

TILAK MAHARASHTRA VIDYAPEETH, PUNE
MASTER OF COMPUTER APPLICATIONS
(Specialization in Data Science)
EXAMINATION :MAY-2024
SEMESTER - II
Sub: Big Data Analytics(MCDS23-203)

Date :25/05/2024

Total Marks : 60

Time: 10.00 am to 12.30 pm

Instruction:

1. All questions are compulsory unless and otherwise stated.
2. Bold figures to the right of every question are the maximum marks for that question.
3. Candidates are advised to attempt questions in order.
4. Answers written illegibly are likely to be marked zero.
5. Use of scientific calculators, Log tables, Mollier Charts is allowed.
6. Draw neat and labelled diagram wherever necessary.

Q.1. Answer the following in 2-3lines. (Any 5) (10)

1. What is Big Data?
2. What are the Core Components of Hadoop?
3. What is Yarn?
4. What are the modes of pig execution?
5. List out sources of data explosion.
6. What is the importance of big data?
7. What is the role of job tracker?

Q.2. Answer the following in short. (Any 4) (20)

1. Explain characteristics of big data.
2. Write a note on: Batch processing in Hadoop.
3. Write a short note on : Node Manager
4. Write a short note on Pig Join
5. What are the various V's in big data?
6. What is schema?

Q.3. Answer the following in detail. (Any 3) (30)

1. Describe Types of Data in big data.
2. Describe Hadoop architecture in detail.
3. Explain input file formats in details.
4. Describe various data types in pig.
5. What is sqoop and what are measured operations performed in sqoop?
