Vidya Vikas Mandal's
Shree Damodar College of Commerce & Economics, Margao-Goa
TY BCA Semester-VI Semester End Examination, June 2022
Data Science Concepts (CAD-110)

Duration: 2 Hours   Max Marks: 60

Instructions: *i) All Questions are compulsory*
ii) Figures to the right indicate full marks

Q. 1. A) State TRUE or FALSE   (5x1=05)

i) Artificial Intelligence (AI) is a concept where in Computers or systems mimic human processes or decision-making abilities.

ii) Final results of Inferential Statistics are shown in the form of charts, tables and graphs.

iii) Supervised learning is a machine learning approach that is defined by its use of labeled data sets.

iv) Linear Discriminant Analysis (LDA) is a dimensionality reduction technique.

v) Boosting aims to decrease variance and not the bias.

Q.1. B) Define the following terms.   (5x1=05)

i) Data Warehouse
ii) Over fitting
iii) Data Wrangling.
iv) Unsupervised Learning
v) Binomial Logistic Regression.

Q.2 Answer the following:

(a) List the different stages in the Life cycle of Data Science.   (2)
(b) Compare and contrast Descriptive Statistics v/s Inferential Statistics.   (3)
(c) Explain any five applications of Data Visualization.   (5)

Q.3. Answer the following:

(a) List any two characteristics of filter method.   (2)
(b) Discuss the types of Logistic Regression   (3)
(c) Explain in detail the data cleaning steps involved in Machine learning.   (5)

Q.4. Answer the following:

(a) Define Data Visualization and list any two data visualization techniques.   (2)
(b) Differentiate between Supervised learning and Unsupervised learning.   (3)
(c) Explain Feature Selection techniques in detail.   (5)

Q.5. Answer the following:

(a) Give a point of difference between Simple Linear Regression and Multiple Linear Regression.   (2)
(b) Describe K-Nearest Neighbours (KNN) Algorithm. Write down the steps of the algorithm.   (3)

(c) For the below given data set apply Naive Bayes Algorithm and predict the outcome for (5)
   Car = { Red, Domestic, SUV}

Colour	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

(Note: Predict whether the given car is stolen or not)

Q.6. Answer the following.
   (a) List any two features of Random Forest Algorithm. (2)
   (b) Write the steps involved in Boosting Algorithm. (3)
   (c) Explain Decision Tree. Illustrate it with the help of an example. (5)

END