Reg. No: ⬚⬚⬚⬚⬚⬚⬚⬚⬚⬚

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)
MCA II Year II Semester Supplementary Examinations March-2021
BIG DATA ANALYTICS

Time: 3 hours　　　　　　　　　　　　　　　　　　　　　　　　Max. Marks: 60

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

UNIT-I

1	a	Discuss the following in detail about conventional challenges in big data.	6M
	b	Describe the prediction error and regression techniques.	6M

OR

2	a	Describe any five characteristics of Big Data.	6M
	b	Define Arcing classifier, Bagging predictors and explain in detail.	6M

UNIT-II

3	a	What are filters in Big Data? Explain Bloom Filter with example.	6M
	b	Explain about FM algorithm and its application	6M

OR

4	a	Explain the three categories of Prediction methodologies.	6M
	b	Write a short note on finding most popular elements using decaying window.	6M

UNIT-III

5	a	Explain the map reduce data flow with single reduce and multiple reduce.	6M
	b	How map reduce job works with classic java stream.	6M

OR

6	a	Explain how map reduce jobs run on YARN.	4M
	b	Discuss the various types of map reduce & its formats.	8M

UNIT-IV

7	a	What are the different types of Hadoop configuration files? Discuss.	8M
	b	What is benchmarking how it works in Hadoop.	4M

OR

8	a	What are the Important Hadoop daemon properties?	6M
	b	How does security is done in Hadoop.Justify.	6M

UNIT-V

9	What is Zookeeper? Explain its features with applications.	12M

OR

10	a	Explain in detail IBM infosphere Big insights and Streams.	6M
	b	What is HiveQL? Explain its features.	6M

*** END ***