



**VI SEMESTER B.TECH. (COMPUTER SCIENCE & ENGINEERING)**

**MAKEUP EXAMINATIONS, JUNE/JULY 2019**

**SUBJECT: BIG DATA MODELING AND MANAGEMENT [CRA-4005]**

**REVISED CREDIT SYSTEM**

**(14/06/2019)**

Time: 3 Hours

MAX. MARKS: 50

**Instructions to Candidates:**

- ❖ Answer **ALL FIVE** questions.
- ❖ Missing data may be suitable assumed.

- 1A.** Explain the technology that is contributing to the catalyzing power of data science and that led to the launch of the big data era. **2M**
- 1B.** How can Big Data help to monitor, predict and manage a firestorm and what makes this a Big Data problem? **5M**
- 1C.** Explain the big data characteristics with respect to the velocity and why real time processing is important? **3M**
- 2A.** With a neat diagram, explain the different steps in Data Science? **3M**
- 2B.** What is a commodity cluster and how does it achieve data level parallelism? Describe fault tolerance and the solutions implemented for fault tolerance? **5M**
- 2C.** Describe what a right programming model is. Give an example for big data programming model and justify the same. **2M**
- 3A.** What is Storage Area Network? List all the advantages and disadvantages of SAN. **3M**
- 3B.** Explain the two divisions of data operations. **2M**
- 3C.** Explain semi structured data model with an example **5M**

- 4A.** Write note on structural constraints with an example. **3M**
- 4B.** With example, explain the different types of graph. **4M**
- 4C.** How format of CSV file helps in storing contents of relations? **3M**
- 5A.** Explain the different challenges in analyzing streaming data systems. **4M**
- 5B.** Explain any six desirable characteristics of a Big Data Management system. **3M**
- 5C.** Explain the schema-on write and schema-on read used in data warehouse and Data Lake. **3M**