Total No. of Questions : 8]

SEAT No. :

P7290

[Total No. of Pages : 2

[6181]-393

B.E. (Electronics & Computer Engineering)
DATA SCIENCE & VISUALIZATION
(2019 Pattern) (Semester-VII) (410341)

Time : 2½ Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) ***Answer Q1 or Q2, Q3 or Q4, Q5 or Q6, Q7 or Q8.***
2) ***Figures to the right side indicate full marks.***
3) ***Assume suitable data,if necessary.***

Q1) a) What is data analysis? What are the issues in data analysis? What are the advantages of data analysis. **[6]**

b) What is Apriori Algorithm? Explain components of Aprior algorithm. Write the steps in Apriori Algorithm. **[6]**

c) What is K-Means algorithm? Explain in detail working of K-Means algorithm, State the application where the is preferred. **[8]**

OR

Q2) a) Explain Inertia, Dunn Index and Silhouette Score evaluation metrics for clustering. **[6]**

b) What is linear regression? Describe advantages and disadvantages of linear regression. **[6]**

c) What is the Naive Bayes algorithm? Explain the Naive Bayes algorithm with Bayes theorem and suitable example. Write the various applications of Naive Bayes algorithms **[8]**

Q3) a) Explain the term Entropy, Information Gain and Gini Index with respect to Decision Tree. **[8]**

b) What is Random Forest algorithm? Write the important features of Random Forest algorithm. Differentiate between Decision Tree and Random Forest. **[8]**

OR

P.T.O.

Q4) a) With the help of diagram explain working of Decision Tree. Explain different terminologies used in Decision Tree. Explain the classification and regression Decision Trees. **[8]**

b) What is back propagation in neural networks? Why back propagation is required in neural networks? How does back propagation work in a neural network? Write advantages of using the back propagation in neural networks. **[8]**

Q5) a) What is a Pie Chart? What is the purpose of Pie Chart? Write advantages & disadvantages of a Pie Chart. **[8]**

b) What is Histogram plot? When to use Histogram plot? Write difference between Bar plot and Histogram plot. **[8]**

OR

Q6) a) Draw the box plot for the given set of data **[8]**

25, 28, 29, 29, 30, 34, 35, 35, 37, 38

b) What is a Data Dashboard? What are the benefits of using data dashboards? How to create an effective data dashboard? **[8]**

Q7) a) What is the Self-Organizing Maps algorithm? Draw architecture and explain working of Self-Organizing Maps algorithm. Write applications of Self-Organizing Maps algorithm. **[9]**

b) What is Principal Component Analysis? Explain various steps for Principal Component Analysis algorithm. Write advantages and disadvantages for Principal Component Analysis. **[9]**

OR

Q8) a) What is meant by clustering high-dimensional data? Explain the different challenges of clustering high-dimensional data. Explain different methods for clustering high-dimensional data. **[9]**

b) Explain Parallel Coordinates plot for data visualization in detail and compare the following:

i) Parallel Coordinate plots and Scatter plots.

ii) Parallel Coordinate plots and Bar plots **[9]**