

- Q.6 Attempt any two:
- What is the importance of the F-test in a linear model? **5**
 - What is a residual in linear regression? How is it used in model evaluation? **5**
 - Explain confidence intervals on the regression coefficients. **5**

Total No. of Questions: 6

Total No. of Printed Pages:4

Enrollment No.....



Faculty of Science
End Sem Examination Dec-2023
BC3CO53 Machine Learning

Programme: B.Sc.

Branch/Specialisation: Computer
Science**Duration: 3 Hrs.****Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

- Q.1
- The real-world machine learning use case are- **1**
 - Digital assistants
 - Chatbots
 - Fraud detection
 - All of these
 - Identify the type of learning in which labelled training data is used- **1**
 - Supervised learning
 - Unsupervised learning
 - Reinforcement learning
 - None of these
 - What is linear regression? **1**
 - A technique used to model and analyze the relationship between two or more variables
 - A type of regression analysis that uses a linear function to model the relationship between a dependent variable and one or more independent variables
 - A statistical measure used to determine the strength of the relationship between two variables
 - All of these
 - What is a residual in the context of linear regression? **1**
 - The difference between the predicted value of the dependent variable and the actual value
 - The difference between the predicted value of the independent variable and the actual value
 - The difference between the predicted value of the dependent variable and the mean of the dependent variable
 - The difference between the predicted value of the independent variable and the mean of the independent variable

[2]

- v. Which of the following methods is used to select the best subset of predictor variables in multiple linear regression? **1**
 (a) Forward selection
 (b) Backward elimination
 (c) Stepwise regression
 (d) All of these
- vi. Which of the following methods do we use to find the best-fit line for data in linear regression? **1**
 (a) Least square error
 (b) Maximum likelihood
 (c) Logarithmic loss
 (d) Both (a) and (b)
- vii. What is multicollinearity? **1**
 (a) A condition in which two or more independent variables are highly correlated
 (b) A condition in which two or more dependent variables are highly correlated
 (c) A condition in which the independent variables are not correlated with the dependent variable
 (d) A condition in which the residuals are not normally distributed
- viii. What is the goal of linear regression? **1**
 (a) To predict an outcome variable based on one or more predictor variables
 (b) To classify data into different categories
 (c) To group similar data points together
 (d) To identify relationships between variables
- ix. What is the coefficient of determination? **1**
 (a) A measure of how well the model fits the data
 (b) A measure of the strength of the relationship between the dependent and independent variables
 (c) A measure of the strength of the relationship between the independent variables
 (d) A measure of the strength of the relationship between the dependent and independent variables, adjusted for the number of independent variables

[3]

- x. How is the coefficient of determination calculated? **1**
 (a) By dividing the sum of the squared residuals by the sum of the squared differences between the actual and predicted values of the dependent variable
 (b) By dividing the sum of the squared residuals by the sum of the squared differences between the actual and predicted values of the independent variable
 (c) By dividing the sum of the squared differences between the actual and predicted values of the dependent variable by the sum of the squared residuals
 (d) By dividing the sum of the squared differences between the actual and predicted values of the independent variable by the sum of the squared residuals
- Q.2 i. What is machine learning? How it is different from human learning? **2**
 ii. What are the different techniques possible in machine learning? **3**
 iii. What is difference between under-fitting and over-fitting in context of model performance? Explain with an example. **5**
- OR iv. What is accuracy, precision, recall and F1-score? Explain with an example. **5**
- Q.3 i. What is linear regression? How does it work? **2**
 ii. What is the difference between a parametric and non-parametric regression model? **8**
- OR iii. What is the primary difference between R squared and adjusted R squared? **8**
- Q.4 i. What is the difference between simple and multiple linear regression? **3**
 ii. Explain with an example the term maximum likelihood estimation. **7**
- OR iii. Explain the properties of the least squares estimators. **7**
- Q.5 i. How do you determine the best fit line for a multiple linear regression model? **4**
 ii. What are the basic assumptions required to validate the multiple regression model? **6**
- OR iii. What is multicollinearity and how does it affect linear regression analysis? **6**

Scheme of Marking

Machine Learning-BC3CO53(T)

Q.1	i)	The real-world machine learning use case are: d) All of the above (Answer)	1
	ii)	Identify the type of learning in which labelled training data is used. a) Supervised Learning (Answer)	1
	iii)	What is linear regression? d) All of the above (Answer)	1
	iv)	What is a residual in the context of linear regression? a) The difference between the predicted value of the dependent variable and the actual value (Answer)	1
	v)	Which of the following methods is used to select the best subset of predictor variables in multiple linear regression? d) All of the Above (Answer)	1
	vi)	Which of the following methods do we use to find the best-fit line for data in Linear Regression? a) Least Square Error (Answer)	1
	vii)	What is multicollinearity? a) A condition in which two or more independent variables are highly correlated (Answer)	1
	viii)	What is the goal of linear regression? a) To predict an outcome variable based on one or more predictor variables (Answer)	1
	ix)	What is the coefficient of determination? a) A measure of how well the model fits the data (Answer)	1
	x)	How is the coefficient of determination calculated? c) By dividing the sum of the squared differences between the actual and predicted values of the dependent variable by the sum of the squared residuals (Answer)	1
Q.2	i.	What is Machine Learning It is different from human learning	1 mark 1 mark
	ii.	What are the different techniques possible in Machine Learning? 3 points 1*3	3
	iii.	What is difference between Under-fitting and Over-fitting in context of model performance, explain with an example? 2 points + example (2.5+2.5)	5
OR	iv.	What is Accuracy, Precision, Recall and F1-Score, explain with an example?	5

Definition + example (1.25*4)

Q.3	i.	What is linear regression how does it work	1 mark 1 mark	2
	ii.	What is the difference between a parametric and non-parametric regression model? 4 points each (4*2)		8
OR	iii.	What is the primary difference between R squared and adjusted R squared? 4 points each (4*2)		8
Q.4	i.	What is the difference between simple and multiple linear regression? 3 points each (4*2)		3
	ii.	Explain with an example the term Maximum Likelihood Estimation. Explanation		7
OR	iii.	Explain the Properties of the Least Squares Estimators. Explanation		7
Q.5	i.	How do you determine the best fit line for a Multiple linear regression model? Explanation		4
	ii.	What are the basic assumptions required to validate the Multiple regression model? Explanation		6
OR	iii.	What is multicollinearity and how does it affect linear regression analysis? Explanation		6
Q.6		Attempt any two:		
	i.	What is the importance of the F-test in a linear model? Explanation		5
	ii.	What is a residual in linear regression and how is it used in model evaluation? Explanation		5
	iii.	Explain Confidence Intervals on the Regression Coefficients Explanation		5

[2]

[3]