Tribhuvan University
Institute of Science and Technology
2079
✡

Master Level / Second Year /Third Semester/ Science
Data Science (MDS 603)
(Techniques for Big Data)

Full Marks: 45
Pass Marks: 22.5
Time: 2 hours

Candidates are required to give their answers in their own words as for as practicable.
Attempt All Questions

Group A **[5 × 3 = 15]**

1. What are the different characteristics of big data?
2. How is functional programming language different than imperative programming language?
3. What are the core components of Hadoop Ecosystem?
4. What are structured, unstructured and semi structured data?
5. How Spark is different than Map Reduce? Compare Hive QL with SQL.

Group B **[5×6 = 30]**

6. Explain the scopes and applications of big data analytics in the different sectors with example.
7. What is map reduce? Explain its execution overview with reference to word frequency count example.
8. What are the different daemons running in a Hadoop cluster? Explain how these daemons work in Master/Slave Architecture.

OR

What are the different configuration modes to setup the Hadoop? Which setup mode is preferred for development and why?

9. Explain the limited taxonomy of NoSQL data base with example.

OR

What are the limitations of distributed databases like NoSQL? Explain it with reference to CAP Theorem.

10. Explain about resilient distributed dataset and data frames in spark.

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
Second assessment 2079

Subject: **Techniques for Big Data**
Course No: **MDS 603**
Level: **MDS/II Year/III Semester**

Full Marks: **45**
Pass Marks: **22.5**
Time: **2hrs**

Candidates are required to give their answer in their own words as far as practicable.

Attempt ALL questions

Group A [5x3=15]

1. What is NoSQL? How is it different than SQL?
2. Compare and Contrast HBase with HDFS.
3. What is Spark? How is it different than MapReduce?
4. Compare Pig Latin with SQL.
5. Differentiate between Hive QL and SQL

Group B [5x6=30]

6. Explain the CAP Theorem with example.

OR

What are the different types of NoSQL Database? Explain with example.

7. Can Mongo DB be considered as an alternate of RDBMS? Compare it against SQL with example.
8. How could SQL queries be executed in Spark. Explain with example.
9. What are the different execution modes in Apache Pig? Explain.
10. Explain about Hive Shell, Hive Services and Hive Metastore.

OR

Explain the user defined functions in Apache Pig and Apache Hive.

***