COURSE CODE (CREDITS): 22M11CI213 (3)   MAX. MARKS: 15

COURSE NAME: Big Data Analytics

COURSE INSTRUCTOR: Deepak Gupta   MAX. TIME: 1 Hr.

Note: 1) All questions are compulsory. Marks and COs are indicated against each question. 2) Attempt questions in the given sequence. 3) Be precise in your answers. 4) Write neatly.

Q 1. Discuss the concept of big data along with various challenges and opportunities.   [3]   CO1

Q 2. Apache Spark has already become the de facto framework for big data analytics. Support   [3]   CO1
your explanation with relevant examples.

Q 3. Differentiate data lake and data warehouse in terms of data type, schema and processing [3]   CO1
model with suitable examples.

Q 4. Draw a block diagram to explain the concepts of NameNode, DataNode and metadata in   [3]   CO2
HDFS.

Q 5. Consider a set of 3 webpages with initial rank 1/3. Apply Page Rank algorithm to calculate [3] CO4
the final page rank of each page.