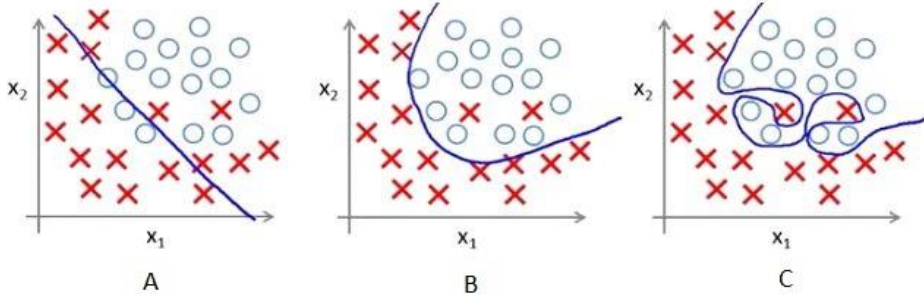Roll No_____   Name_____   Group_____

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology

UML501 (Machine Learning) Quiz -I (September 12, 2022)

Time: 12 Minutes   **MM: 10**

Note: Overwritten and multiple answers will not be evaluated. Use of pencil is not allowed. Answer to be written in the **Answer** column only. There will be deduction 0.25 marks for wrong answers.

		Question	Answer
Q1		What may happen when you add very large penalty in the loss function of Ridge regression? A. Some of the coefficient will become zero B. **Some of the coefficient will approach zero but not zero** C. Both A and B, depending on the situation D. None of these	
Q2		Because of low bias and high variance, we get _____ in the model A. High error B. Perfectly fitting C. Underfitting D. **Over fitting**	
Q3		Which is best suited to test linear relationship of dependent and independent continuous variables? A. Bar Chart B. Histogram C. **Scatter Plot** D. None of the mentioned	
Q4		User rating given to a movie on a scale of 1-10, can be considered as an attribute of type? A. Nominal B. **Ordinal** C. Interval D. Ratio	
Q5		Suppose you have m=14 training examples with n=3 features (excluding the additional all-ones feature for the intercept term, which you should add). The normal equation is $\theta = (X^TX)^{-1}X^Ty$. For the given values of m and n, what are the dimensions of θ, X, and y in this equation? A. X is 14×3, y is 14×1, θ is 3×3 B. X is 14×4, y is 14×4, θ is 4×4 C. **X is 14×4, y is 14×1, θ is 4×1** D. X is 14×3, y is 14×1, θ is 3×1	
Q6		Following are the three scatter plots (A, B, C) and hand drawn decision boundaries for learning algorithms.	

A   B   C

Which of the following above figure shows that the decision boundary is overfitting the training data?
    a) A
    b) B
    c) C
    d) None of these

Q7	The relationship between number of beers consumed (x) and blood alcohol content (y) was studied in 16 male college students by using least squares regression. The following regression equation was obtained from this study: y = -0.0127 + 0.0180 x The above equation implies that: A. Each beer consumed increases blood alcohol by 1.27% B. On average it takes 1.8 beers to increase blood alcohol content by 1% C. Each beer consumed increases blood alcohol by an average amount of 1.8% D. each beer consumed increases blood alcohol by exactly 0.018	
Q8	The correlation between the number of years an employee has worked for a company and the salary of the employee is 0.75. What can be said about employee salary and years worked? A. There is no relationship between salary and years worked. B. Individuals that have worked for the company the longest have higher salaries. C. Individuals that have worked for the company the longest have lower salaries. D. The majority of employees have been with the company a long time.	

Q9	Suppose you have trained a polynomial model $y = w_0 + w_1 x + \ldots + w_M x^M$ via least-squares regression, and you find that it has a low training error, but a high testing error. Which of the following is likely to reduce the testing error?
	
	    A. Decreasing M and re-training.
	    B. Increasing M and re-training.
	    C. Dividing each of the learnt weights by 2.
	    D. All options are correct	

Q10	Which of the following statement is true about outliers in Linear regression?
	    A **Linear regression is sensitive to outliers**
	    B Linear regression is not sensitive to outliers
	    C Can't say
	    D None of these	