

04/06/2025 BE CSE-AIML SEM-VII C-SCHEME BDA QP CODE: 10080908

Time: 03 Hours

Marks: 80

Note: 1. Question 1 is compulsory

2. Answer any three out of the remaining five questions.

3. Assume any suitable data wherever required and justify the same.

- Q.1 (a) Compare traditional data and big data. 05
- (b) What are the advantages and limitations of Hadoop 05
- (c) Differentiate between SQL vs NoSQL 05
- (d) List and explain Distance measures for Big Data 05
- Q.2 (a) Draw Hadoop Ecosystem and briefly explain its components 10
- (b) Write the functions of the components and execution steps in Map Reduce 10
- Q.3 (a) Explain Selection and Projection algebraic operation using MapReduce. 10
- (b) Explain Key-value store and Document Store NoSQL architectural pattern with example. 10
- Q.4 (a) Draw a neat sketch, explain the architecture of the data-stream management system 10
- (b) Explain DGIM algorithm for counting ones in a stream with example 10
- Q.5 (a) Explain Page rank using Map reduce, also explain spider traps and dead ends 10
- (b) Explain Movie recommendation using Content -based filtering. 10
- Q.6 **Write short notes on any two (any 2)** 20
- (a) Bloom Filter with analysis
- (b) Cure Algorithm
- (c) Clustering of Social-Network Graphs.
- (d) Four ways that NoSQL systems handle big data problems.
