


<b>Name:</b>  <b>Enrolment No:</b>		
<b>UNIVERSITY OF PETROLEUM AND ENERGY STUDIES</b> <b>End Semester Examination, December 2021</b>		
<b>Course: Big Data Analytics</b> <b>Program: MBA Business Analytics</b> <b>Course Code: DSBA 8002</b>		<b>Semester: III</b> <b>Time 03 hrs.</b> <b>Max. Marks: 100</b>
<b>SECTION A</b> <b>1. Each Question will carry 2 Marks</b> <b>2. Instruction: Complete the statement / Select the correct answer(s)</b>		
S. No.	Questions	CO
Q1	What are the five Vs of Big Data?	CO1
Q2	Which of the following aren't the sources of Big Data? Only one option is correct.  a) Users b) Applications c) System d) Sensors e) Scanners f) Social Media g) None of these	CO2
Q3	Statistics + Big Data + Machine Learning = ?	CO3
Q4	According to analysts, for what can traditional IT systems provide a foundation when they're integrated with big data technologies like Hadoop?  a) Big data management and data mining b) Data warehousing and business intelligence c) Management of Hadoop clusters d) Collecting and storing unstructured data	CO3
Q5	What are the main components of Big Data?  a) MapReduce b) HDFS c) YARN d) All of these	CO1
Q6	What are the different features of Big Data Analytics?  a) Open-Source b) Scalability c) Data Recovery d) All the above	CO4
Q7	All of the following accurately describe Hadoop, EXCEPT:	CO2

	a) Open source b) Real-time c) Java-based d) Distributed computing approach	
Q8	_____ hides the limitations of Java behind a powerful and concise Clojure API for Cascading.  a) Scalding b) Cascalog c) Hcatalog d) Hcalding	CO4
Q9	Facebook Tackles Big Data With _____ based on Hadoop.  a) Project Prism b) Prism c) ProjectData d) ProjectBid	CO1
Q10	What is a unit of data that flows through a Flume agent?  a) Record b) Event c) Row d) Log	CO4

**SECTION B**

- 1. Each Question will carry 5 Marks**
- 2. Instruction: Write short or brief notes**

Q11	What do you mean by structured vs unstructured vs semi-structured big data?	CO2
Q12	What are linear regression and logistic regression? Give an example of each.	CO4
Q13	What are the advantages and challenges of using Big Data?	CO4
Q14	Explain the data source layer, security layer, and visualization layer in the big data architecture.	CO3

**SECTION C**

- 1. Each Question will carry 10 Marks**
- 2. Instruction: Write medium answers**

Q15	Draw the CRISP-DM (Cross-Industry Standard Process for Data Mining)  OR  What are the advantages and disadvantages of using artificial intelligence?	CO4
Q16	Explain the data loading process in Weka.	CO4
Q17	Define the term 'classifiers' in Weka. Explain their usage.	CO3

**SECTION D**

- 1. Each Question will carry 15 Marks**

**2. Instruction: Write long answers**

Q18	Explain in detail the applications of artificial intelligence in various domains. Mention at least 6 points.  OR  What are the four cloud deployment models? Explain each with the help of labeled diagrams.	CO4
Q19	Explain the decision tree algorithm and the clustering algorithm.	CO4