SET-A

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, December 2018

Course: Big Data Analytics (CSIB 385)   **Semester: VII**
Programme: B.TECH (CSE) in BAO   **Time: 03 hrs.**
Max. Marks: 100
Instructions:

	SECTION A		
S. No.		**Marks**	**CO**
Q 1	EXPLAIN RDBMS Vs HADOOP	4	CO2
Q2	Give Full form of Jaql?	4	CO3
Q3	Discuss Hadoop consist of which 3 sub projects?	4	CO4
Q4	Explain Traditional Computing versus Stream Computing	4	CO5
Q5	Design Principles of Hadoop lets the system handle which issues automatically	4	CO2
	SECTION B		
Q6	Diagrammatically represent MapReduce Data Flow.	10	CO2,3
Q7	What is Jaql? Why Jaql? Elaborate about Jaql and MapReduce	10	CO 3
Q8	A. Stream Computing is oriented towards which kind of data?	2	CO1,4, 5
	B. Descrive how Scaling in Infosphere Streams is a achieved	2	
	C. out of accelerator and processing element which one is a basic deployment unit.	2	
	D. Name and illustrate which type of operators is used to create a stream from data flowing from the external source	2	
	E. By using accelerators we were able to do what with data?	2	
Q9	A) Explain Big Data, B) Illustrate 3 V's in Big Data. C) Define the Risks Of Big Data, or C) Elaborate about Structure Of Big Data;	10	CO1
	SECTION-C		
Q10	SPL standard toolkit operators are grouped in three Categories. (A) Firstly elaborate two Adapter operators.	20	CO4,5

	(Attempt any one from b and c sub questions)		
	(B) Elaborate any five Relational operators. (C) Elaborate any five Utility operators		
Q11	Describe any ten HDFS Command	**20**	**CO2,3**

SET-B

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, December 2018

Course: Big Data Analytics (CSIB 385)   **Semester: VII**

Programme: B.TECH (CSE) in ECRA   **Time: 03 hrs.**

Max. Marks: 100

Instructions:

SECTION A

S. No.		Marks	CO
Q 1	What is Big Data? Discuss big data in healthcare, transportation and medicine	4	CO1
Q 2	Write a Short note on Hadoop Ecosystem also explain various elements of Hadoop.	4	CO2
Q 3	Differentiate: Apache Pig Vs Map Reduce.	4	CO2
Q 4	Explain "Map Phase" and "Combiner Phase" in Map Reduce.	4	CO2
Q 5	What do you mean by HiveQL Data Definition Language?	4	CO3

SECTION B

Q 6	Discuss the concept of regions in HBase and Storing Big data with HBase.	8	CO3
Q 7	Explain Pig data Model in detail and Discuss how it will help for effective data flow	8	CO4
Q 8	Write a short note on NoSQL databases. List the differences between NoSQL and relational databases?	8	CO4
Q 9 a	Explain Hive components in detail. Also list the features of hive.	8	CO3
	OR		
Q 9 b	Discuss Hadoop YARN in detail with failures in classic Map-reduce.	8	CO3
Q 10 a	Explain Jaql in detail with the help of example.	8	CO3
	OR		
Q 10 b	Explain working of Hive with proper steps and diagram.	8	CO4

SECTION-C

Q 11	Draw and explain HDFS Architecture. Explain the functions of Name Node and Data Node. What is Secondary Name Node? Is it a substitute to the Name node	20	CO1
Q 12 a	Explain in detail Stream programming language (SPL) and also its data flow network. Describe in brief the various operators used in SPL.	20	CO5
	OR		

| Q 12 b | What is Spark? State the advantages of using Apache Spark over Hadoop Map Reduce for big data processing with example. | 20 | CO4 |