

Roll No.

Total Printed Pages - 9

F- 1064

M.Sc. (IT) (Fourth Semester)
EXAMINATION, May-June, 2022
(New Course)
Big Data Analytics
(403)

*Time : Three Hours]**[Maximum Marks:100**[Minimum Pass Marks :40***Note: Attempt all sections as directed.**

Section - A
(Objective Type Questions)

(1 mark each)**Note: Attempt all questions.****Choose the correct answer.**

1. Discovering interesting patterns and knowledge from massive amount of data is called-
 - (A) Data Warehouse
 - (B) Data Analysis
 - (C) Knowledge Discovery Databases
 - (D) Data Mining

2. Which is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in decision making process?
 - (A) Knowledge Discovery Databases
 - (B) Web Mining
 - (C) Data Mining
 - (D) Data Warehouse
3. Structured and systematic analysis at each step before proceeding to the next is called:
 - (A) Spiral
 - (B) Interactive
 - (C) Prototype
 - (D) Waterfall
4. Which is not part of Post-Processing?
 - (A) Pattern selection
 - (B) Pattern interpretation
 - (C) Pattern evaluation
 - (D) Pattern discovery
5. Which is not data classification?
 - (A) Semi Structured Data
 - (B) Hybrid Structured Data
 - (C) Structured Data
 - (D) Unstructured Data

P.T.O.**F- 1064**

[3]

6. Which are not main components of Big Data?
 - (A) YARN
 - (B) SQL Server
 - (C) HDFS
 - (D) MapReduce
7. Big data analytics technologies are necessary to:
 - (A) Formulate eye - catching charts and graphs
 - (B) Extract valuable insights from the data
 - (C) Integrate data from internal and external sources
 - (D) None of these
8. Which is the data model built with small data-
 - (A) Physical
 - (B) Machine Learning
 - (C) Artificial and Expert
 - (D) Relational
9. Which was inspired by Google MapReduce and Google File System papers?
 - (A) Mapreduce
 - (B) YARN
 - (C) Hadoop
 - (D) Cruise control

F-1064

P.T.O.

[4]

10. Which of the following genres does Hadoop produce?
 - (A) RDBMS
 - (B) JAX-RS
 - (C) Distributed file system
 - (D) Java Message Service
11. Which is the architectural center of Hadoop that allows multiple data processing engines?
 - (A) Chuckwa
 - (B) Hive
 - (C) YARN
 - (D) Incubator
12. Which Tracker allocates work to the tracker nearest to the data with an available slot.
 - (A) Task
 - (B) Process
 - (C) Job
 - (D) None of these
13. In which of the following Data life cycle is managed in MongoDB schema Design?
 - (A) Insert, Update and Delete
 - (B) Grant, Permission and View
 - (C) Read and write ratio, Query type
 - (D) Application Side

F-1064

[5]

14. Which is used for storing documents and making remote procedure calls in MongoDB?
- (A) BSON
 - (B) JSON
 - (C) Native
 - (D) Divers
15. Which document are stores in object data type?
- (A) Embedded documents
 - (B) Database documents
 - (C) Query documents
 - (D) Web documents
16. Which is a not session command?
- (A) killCursors
 - (B) endSession
 - (C) commitTransaction
 - (D) killSessions
17. Which is not part of Pig Latin Data Model?
- (A) Tuple
 - (B) Bag
 - (C) Field
 - (D) Relationship

F-1064

P.T.O.

[6]

18. Which is not Apache Pig Execution Mechanism?
- (A) Interactive Mode
 - (B) Batch Mode
 - (C) Embedded Mode
 - (D) Collection Mode
19. You can run Pig in interactive mode using the _____ shell.
- (A) PigRunner
 - (B) File System
 - (C) HDFS
 - (D) Grunt
20. Which is not part of Hive Data Model?
- (A) Table
 - (B) Buckets
 - (C) Partition
 - (D) Bag

Section - B

(Very Short Answer Type Questions)

(2 marks each)

Note: Attempt all questions. Answer using 2-3 sentences.

1. What is Data Cube?
2. What is attributes?

F-1064

[7]

3. What is difference between big data and small data?
4. What is Big data Types?
5. What is Hadoop Distributed File System?
6. What is Hadoop ecosystem?
7. Why are use NoSQL?
8. What is MongoDB?
9. What is Local Mode in Apache Pig?
10. What is Atom?

Section - C

(Short Answer Type Questions)

(3 marks each)

Note: Attempt all questions. Answer precisely using <75 words.

1. What do you mean by KDD? Explain it.
2. What are OLAP operations?
3. What are the importance of Big Data?
4. What are the challenges of Conventional System?
5. What are Hadoop core components?
6. What are features of Hadoop?
7. What are advantages of NoSQL?
8. What are the difference between RDBMS and MongoDB?
9. What is Apache Pig Execution?
10. What are the Data type of Hive?

[8]

Section - D

(Long Answer Type Questions)

(6 marks each)

Note: Attempt all questions. Answer precisely using 150 words.

1. What do you mean by Data Warehouse? Explain the Data Warehouse Applications.

OR

Explain the following:

- (i) Architecture of Data Mining System
- (ii) Data Mining Functionality

2. What do you understand by Data Classification? Write the different type of Data.

OR

Explain the following:

- (i) Big data Analytics Applications
- (ii) Big data Processing Architectures

3. What do you understand by Hadoop Distributed File System data Storage? Explain the Hadoop Physical Organization.

OR

Explain the following:

- (i) MapReduce Framework
- (ii) Hadoop YARN

[9]

4. What do you mean by NoSQL? Explain the different types of NoSQL database.

OR

Explain the following:

- (i) Features of MongoDB
- (ii) MongoDB Commands

5. What is big data programming languages? Explain the characteristics and limitation of HIVE.

OR

Explain the following:

- (i) Hive Built-in functions
- (ii) Pig Architecture