

Name:

Enrolment No:



**UNIVERSITY OF PETROLEUM AND ENERGY STUDIES**  
**End Semester Examination, May 2019**

**Course: Business Intelligence**  
**Program: B.Tech (OGI)**  
**Course Code: CSIB 383**

**Semester: 8<sup>th</sup>**  
**Time 03 hrs.**  
**Max. Marks: 100**

**Instructions: (All questions are compulsory)**

**SECTION A**

		<b>Marks</b>	<b>CO</b>
Q 1	Explain Centralized and Decentralized BI.	4	CO1
Q 2	Describe Outlier Analysis and its applications.	4	CO3
Q 3	Discuss the issues related to Data Mining.	4	CO2
Q 4	Compare OLAP and OLTP.	4	CO1
Q 5	Explain Data Warehouse (DWH). What are the advantages and disadvantages of using DWH?	4	CO1

**SECTION B**

Q 6	What is K-means clustering? Use K-means algorithm to create three clusters for the given set of values {1, 5, 7, 10, 19, 20, 31, 39, 40}.	10	CO4
Q 7	Explain the classifier accuracy evaluation techniques.	10	CO5
Q 8	What is structured data? Explain its sources and how it can be managed?	10	CO2
Q 9	Write short notes on: a) Regression b) Analytics c) Risk Management and Mitigation	10	CO3
	OR		
	Design a BI system for fraud detection in Banking by describing all the steps from Data Collection to Decision Making.	10	CO3

**SECTION-C**

Q 10	Build a prototype dashboard which will provide a hospital with features and performance that meet their objectives. Use any Data mining technique.	20	CO4
Q 11	Explain Density and Distribution Clustering. Consider the similarity matrix given below.	20	CO5

Show the hierarchy of clustering created by the single-link clustering algorithm.

	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>	<b>P5</b>	<b>P6</b>
<b>P1</b>	0.70	0.40	0.65	0.05	0.20	1.00
<b>P2</b>	0.40	0.65	0.35	0.20	1.00	0.70
<b>P3</b>	0.65	0.05	0.20	1.00	0.70	0.40
<b>P4</b>	0.05	0.20	1.00	0.85	0.40	0.65
<b>P5</b>	0.20	1.00	0.70	0.40	0.65	0.05
<b>P6</b>	0.05	0.35	0.40	0.65	0.85	1.00

OR


Write short notes on the following:

- a) Spatial Data Mining
- b) Periodic crawler vs Incremental crawler
- c) Data Integration and Transformation
- d) Centroid Clustering

**20**

**CO5**

**Model Question Paper (Blank) is on next page**

Name:			
Enrolment No:			
<b>UNIVERSITY OF PETROLEUM AND ENERGY STUDIES</b> <b>End Semester Examination, May 2019</b>			
Course: Business Intelligence		Semester: 8 <sup>th</sup>	
Program: B.Tech (OGI)		Time 03 hrs.	
Course Code: CSIB 383		Max. Marks: 100	
<b>Instructions: (All questions are compulsory)</b>			
<b>SECTION A</b>			
		<b>Marks</b>	<b>CO</b>
Q 1	Briefly explain all the data stream methodologies.	4	CO1
Q 2	Explain OLAP operations also state its applications.	4	CO5
Q 3	How mining sequence patterns in transactional Databases is done?	4	CO3
Q 4	Compare Database and Data-warehouse.	4	CO1
Q 5	Explain SAAS. What are the advantages and disadvantages of using SAAS?	4	CO1
<b>SECTION B</b>			
Q 6	What is K-means clustering? Use K-means algorithm to create three clusters for the given set of values {2,4,6,7,9,10,11,15,29,40}.	10	CO4
Q 7	Explain the use of BI for Process improvement.	10	CO2
Q 8	What is semi-structured data? Explain its sources and how it can be managed?	10	CO2
Q 9	Write short notes on: d) Classification e) Prediction f) Clustering	10	CO3
OR			
	Design a BI system for fraud detection by describing all the steps from Data Collection to Decision Making.	10	CO3
<b>SECTION-C</b>			
Q 10	Build a BI report which will provide ecommerce company with features and performance that meet their objectives. Use any Data mining technique.	20	CO4
Q 11	Explain Hierarchical clustering. Consider the similarity matrix given below. Show the hierarchy of clustering created by the single-link clustering algorithm.	20	CO5

	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>	<b>P5</b>	<b>P6</b>
<b>P1</b>	1.00	0.70	0.65	0.40	0.20	0.05
<b>P2</b>	0.70	1.00	0.95	0.70	0.50	0.35
<b>P3</b>	0.65	0.95	1.00	0.75	0.55	0.40
<b>P4</b>	0.40	0.70	0.75	1.00	0.80	0.65
<b>P5</b>	0.20	0.50	0.55	0.80	1.00	0.85
<b>P6</b>	0.05	0.35	0.40	0.65	0.85	1.00

OR

Write short notes on the following:

- e) Crawlers
- f) Harvest System
- g) Virtual Web View

Also explain Web content mining with respect to above mentioned programs.

**20**

**CO5**