Enrollment No: _____ Exam Seat No: _____ C. U. SHAH UNIVERSITY Winter Examination-2019

Subject Name : Business Intelligence and Big Data Analytics

Subject Code : 4TE07BIB1			Branch: B.Tech (CE)		
Seme	ester: 7	7 Date : 18/11/2019	Time : 10:30 To 01:30	Marks : 70	
Instru (1 (2 (3 (4	1) Use 2) Instr 3) Drav 4) Assu	of Programmable calculator & any outcomes written on main answer boo v neat diagrams and figures (if nece ume suitable data if needed.	other electronic instrument is prook are strictly to be obeyed. essary) at right places.	ohibited.	
Q-1 Atter	a) b) c) d) e) f) g) h) i) j) k) l) m) n) npt any	Attempt the following questions What is grid computing? What is neural network? Define E-commerce. What is NoSQL? Where it can be What is S3? Give definition of Business Intelli Give definition of data visualization What is cluster analysis? What is cluster analysis? What is Big Data? What is OLAP? What is market basket analysis? What is Zookeeper? What is Apache Spark? What is the default HDFS block sit four questions from Q-2 to Q-8	: used? gence on. ize?	(14)	
Q-2	(a) (b)	Attempt all questions Explain four V's of big data in det List out and explain best practices	tail. for big data analytics.	(14)	
Q-3	(a) (b)	Attempt all questions Write a short note on Hadoop. Explain Descriptive analytics, pre- analytics in detail.	dictive analytics and perspective	(14)	
Q-4	(a) (b)	Attempt all questions Explain basic clustering methods in What are the benefits of Big Date How Big Data Analytics can be cities.(Discuss one application)	in detail. a? Discuss challenges under Bi e useful in the development of	(14) g Data. f smart Page 1 of 2	



Q-5		Attempt all questions	
-	(a)	Explain Map reduce architecture and it's working in detail.	(07)
	(b)	What is supervised and unsupervised learning?	(03)
	(c)	Explain CAP theorem in brief.	(04)
Q-6		Attempt all questions	
-	(a)	Explain the architecture of Hive. List out features of Hive.	(07)
	(b)	How RDBMS differs from HBase?	(04)
	(c)	Write down the features of HDFS.	(03)
Q-7		Attempt all questions	(14)
-	(a)	Explain K-means algorithm in detail with suitable example.	
	(b)	Explain CRUD operations in MongoDB.	
0-8		Attempt all questions	(14)

Attempt all questions
(a) You are given the transaction data shown in the Table below from a fast food restaurant. There are 9 distinct transactions (order:1-order:9) and each transaction involves between 2 and 4 meal items. There are a total of 5 meal items that are involved in the transactions. For simplicity we assign the meal items short names (M1 –M5) rather than the full descriptive names (e.g., Big Mac).

Meal Item	List of Item IDs	Meal Item	List of Item IDs
Order:1	M1, M2, M5	Order:6	M2, M3
Order:2	M2, M4	Order:7	M1, M3
Order:3	M2, M3	Order:8	M1, M2, M3, M5
Order:4	M1, M2, M4	Order:9	M1, M2, M3
Order:5	M1, M3		

For all of the parts below the minimum support is 2/9(.222) and the minimum confidence is 7/9(.777). Note that you only need to achieve this level, not exceeds it. Apply the Apriori algorithm to the dataset of transactions and identify all frequent k-item sets. Show all of your work. You must show candidates but can cross them off to show the ones that pass the minimum support threshold.

(b) Write a short note on intellectual property challenges for big data.

