Sl. No.

No. of Printed Pages : **2**

21CSC3C10L

M.Sc. III Semester Degree Examination, April/May - 2023

COMPUTER SCIENCE

Data Analytics

Time : 3 Hours Maximum Marks : 70

Instruction : *Answer* ***any five*** *questions (Question* ***No. 1*** *is* ***Compulsory****)*

1. (a) Define Big Data. Discuss on Big Data Architecture. **7**

(b) For the given data set, construct a predictive model to classify someone as Good or Bad. **7**

1	Good/Bad	Line
2	Good	You are awesome
3	Bad	You look ugly
4	Good	This is so nice
5	Good	I love this tea

2. (a) Enlist the causes for missing values in data set. **7**

(b) Construct box plot diagram by considering your own data for multivariate analysis. **7**

3. (a) Discuss K-Means technique. **7**

(b) Write an algorithm for Association rule generation from a frequent Item set. **7**

4. (a) Demonstrate model validation with an example. **7**

(b) What are the parameters influenced by our requirements for the selection of Technique and Model Selection ? **7**

5. (a) Explain binary classification process with an example. **7**

(b) Discuss on Case-based reasoning with a neat diagram. **7**

6. (a) What are the steps taken to resolve missing values, Redundant data, Inconsistent data, Noisy data ? **7**

(b) List out the advantages and disadvantages of DBSCAN, Agglomerative hierarchical clustering. **7**

P.T.O.

7. (a) Why should we use different data to train and to test a model ? **7**

(b) Explain the classification process in Analytics. **7**

8. (a) Explain the measures of classification. **7**

(b) Write a note on Predictive Analytics. **7**

- o 0 o -