Vidyalankar Institute of Technology
Semester 7 – CMPN - Mid Semester Assessment – 1

Date: 05/08/2024	Machine Learning	30 Marks /1 hour

1	**Solve any two (5 marks each)**		**CO**
	A	Justify why Linear Regression is not used for performing classification.	CO1
	B	Elaborate on: (i) Karl Pearsons Coefficient of Co relation. (ii) R square method.	CO2
	C	Discuss the importance of the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) in evaluating binary classifiers. How do these metrics help in comparing models?	CO2
2	**Solve any two (5 marks each)**		
	A	Outline the steps involved in developing a Machine Learning application for predicting housing prices.	CO1
	B	Justify the significance of the F1 score in imbalanced classification problems. How does it address the limitations of using precision or recall alone? Provide a scenario where the F1 score might still be misleading.	CO2
	C	Justify the bias-variance trade-off in the context of model complexity. How does this trade-off influence the choice of model in a real-world ML application? Provide examples to support your explanation.	CO2
3	**Solve anyone (10 marks each)**		
	A	For the dataset given below, construct a decision tree using Gini Index, and determine which attribute is a root attribute.	CO2

Sr. No.	Weather	Parent	Money	Decision
1	Sunny	Yes	Rich	Cinema
2	Sunny	No	Rich	Tennis
3	Windy	Yes	Rich	Cinema
4	Rainy	Yes	Poor	Cinema
5	Rainy	No	Rich	Stay In
6	Rainy	Yes	Poor	Cinema
7	Windy	No	Poor	Cinema
8	Windy	No	Rich	Shopping
9	Windy	Yes	Rich	Cinema
10	Sunny	No	Rich	Tennis

	B	Justify the significance of Cross Validation. List the type and discuss k-fold cross validation in detail? CO3

CO1	Gain knowledge about basic concepts of Machine Learning and understand the difference between supervised and unsupervised techniques
CO2	To select, apply and evaluate an appropriate machine learning model for the given
CO3	Ability to understand regression techniques.