CAD110

Total No. of Printed Pages: 2

BCA (Semester – VI)
EXAMINATION MAY - 2023
Data Science Concepts

[Time:2 Hours]　　　　　　　　　　　　　　　　　　　　　　　　[Max. Marks:60]

Instructions:　i) Question 1-6 are compulsory
　　　　　　　　　ii) Figures to the right indicate full marks.

Q1　　**A. Match the following:**　　　　　　　　　　　　　　　　　　5*1=5

Column A	Column B
i. Association rule mining	a. Wrapper method
ii. Forward selection	b. Uses data to make suggestions
iii. Strong learners	c. Assume all feature are independent and unrelated
iv. Prescriptive ML System	d. Aprori Algorithm
v. Naïve bayes Algorithm	e. High prediction accuracy

　　　B) Answer the following
　　　　i. What is Dimensionality Reduction?　　　　　　　　　　　5*1=5

　　　ii. What is Market Basket Analysis?

　　　iii) What is Data Science?

　　　iv. What is Data Analysis?

　　　v. What is Classification?

Q2　　A. List any 2 applications of Clustering.　　　　　　　　　　　2

　　　B. State any 3 features of machine learning.　　　　　　　　　3
　　　C. Explain the K-Nearest Neighbour algorithm.　　　　　　　　5

Q3　　A. Give any two examples of data visualization.　　　　　　　2

　　　B. Explain the terms Support and Confidence using the dataset given below.　　3

　　　Customer 1: Apples, Oranges, Grapes

　　　Customer 2: Oranges, Apples, Gum

　　　Customer 3: Oranges, Bananas, Grapes

　　　Customer 4: Gum, Bananas, Bread

　　　Customer 5: Apples, Oranges, Bread

1

C. Explain Artificial Neural Network.   5

Q4   A. Give 2 points of difference between Descriptive Statistics and Inferential Statistics.   2

B. List any 3 disadvantages of Decision Trees.   3

C. Give 5 points of difference between K-Means and K-Medoid clustering algorithm.   5

Q5   A. List any 2 steps involved in data cleaning process.   2

B. State the 3 categories of Logistic Regression.   3

C. Give 5 points of difference between Supervised Learning and Unsupervised Learning.   5

Q6   A. How is the final output calculated in Random Forest Algorithm?   2

B. Define Feature Selection and list any 2 advantages.   3

C. Explain the Ensemble Learning Techniques used in Machine Learning.   5