Sardar Patel Institute of Technology

Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai-400058, India

(Autonomous College Affiliated to University of Mumbai)

Mid Semester Examination
September 2019

Max. Marks: 20	Duration: 1 Hr
Class: M.Tech	Semester:I
Course Code: CE912	Branch: Computer

Name of the Course: Big Data Analytics and Management

Instructions:
(1) All questions are compulsory
(2) Assume suitable data if necessary
(3) Draw neat diagram wherever required.

Q No.		Max. Marks	CO	BL
Q.1	Illustrate the Hadoop Ecosystem and its core component at the input level, storage level, processing level, output level. What are the data injection tools for Structured, unstructured and semi structured data.	4+3	CO1	L2
Q.2	Elaborate the the concept of checkpointing by taking into consideration fsimage and edits-log files OR Consider a Job of 1 TB of data. Scenario 1 : one machine with 4 I/O channels each channel of 100 Mbps Scenario 2 : Clusture of 10 machines with 4 I/O channels Each Channel- 100 Mbps What is the amount of time required in both the scenarioes to read 1 TB of data	3	CO2	L3
Q.3	Write the rules followed when streams is represented by buckets in DGIM Algorithm. How an input bit stream 1010110001011110110010110 is partitioned into buckets. What happens when a new input stream 10101011 is added to it?	3+2+2	CO4	L3
Q.4	Draw the Structure of web and explain the components of it. OR Define Page rank. How spider traps and dead ends affect the Page rank.	1+2	CO4	L2