

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST-I (EXAMINATION - MARCH 2023)

M.Tech. - II Semester (CSE)

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 and its characteristics – velocity, variability and veracity. [3] CO1
- Q 2. Apache Spark is a streaming platform. How does it differ from Apache Hadoop? Support your explanation with relevant examples. [3] CO1
- Q 3. Discuss following big data challenges with suitable examples. [3] CO1
- (a) Data storage
 - (b) Data privacy and ethics
 - (c) Data visualization
- Q 4. How does HDFS implement data/block replication? Explain it by considering NameNode, DataNodes and metadata. [3] CO2
- Q 5. Explain different phases of MapReduce by considering a word count example. [3] CO2