Total No. of Questions : 6]

SEAT No. :

P546

[Total No. of Pages : 2

APR - 18/TE/Insem. - 149

T.E. (Information Technology)

DATA SCIENCE AND BIG DATA ANALYTICS

(2015 Pattern) (Semester - II)

Time : 1 Hour] *[Max. Marks : 30*

Instructions to the candidates:

1) ***Answer Q1 or Q2, Q3 or Q4, Q5 or Q6.***
2) ***Draw neat diagrams wherever necessary.***
3) ***Figures to the right indicate full marks.***
4) ***Assume suitable data, if necessary.***

Q1) a) Justify your answer with example "Data science and Big Data are same or different". **[5]**

b) Explain the role of Shared Everything & Shared Nothing architecture in Big Data. **[5]**

OR

Q2) a) Define with example Big Data with 5V's. **[5]**

b) Enlist the impact of learning approaches in Big Data? Explain different kinds of learning approaches. **[5]**

Q3) a) Find the first three powers of following transition matrix using Markov chain. $D=\begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$ **[6]**

b) Determine distinct elements in below input stream of integers using Flajolet Algorithm. Consider Hash function h(X) = 6X + 1, X = 1, 3, 2, 1, 2, 3, 4, 3, 1, 2, 3, 1. **[4]**

OR

P.T.O.

Q4) a) Solve following problem using Markov chain. **[6]**

In the Dark Ages, Harvard, Dartmouth and yale admitted as per below scenario. Assume that, 80% of the sons of Harvard men went to Harvard and rest went to Yale. 40% of the sons of Yale men went Yale and rest split evenly between Harvard and Dartmouth. Of the Sons of Dartmouth men, 70% went to Dartmouth, 20% to Harvard and 10% to Yale.

i) Find the probability that the grandson of a man from Harvard went to Harvard.

ii) Modify the above by assuming that the son of a Harvard man always went to Harvard. Again find the probability that the grandson of a man from Harvard went to Harvard.

b) Explain Bloom filter with proper example. **[4]**

Q5) a) Explain Hadoop Ecosystem in detail. **[6]**

b) Differentiate between SQL and NoSQL Databases with example. What is the need to develop Big Data applications using NoSQL databases? **[4]**

OR

Q6) a) Explain HDFS Read & Write operations in detail. **[6]**

b) What is the need of Map-Reduce in Big Data? Define the architecture of Map-Reduce on Hadoop. **[4]**