata? Explain.

3. Illustrate Online Analytical Processing (OLAP).

4. What is impromptu? Explain.

5. List out the steps involved in the process of knowledge discovery.

6. Explain the greedy algorithm.

7. List out the applications of clustering.

PART B — ($5 \times 10 = 50$ marks)

Answer any FIVE questions.

8. Explain the two types of approaches in data warehouse.

9. Explain the multidimensional data model in detail.

10. Explain MOLAP and ROLAP in detail.

11. Explain the data mining task primitives.

12. List out and explain the major issues in data mining.

13. Illustrate the Bayesian theorem.

14. Explain the K-Means Clustering Method.

———————

DIP-C-187 DDS–03

DIPLOMA EXAMINATION – DECEMBER 2020

BIG DATA ANALYTICS

Time : 3 hours Maximum marks : 75

PART A — (5 × 5 = 25 marks)

Answer any FIVE questions.

1. What is Big Data?: Discuss any three big data applications.

2. Explain matrix vector multiplication on MapReduce.

3. Discuss the two classes of big data technologies.

4. List out the major components of Hadoop and explain.

5. Explain any five common Hadoop shell commands.

6. Discuss about the capacity scheduler.

7. Explain the Hive architecture.

PART B — (5 × 10 = 50 marks)

Answer any FIVE questions.

8. Explain the seven drivers of Big Data.

9. How to move the data in and out of the Hadoop.

10. What is serialization? Explain.

11. Illustrate the HDFS architecture.

12. Give a detailed account on HDFS Federation.

13. Explain the yarn applications in detail.

14. List out the different types of queries and explain any four in detail.

———————

DIP-C-188 **DDS–04**

DIPLOMA EXAMINATION - DECEMBER – 2020

DATA SCIENCES AND BIG DATA ANALYTICS

PROGRAMMING WITH PYTHON

Time:3 Hours **Maximum Marks:75**

PART A (5*5 = 25Marks)

Answer any FIVEquestions.

1 . Explain raising an exception.

2 . Briefly explain shallow copy with an example.

3 . Enumerate on Dynamic type checking with an example.

4 . Explain with an example on Local scope and Global scope.

5 . List the types of Inheritance.

6 . Discriminate on Slot in python with an example.

7 . Explain String template with an example.

PART B (5*10 =50Marks)

Answer any FIVE questions.

8 . Describe Single Line Docstring and Multiline Docstring .

9 . Write a program to sort elements in Dictionary.

10. Write a program for replace and converting upper string to lower string and viceversa.

11. Write a program for switch statement with an example.

12. Enumerate on Inheritance with an example.

13. Explain the basic concepts of Object oriented programming

14. Elloborate on Opening files in python.

DIP-C-189 DDS–05

DIPLOMA EXAMINATION – DECEMBER 2020S

STATISTICAL ANALYSIS WITH R

Time : 3 hours Maximum marks : 75

PART A — (5 × 5 = 25 marks)

Answer any FIVE questions.

1. What is Qualitative data? Explain with example.

2. Explain the ANOVA based on any one basic experimental design.

3. Illustrate – Sign test.

4. Discuss Estimated Simple Regression Equation.

5. Illustrate bootstrap for lowness curve.

6. List out and explain the basic functions of statistical package R.

7. How to create plots in R? Explain.

PART B — (5 × 10 = 50 marks)

Answer any FIVE questions.

8. Explain the histogram with example.

9. Discuss the following.

   (a) Mean

   (b) Median

   (c) Quertile

   (d) Percentile

10. Explain about Standardized Residual.

11. Elaborate Prediction Interval for MLR.

12. Give a detailed account on permutation test.

13. Explain the Apply function with example.

14. Explain the Randomized Block Design with example.

———————