U.S.N

P.E.S. College of Engineering, Mandya - 571 401
(An Autonomous Institution affiliated to VTU, Belagavi)

Second Semester, M. Tech - Computer Science and Engineering (MCSE)
Semester End Examination; June - 2017
Managing Big Data

Time: 3 hrs   Max. Marks: 100

Note: Answer **FIVE** full questions, selecting **ONE** full question from each unit.

UNIT - I

1 a.	Apply Big data analytic framework to real time fraud detection problems. Also identify the potential data sources.	10
b.	Analyze the operations of Distributed file system and RDBMS.	10
2 a.	Describe the impact of Big Data on advertising industry.	10
b.	Explain Big Data and Hadoop open source technologies.	10

UNIT - II

3 a.	Construct a two level aggregate structure for customer information in a column-family structure.	10
b.	Explain the data models relationships and data base types.	10
4 a.	Illustrate with an example, Graph Database structure and relationship.	10
b.	Compare sharding and replication techniques of distribution of data.	10

UNIT - III

5 a.	Differentiate between Hadoop streaming and Hadoop Pipes.	10
b.	Discuss in detail the design issues of Hadoop distributed file system.	10
6 a.	List the different failures and their effects that need to be considered for running Map-Reduce program on YARN.	10
b.	Explain Hadoop primitives for data I/O.	10

UNIT - IV

7 a.	Explain the process involved in decomposing a problem into Map-Reduce Jobs.	10
b.	Explain the steps involved in Mapper testing and Reducers testing.	10
8 a.	Give the systematic way of Hadoop running a Map-Reduce job using YARN.	10
b.	Explain job scheduling for multi user Hadoop and Map-Reduce system.	10

UNIT - V

9 a.	Explain H Base and their data model and implementations.	10
b.	Illustrate with an example, the Cassandra data model.	10
10 a.	Explain in detail pig scripting platform and pit data model.	10
b.	List the primitive data types of Apache Hive and Hive QL data manipulation.	10

∗ ∗ ∗ ∗