C.U.SHAH UNIVERSITY
Summer Examination-2018

Subject Name: Business Intelligence and Big Data Analytics

Subject Code: 4TE07BIB1   **Branch: B.Tech (CE)**

Semester: 7   **Date: 26/03/2018**   **Time: 10:30 To 01:30**   **Marks: 70**

Instructions:
1. (1) Use of Programmable calculator & any other electronic instrument is prohibited.
2. (2) Instructions written on main answer book are strictly to be obeyed.
3. (3) Draw neat diagrams and figures (if necessary) at right places.
4. (4) Assume suitable data if needed.

Q-1   **Attempt the following questions:**

a)	What do the four V's of Big Data denote?	**01**
b)	Differentiate between Structured and Unstructured data.	**01**
c)	What are the main components of a Hadoop Application?	**01**
d)	What is Big Data?	**01**
e)	What is a block in HDFS?	**01**
f)	What is commodity hardware?	**01**
g)	What is the role of Zookeeper in HBase architecture?	**01**
h)	What are the most commonly defined input formats in Hadoop?	**01**
i)	What are real-time industry applications of Hadoop?	**01**
j)	Explain Metastore in Hive.	**01**
k)	What is Zookeeper?	**01**
l)	Write difference between Hive and RDBMS.	**01**
m)	How is Hadoop different from other parallel computing systems?	**01**
n)	What all modes Hadoop can be run in?	**01**

Attempt any four questions from Q-2 to Q-8

Q-2   **Attempt all questions**

a)	Write a note on NoSQL databases. List the differences between NoSQL and relational databases.	**07**
b)	What are the advantages of Hadoop? Explain Hadoop Architecture and its Components with proper diagram.	**07**

Q-3   **Attempt all questions**

a)	Explain working of Hive with proper steps and diagram.	**07**
b)	What are the benefits of Big Data? Discuss challenges under Big Data. How Big Data Analytics can be useful in the development of smart cities.(Discuss one application)	**07**

Q-4 **Attempt all questions**

a) Write Map Reduce code for counting occurrences of specific words in the input **07**
text file(s). Also write the commands to compile and run the code.

b) Define HDFS. Discuss the HDFS Architecture and HDFS Commands in brief. **07**

Q-5 **Attempt all questions**

a) What is Hbase? Discuss in detail the data model and Implementation Aspect. **07**

b) How big data analysis helps businesses increase their revenue? Give example. **07**

Q-6 **Attempt all questions**

a) Explain the steps of the "Apriori Algorithm" for mining Frequent Item set. **07**

b) Explain working of following phases of Map Reduce with one common example. **07**
(i) Map Phase (ii) Combiner Phase (iii) Shuffle and Sort Phase (iv) Reducer
Phase

Q-7 **Attempt all questions**

a) What is Market Basket Analysis? Explain Association Rules with Confidence & **07**
Support.

b) Explain Multilayer feed-forward neural network with example. **07**

Q-8 **Attempt all questions**

a) Explain Visual Data Analysis Techniques in Details. **07**

b) Write the steps of the K-means Clustering Algorithm. Also state its limitation. **07**