

**SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR**  
(AUTONOMOUS)

**MCA II Year II Semester Regular & Supplementary Examinations April-2026**  
**BIG DATA ANALYTICS**

Time: 3 Hours

Max. Marks: 60

(Answer all Five Units 5 x 12 = 60 Marks)

**UNIT-I**

1 Explain about analytic processes and tools in big data. CO1 L2 12 M

OR

2 a What is bootstrapping? outline its importance. CO6 L2 6 M

b Clearly explain about modern data analytic tools. CO1 L2 6 M

**UNIT-II**

3 a Explain the different applications of data streams in detail. CO5 L2 6 M

b What are filters in Big Data? Explain Bloom Filter with example CO1 L1 6 M

OR

4 Explain the following. i). FM algorithm and its application CO2 L2 12 M

ii). AMS algorithm and its applications

**UNIT-III**

5 a How Hadoop streaming is suited with text processing? Explain. CO4 L1 6 M

b How failures are handled in Map Reduce Applications? CO3 L1 6 M

OR

6 a Clearly explain how map reduce jobs run on YARN. CO4 L2 6 M

b Discuss the various types of MapReduce & its formats. CO3 L2 6 M

**UNIT-IV**

7 a What is Cluster? Explain the setting up a Hadoop cluster. CO3 L1 6 M

b What are the additional configuration properties to set for HDFS? CO3 L1 6 M

OR

8 a Describe the Hadoop specification followed by facebook. CO1 L1 6 M

b What are the security issues that may arise in the hadoop environment? Explain. CO1 L2 6 M

**UNIT-V**

9 a Describe in detail HIVE services. CO5 L1 6 M

b Explain in detail IBM infosphere Big insights and Streams. CO6 L2 6 M

OR

10 a What is HiveQL? Explain its features. CO5 L2 6 M

b Explain the features of HBase in a brief manner. CO5 L2 6 M

\*\*\* END \*\*\*