Plot No. 2, Knowledge Park-III, Greater Noida (U.P.)–201306

POST GRADUATE DIPLOMA IN MANAGEMENT (2019 -21)
MID TERM EXAMINATIONS (TERM -V)
Academic Session- 2020-21

Subject Name: Big Data Analytics	Time: 01.30 hrs
Sub. Code: PGIT-03	Max Marks: 20

Note:
1. Writing anything except Roll Number on question paper will be deemed as an act of indulging in unfair means and action shall be taken as per rules.
2. All questions are compulsory in Section A, B & C. Section A carries 1 Case Study of 8 marks. Section B carries 3 questions of 2 marks each and Section C carries 2questions of 3 marks each.

<u>**SECTION – A**</u>　　　　　　　**04+04 = 08 Marks**

Q. 1: Case Study:

Bank of America – Using Data to Leverage Customer Experience

10 years ago, Bank of America was one of the first financial companies to provide mobile banking to its customers. Recently, BoA launched Erica which is their first virtual financial assistant. It is considered as the world's finest innovation in finance domain. Currently, Erica is serving as a customer advisor to more than 45 million users around the world. Erica also makes use of Speech Recognition to take customer inputs, which is a technological advancement in the field of Data Science.

Furthermore, several other banks like BoA are making use of Data Science and predictive analytics. Using data science, banking industries are able to detect frauds in payments and customer information. It also prevents frauds regarding insurances, credit cards, and accounting. In order to minimize the losses, a bank needs to detect fraud sooner. In order to carry this out, banks employ data scientists to use their quantitative knowledge where they apply algorithms like association, clustering, forecasting, and classification.

Risk modeling is another important area that is supervised by the banks to regulate financial activities. Using Machine Learning, banks are able to minimize risk modeling. Through analytical solutions, banks can make data-driven decisions that are based on transparency and risk analysis. Furthermore, Bank of America detected the high-risk accounts using this technology of big data.

Various banks like BoA are understanding their customers through an intelligent customer segmentation approach. Through various data-mining techniques, banks are able to segment their customers in the high-value and low-value segments. There are various techniques that a data scientist makes use of such as clustering, logistic regression, decision trees to help the banks to understand the Customer Lifetime Value (CLV) and take group them in the appropriate segments.

(A) Discuss the importance of Big Data Analytics for Banks which could result in more efficiency and effectiveness of their processes?

(B) What are the threats for Big Data Analytics in the organisations like BoA which could generate new problems? How to handle them strategically?

<u>**SECTION – B**</u>　　　　　　　**02×03 = 06 Marks**

Q. 2: Discuss the CRISP-DM model stages for data mining in traditional environment.

Q. 3: How do we create a Matrix in R Programming? Write the command for creating the following matrix:

Ravi	Rashi	Rakhi
Sanjay	Sakshi	Saksham

Q. 4: What is the difference between a Vector and List? Write the command for accessing data from a Vector.

SECTION – C   03×02 = 06 Marks

Q.5. Discuss the major characteristics of Big Data with suitable example.

Q. 6. Write the command to create a dataframe for 05 records of employees with four variable with the structure given below. Also show the final shape of dataframe:

```
'data.frame':  5 obs. of  4 variables:
 $ employeeID: num  1001 1002 1003 1004 1005
 $ empNAME   : chr  "Raj" "Vidisha" "Enam" "Shubhendu" ...
 $ empSALARY : num  20000 25000 30000 28000 35000
 $ empJOIN   : Date, format: "2019-01-01" ...
```

Mapping of Questions with Course Learning Outcome

COs	Question Number(s)	Total Marks Allocated to the CO
CO1	1 (a)	**04**
CO2	1 (b)	**04**
CO3	2,5	5
CO4	3,4,6	7
CO5	Not covered in Mid Term	

After completion of course, students should be able to:

- Understand the evolution of data, its storage and applications

- Differentiate between generic data and Big Data

- Understand the Big Data ecosystem, along with its various components

- Apply Big Data Analytics Software & Tools for solving the business problems

- Apply Data Visualization to overcome business challenges

Note: Font: Times New Roman, Font size: 12.