| **Examination** | **: B.Tech Semester VII** | **Seat No.** | **:** |
| --- | --- | --- | --- |
| **Date** | **: 08-09-2022** | **Day** | **: Thursday** |
| **Time** | **: 9:00 to 10:15 a.m.** | **Max. Marks** | **: 36** |

| **INSTRUCTIONS:** | |
| --- | --- |
| 1. | Figures to the right indicate maximum marks for that question. |
| 2. | The symbols used carry their usual meanings. |
| 3. | Assume suitable data, if required & mention them clearly. |
| 4. | Draw neat sketches wherever necessary. |

Blooms Taxonomy levels : R-Remembering, U- Understanding, A-Applying, N-Analyzing, E- Evaluating, C-Creating

| **Q.1** |  | **Do as directed.** | | **[12]** |
| --- | --- | --- | --- | --- |
| **CO5** | **A** | (a) | Illustrate similarity at configuration level while installing HBase with respect to Hadoop. | [2] |
| **CO5** | **N** | (b) | Is HBase classified as a type of columnar database? Support your case with advantages/disadvantages to boost/weaken data analytics. | [2] |
| **CO5** | **R** | (c) | Describe the purpose of the column family concept in general. | [2] |
| **CO6** | **U** | (d) | Discuss the benefits of using apache maven as a tool. | [2] |
| **CO6** | **C** | (e) | Derive the use of terasort for benchmarking hadoop. | [2] |
| **CO6** | **R** | (f) | What is test driven development (TDD)? | [1] |
| **CO6** | **R** | (g) | Describe the mentioned terms with respect to HBase: Zookeeper, regionserver. | [1] |
|  |  |  |  |  |
| **Q.2** |  | Attempt ***Any TWO*** from the following: | | **[12]** |
| **CO4** | **E** | (a) | Summarize Cassandra consistency levels and describe consistency in general with respect to NoSQL. | [6] |
| **CO4** | **E** | (b) | Convince that Cassandra’s features like Hinted handoff, replication factor and gossip-failure detection makes it robust. | [6] |
| **CO4** | **E** | (c) | Summarize usage of CQLSH with 3-4 demo commands. | [6] |
|  |  |  |  |  |
| **Q.3** |  | Attempt the following: | | **[12]** |
| **CO6** | **N** | (a) | Analyze the similarity and differences in the working of Map Reduce programming framework prototype with respect to Apache Hadoop and mongoDB. | [6] |
| **CO2** | **C** | (b) | Generate mongoDB Query Language (MQL) statements to achieve following requirements (assume appropriate required information and mention it clearly):   * create and utilize a collection within your db and insert few varied documents * delete selected documents by default primary key * upsert sample document/s as inactive via a field(status) and value(inactive) | [6] |
|  |  |  |  |  |
|  |  | **OR** | |  |
|  |  |  | |  |
| **Q.3** |  | Attempt the following: | | **[12]** |
| **CO1** | **N** | (a) | Analyze and explain with example, mongoDB data modeling normalized and unnormalized way of document/s handling. | [6] |
| **CO2** | **C** | (b) | Build a solution using mongoDB to perform statistics level analytics (i.e. aggregation) for following requirements:   * import a csv via command line * group by attribute like department and display average salary department level * Display document/s is either not having a salary field at all or NULL. | [6] |