UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, May 2020

Course: Information Security Intelligence & Compliance Analysis Semester: VIII
Program: B. Tech (CSE+CSF) Time 03 hrs.
Course Code: CSIB 448 Max. Marks: 100

Instructions:

SECTION A (60 Marks)

S. No.		Marks	CO

Q 1	Listed below are some of the various characteristics of Big data. Pick the odd one out.
	    a. Handles a large amount of data
	    b. Distributes redundant data storage
	    c. Is a system that does not allow data processing capabilities.
	    d. Process tasks in parallel
		2	CO1

Q 2	The three 'V' associated with big data analytics are:
	    a. Variability
	    b. Volume
	    c. Velocity
	    d. Variety
		2	CO1

Q 3	Which one of these will not fall in the category of structured data:
	    a. Spreadsheet data
	    b. Sensor Data
	    c. Satellite Data
	    d. Data from a airline reservation system
		2	CO2

Q 4	How is big data different from traditional data sources?
	    a. Big data can be an entirely new source of data
	    b. Only factor is the inability of traditional databases to hold this data
	    c. High speed of incoming data
	    d. All of these
		2	CO2

Q 5	What is the aim of Big Data Analytics?
	    a. Cost Reduction
	    b. New Products and Services
	    c. Faster and better decision making
	    d. All of these
		2	CO1

Q 6	Which of these is not a function of SAS?
	    a. Information Recovery
	    b. Statistical analysis
	    c. Log Generation
	    d. Project Management
		2	CO3

Q 7	Most influential and established tools for security analytics
	  a. Microsoft Access
	  b. Python Programming
	  c. Splunk
	  d. Microsoft excel
		2	**CO2**

Q 8	Examples pertaining to Big Data Analytics are:
	  a. Logs from web server
	  b. Emails from customers and other feedback responses
	  c. Records of mobile phone calls
	  d. All of these
		2	**CO2**

Q 9	Which of these is also called "Google of log files"?
	  a. IDS
	  b. SPLUNK
	  c. Syslog
	  d. None of these
		2	**CO3**

Q 10	Pick the odd one out:
	  a. Syslog
	  b. SNMP
	  c. Code & Scripts
	  d. IDS
		2	**CO3**

Q 11	Which is a tool for log monitoring?
	  a. NAGIOS
	  b. IDS
	  c. IPS
	  d. None of these
		2	**CO3**

Q 12	Data Scientist job includes:
	  a. Clustering of data
	  b. Identifying source of data
	  c. Understanding business challenge
	  d. Cloud Management
		2	**CO2**

Q 13	Who will not be part of a big data team?
	  a. Statistician
	  b. Systems Engineer
	  c. Data Analyst
	  d. Business Analyst
		2	**CO2**

Q 14	Which of the following data sources will aid in real time data analysis?
	  a. Firewall Logs
	  b. Application Logs
	  c. Operating System Logs
	  d. None of these
		2	**CO3**

Q 15	Data mining algorithms require:
	   a. Efficient sampling method
	   b. Storage of intermediate results
	   c. Capacity to handle large amount of data
	   d. All of these
		2	CO1

Q 16	Hidden knowledge can be found by using
	   a. Searching algorithm
	   b. Pattern recognition algorithm
	   c. Searching algorithm
	   d. Clues
		2	CO3

Q 17	What are the technical factors driving Big Data adoption?
	   a. Decreasing cost of storage
	   b. Flexibility and cost effectiveness of data centers and cloud computing
	   c. Development of new frameworks (such as Hadoop)
	   d. All of these
		2	CO2

Q 18	Which if these is not a responsibility of a data explorer?
	   a. Identification of data
	   b. Clustering of data
	   c. Analysis of data
	   d. Mining of data
		2	CO4

Q 19	A business intelligence product named Watson, is developed by:
	   a. Google
	   b. Microsoft
	   c. IBM
	   d. Facebook
		2	CO1

Q 20	Pick the odd one out.
	   a. Registry
	   b. Event Logs
	   c. Schemas
	   d. File System
		2	CO3

Q 21	What must not be considered while analyzing big data?
	   a. Type of data
	   b. Visualization
	   c. Analysis
	   d. Ingestion
		2	CO2

Q 22	Pick the odd one out.
	   a. Watson
	   b. Apache Hadoop
	   c. Cassandra
	   d. Knime
		2	C02

Q 23	The following is not correct about apache Hadoop.
	    a. It is product of Google
	    b. Uses Map reduce programming model
	    c. Hadoop framework is mostly written in Java
	    d. Apache Hadoop framework is composed of different modules
		2	**CO2**

Q 24	Which of the following data types will not play a role in information security analytics strategy?
	    a. User activity
	    b. Firewall Logs
	    c. Amount of data
	    d. Application Logs
		2	**CO4**

Q 25	Required skills of a data scientist are:
	    a. Technical Expertise
	    b. Curiosity
	    c. Statistical Knowledge
	    d. All of these
		2	**CO3**

Q 26	Pick the odd one out.
	    a. Streaming data from sensors
	    b. Unstructured Data
	    c. Map Reduce
	    d. Structured Data
		2	**CO2**

Q 27	Which if these is not a phase of big data adoption process?
	    a. Distributed file system
	    b. Application Requirement
	    c. Cluster Design
	    d. Hardware Architecture
		2	**CO3**

Q 28	Identify the skill or skills that is must for a big data team member
	    a. Mathematics
	    b. Data Analysis
	    c. Data Visualization
	    d. All of these
		2	**CO3**

Q 29	"Leverage past data to understand why something happened or to predict what will happen in the future across various scenarios" is-
	    a. Predictive Analysis
	    b. Prescriptive Analysis
	    c. Descriptive Analytics
	    d. None of these Incorrect
		2	**CO4**

Q 30	Why Big Data matters?
	    a. Improves understanding of your customer
	    b. Enhances decision making
	    c. Enhances business operations
	    d. All of these
		2	**CO3**

	SECTION B (20 Marks)		
Q 1	Highlight the importance of big data visualization. Explain few common platforms or strategies. **OR** Explain the process of big data classification. List down the various criterion of data classification.	**10**	**CO1** **CO4**
Q 2	Compare using example the structured data from unstructured data. Also elaborate the various strategies of data acquisition **OR** Show the importance of biometric security with respect to information retrieval and data accessing.	**10**	**CO3**
	SECTION C (20 Marks)		
Q 1	Explain the various phases involved in creating big data team. Identify the factors affecting network configuration manager	**10**	**CO2**
Q 2	Elaborate the steps involved in securing big data. In context to information security list down the benefits of security analytics.	**10**	**CO4**