

Reg. No:

--	--	--	--	--	--	--	--	--

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)**MCA II Year II Semester Regular Examinations October-2020**
BIG DATA ANALYTICS

Time: 3 hours

Max. Marks: 60

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

UNIT-I

- 1 a Describe the steps involved in support vector based inference methodology. **6M**
b Write a short note on Statistical Inference. **6M**

OR

- 2 a Define and explain the different inferences in big data analytics. **6M**
b Describe the bootstrapping and its importance. **6M**

UNIT-II

- 3 a Explain the different applications of data streams in detail. **6M**
b Explain how to count ones in a window using DGIM algorithm. **6M**

OR

- 4 a What is Real Time Analytics? Discuss their technologies in detail. **6M**
b Explain the three categories of Prediction methodologies. **6M**

UNIT-III

- 5 a What is Hadoop? Explain its components. **6M**
b Explain the failures in Map reduce. **6M**

OR

- 6 a How Hadoop streaming is suited with text processing explain. **4M**
b Write in detail the concept of developing the Map Reduce Application. **8M**

UNIT-IV

- 7 a What is Cluster? Explain the setting up a Hadoop cluster. **8M**
b What are control scripts? Explain the start.dfs.sh script, Start.mapred.sh. **4M**

OR

- 8 a Discuss administering Hadoop with its checking point process diagram **6M**
b Explain Data node directory structure **6M**

UNIT-V

- 9 a Explain Grouping, Join, CoGroup, Cross & Group in data. **6M**
b Explain the process of installing HIVE & features of HIVE. **6M**

OR

- 10 a Give a detail note on Interaction techniques with its applications. **6M**
b Discuss the visual data analysis techniques in detail. **6M**

*** END ***