WALCHAND INSTITUTE OF TECHNOLOGY, SOLAPUR
An Autonomous Institute affiliated to PAH Solapur University, Solapur

Information Technology
Final Year B. Tech. Sem-VII 2024-25 (CBCS pattern)
End Semester Examination

Course: Professional Elective- III- Big Data Analytics **Max. Marks:** 60 **SET: Q**
Course Code : 21ITU7E42T

Day and Date: Monday 02/12/2024 **Time:** 02.30 pm to 05.30 pm

Instructions:
1. All questions are compulsory.
2. Assume suitable data wherever necessary.
3. Draw neat sketches wherever necessary.
4. Q.1 should be solved in first 30 minutes and answers should be written only on the first page of answer book.
5. Supplements will not be provided.
6. Use of nonprogrammable calculator is permitted.

Q. No	Sub Que		Marks
1		**Multiple Choice Questions**	12
	i.	Hive uses _____ for logging. a) logj4 b) log4l c) log4i d) log4j	
	ii.	Which of the following function is used to read data in PIG? a) WRITE b) READ c) LOAD d) None of the mentioned	
	iii.	Pig operates in mainly how many nodes? a) Two b) Three c) Four d) Five	
	iv.	How many formats of SequenceFile are present in Hadoop I/O? a) 2 b) 3 c) 4 d) 5	

v.	Hadoop I/O Hadoop comes with a set of _____ for data I/O. a) Methods b) commands c) classes d) none of the mentioned	
vi.	A _____ serves as the master and there is only one Name Node per cluster. a) Data Node b) Name Node c) Data block d) Replication	
vii.	Running a _____ program involves running mapping tasks on many or all of the nodes in our cluster. a) MapReduce b) Map c) Reducer d) All of the mentioned	
viii.	Point out the correct statement. a) MapReduce tries to place the data and the compute as close as possible b) Map Task in MapReduce is performed using the Mapper() function c) Reduce Task in MapReduce is performed using the Map() function d) All of the mentioned	
ix.	Hive also support custom extensions written in _____ a) C# b) Java c) C d) C++	
x.	Apache HBase is a non-relational database modeled after Google's _____ a) BigTop b) Bigtable c) Scanner d) FoundationDB	
xi.	Which of the following class provides access to configuration parameters? a) Config b) Configuration c) OutputConfig d) None of the mentioned	
xii.	HBase is a distributed _____ database built on top of the Hadoop file system. a) Column-oriented b) Row-oriented c) Tuple-oriented d) None of the mentioned	

WALCHAND INSTITUTE OF TECHNOLOGY, SOLAPUR
An Autonomous Institute affiliated to PAH Solapur University, Solapur

Information Technology
Final Year B. Tech. Sem-VII 2024-25 (CBCS pattern)
End Semester Examination

Course: Professional Elective- III- Big Data Analytics   **Max. Marks**: 60   **SET: Q**

Course Code : 21ITU7E42T

Day and Date: Monday 02/12/2024   **Time:** 02.30 pm to 05.30 pm

Instructions:

1. All questions are compulsory.
2. Assume suitable data wherever necessary.
3. Draw neat sketches wherever necessary.
4. Q.1 should be solved in first 30 minutes and answers should be written only on the first page of answer book.
5. Supplements will not be provided.
6. Use of nonprogrammable calculator is permitted.

Q. No	Sub Que		Marks
Q.2		**Attempt any *Four***	12
	i.	What is a Big data? What are the features Big data?	
	ii.	What are the techniques of Analysing Data with Hadoop?	
	iii.	Illustrate with example how Hadoop Streaming works.	
	iv.	What are the different types of Digital data?	
	v.	Write a short note on Apache Hadoop?	
Q.3		**Attempt any *Two***	12
	i.	Demonstrate Hadoop file system interfaces in HDFS(Hadoop Distributed File System).	
	ii.	What is DFS? Why we need DFS? Elaborate in detail with example.	
	iii.	What is Hadoop? What are the important Features of HDFS(Hadoop Distributed File System)?	

Q.4		**Attempt any _Four_**
	i.	What are the key differences between Hive and Pig?
	ii.	How is Apache Pig different from MapReduce?
	iii.	What are the major components of a Pig execution environment?
	iv.	Explain the different complex data types in Pig.
	v.	What is UDF in Hive? Illustrate in detail?
Q.5		**Attempt any _Two_**
	i.	What is HBASE? Give the difference Between Hbase and Hive.
	ii.	Differentiate between Supervised Learning and Unsupervised Learning.
	iii.	Iilustrate with example : Big Data Analytics with BigR.