National Institute of Technology
Tiruchirappalli, Tamil Nadu – 620 015

Machine Learning Techniques and Practices – CT1 **Date: 24.02.2022**

Duration: 1 Hr **Time:** 10:00 – 11:00 AM

Total Marks: 20

1. If a dataset has the target labels associated with all the samples and the machine learning algorithm also considers all these labels to take a decision, then which category does this learning algorithm belong to? ***(1 M)***

   (a) Supervised **(b)** Unsupervised **(c)** Semi-supervised **(d)** Reinforcement

2. (i) Draw Box-and-Whisker Plot for the values: **4, 7, 9, 8, 12, 80, 15** ***(4 + 2 = 6 M)***

   (ii) Discuss about the spread of data in the above plot using Q1, Q2, Q3 and IQR values.

3. Match the following: ***(2 M)***

A. Feature Binning	(i) Creates separate columns for each unique value that is present in the categorical feature
B. Feature Engineering	(ii) Converts the "n" unique values in categorical features to values between 0 and n-1
C. Label Encoding	(iii) Should be utilized when one wants to replace an existing feature with more meaningful additional features
D. One Hot Encoding	(iv) Should be applied on features that has large number of unique values

   (a) A → (ii); B → (iii); C → (iv); D → (i) **(b)** A → (iii); B → (i); C → (iv); D → (ii)

   (c) A → (i); B → (iv); C → (ii); D → (iii) **(d)** A → (iv); B → (iii); C → (ii); D → (i)

4. (i) Write the names of various binning methods. ***(2 + 3 = 5 M)***

   (ii) Consider the following data and apply any two binning methods *[**Hint:** Bin Size = 3]*

   15, 21, 45, 6, 11, 17, 45, 19, 12, 4, 9, 5

5. (i) Does the feature "Age" in below dataset require Feature Scaling? State the reason.

   (ii) How will you identify whether the features "Speed" and "Acceleration" are related to each other or not (State the reason). Also, draw a rough graph using these two features and identify what

Dr. R. Bala Krishnan, Assistant Professor, Department of CSE

will be the range/value of Pearson Correlation Coefficient. What will be your final decision – That is, under what condition you will be dropping a feature? ***(2 + 4 = 6 M)***

Sl. No.	Name	Age	Speed (in km/hr)	Acceleration (in %)	Rotation of Tyre (in km/hr)	Direction of Motion	Clear Vision (in %)	Class
1.	Bala	1000	100	60	100	Forward	80	Racing
2.	Karthick	25	120	70	120	Forward	70	Racing
3.	Sundar	35	140	75	140	Forward	60	Racing
4.	Rajesh	30	160	85	160	Forward	50	Racing
5.	Kumar	20	180	100	180	Forward	40	Racing

---------- END ----------