**DHARMSINH DESAI UNIVERSITY, NADIAD**

**FACULTY OF TECHNOLOGY**

**B.TECH. SEMESTER VII [CE]**

**SUBJECT: (CE720) BIG DATA ANALYTICS**

*[*

| **Examination** | **: First Sessional** | **Seat No** | **: \_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
| --- | --- | --- | --- |
| **Date** | **: 28/08/2020** | **Day** | **: Friday** |
| **Time** | **: 10:00 TO 11:15 AM** | **Max. Marks** | **: 36** |



| **Q.1** | **Do as directed.** | | | |
| --- | --- | --- | --- | --- |
| (a) | List at least two responsibilities of namenode in Hadoop. | | **[2]** |
| (b) | What is checkpointing in hdfs? | | **[2]** |
| (c) | List out two components of resource manager with their duties. | | **[2]** |
| (d) | Where and how does Big Data and Analytics fit in the realm of Internet of Things? | | **[2]** |
| (e) | Explain the concept of schema on read with any application scenario. | | **[2]** |
| (f) | Describe sharding and policies made available for the same by MongoDB. | | **[2]** |
|  | | | | |
| **Q.2** | Attempt ***Any Two*** from the following questions. | | |  |
| (a) | Explain application workflow in YARN. | | **[6]** |
| (b) | Explain following terms with respect to mapreduce: Task, Task attempt, Failed Task, Killed task, Speculative execution, InputSpilt | | **[6]** |
| (c) | Explain write operation in Cassandra with figure and example. | | **[6]** |
|  | | | | |
| **Q.3** | (a)  (b) | Give overview of hadoop ecosystem and related components.  Write mongoDB queries with solutions to following requirements:  (i) Create a database 'bdaspecial' and collection 'WordCloud'. Also show steps on how to confirm their presence.  (ii) How do you import records from csv file into mongoDB, command line?  (iii) Write a mongoDB query to aggregate and find max salary per group of various branches only for year 2020. Assume suitable data has been imported in step aii. | | **[3]**  **[3]** |
|  | (c) | Write a WordCount Solution using Map-Reduce supported by Hadoop framework and eloborate the process. | | **[6]** |
| **OR** | | | | |
| **Q.3** | (a)  (b) | Describe SQL, NoSQL and NewSQL.  Explain local/host OS file system vs distributed file system (HDFS) with solutions to following requirements:  (i) Create a HDFS directory named 'WordCloud' within home of 'hdfsspecial' HDFS user.  (ii) Copy sample files and directories available in home of 'hostspecial' local/host OS user to HDFS directory 'WordCloud'  (iii) List all the files and directories available in HDFS directory 'WordCloud' located in home of user 'hdfsspecial' HDFS user. | | **[3]**  **[3]** |
| (b) | Write a WordCount Solution using Map-Reduce supported by MongoDB framework and eloborate the process. | | **[6]** |
|  | | |  | |