Linear Regression Quiz

Total points 14/20

Email address *

skkumarlokam@gmail.com

Name *

Siva Lokam

✓ 1. Which of the following methods do we use to find the best fit line for data in Linear Regression? 1/1

- (●) Least Square Error ✓
- ( ) Maximum Likelihood
- ( ) Logarithmic Loss
- ( ) Both A and B

✓ 2. Which of the following evaluation metrics can be used to evaluate a model while modeling a continuous output variable? 1/1

- ( ) Accuracy
- ( ) AUC-ROC
- (●) Mean Squared Error ✓
- ( ) Logloss

✗ 3. Ridge Regularization can be used for variable selection in Linear Regression. 0/1

- ○ True
- ◉ False ✗

Correct answer

- ◉ True

Feedback

Ridge regression is a regularization method intended to penalize features with too high weights

✗ 4. What linear regression algorithm can you use if you have a training set with millions of features? 0/1

- ○ Normal Equation
- ◉ Stochastic Gradient Descent ✗
- ○ Batch Gradient Descent
- ○ None

Correct answer

- ◉ Batch Gradient Descent

✓ 5. Why would you want to use Ridge Regression over Normal Regression? 1/1

- ○ It generalizes the data
- ○ Underfits the data
- ○ Overfits the data
- ◉ Regularizes the data ✓

✓ 6. Is it good to stop Mini-batch Gradient Descent immediately when the validation error goes up? 1/1

- ○ Yes
- ◉ No ✓

✓ 7. In Polynomial Regression, on plotting learning curve we see a huge gap between training error and validation error(this being higher), this is so because of ?? 1/1

- ○ Model Underfitting
- ◉ Model Overfitting ✓
- ○ Generalization of Model
- ○ None

✓ 8. What can you expect an overfit model to have among these? 1/1

- (●) Low Bias, High Variance ✓
- ( ) High Bias, High Variance
- ( ) Low Bias, Low Variance
- ( ) High Bias, Low Variance

✗ 9. What will happen when you fit degree 3 polynomial in linear regression? 0/1

- (●) There are high chances that degree 3 polynomial will under fit the data ✗
- ( ) There are high chances that degree 2 polynomial will over fit the data
- ( ) Can't say
- ( ) None of these

Correct answer

- (●) Can't say

Feedback

Depends on the complexity of data

✓ 10. Suppose you plotted a scatter plot between the residuals and predicted values in linear regression and you found that there is a relationship between them. Which of the following conclusion do you make about this situation? 1/1

- ○ Since there is a relationship the model is good
- ◉ Since there is a relationship the model is not good ✓
- ○ None of these
- ○ Can't say

✗ 11. Can you help me out in knowing the assumptions before building any Linear Regression Model? 0/1

- ○ Linear relationship, Singlevariate normality, Multicollinearity among features, Homoscedasticity
- ○ Linear relationship, Bivariate normality, No Multicollinearity among features, Homoscedasticity
- ○ Linear relationship, Singlevariate normality, No or little Multicollinearity among features, No autocorrelation among features, Heteroscedasticity
- ◉ Linear relationship, Multivariate normality, No or little Multicollinearity among features, No autocorrelation among features, Homoscedasticity ✗

Correct answer

- ◉ Linear relationship, Singlevariate normality, No or little Multicollinearity among features, No autocorrelation among features, Heteroscedasticity

✓ 12. Which of the following statement is true about outliers in Linear regression? 1/1

- ○ Linear regression is not sensitive to outliers
- ◉ Linear regression is sensitive to outliers ✓
- ○ None of these
- ○ Can't say

✗ 13. We can also compute the coefficient of linear regression with the help of an analytical method called “Normal Equation”. Which of the following is/are true about Normal Equation? 0/1

- ☑ There is no need to iterate ✓
- ☑ We don’t have to choose the learning rate ✓
- ☐ It becomes slow when number of features is very large

Correct answer

- ☑ There is no need to iterate
- ☑ We don’t have to choose the learning rate
- ☑ It becomes slow when number of features is very large

Feedback

Normal equation method uses matrices and hence:
There's no need to iterate
No learning rate since there's no optimizing function like gradient descent
Computation becomes very complex as the number of features increases due to matrices

✓ 14. Which among the following is true about Residuals? 1/1

- ( ) Higher is better
- (•) Lower is better ✓
- ( ) Can't say
- ( ) None of these

✓ 15. Can all data points lying above IQR > 3 be treated as outliers? 1/1

- ( ) Yes
- ( ) No
- (•) Maybe ✓
- ( ) None of these

✓ 16. When the correlation coefficient , r , is close to one: 1/1

- ( ) there is no relationship between the two variables
- (•) there is a strong linear relationship between the two variables ✓
- ( ) it is impossible to tell if there is a relationship between the two variables
- ( ) the slope of regression line will be close to one

Given the following data pairs (x, y) :- (1, 1.24), (2, 5.23), (3, 7.24), (4, 7.60), (5, 9.97), (6, 14.31), (7, 13.99), (8, 14.88), (9, 18.04), (10, 20.70) Answer the questions below

✓ 17. Find the regression equation. 1/1

- ○ y = 0.490 x - 0.053
- ○ y = 2.04 x
- ◉ y = 1.98 x + 0.436 ✓
- ○ y = 0.49 x

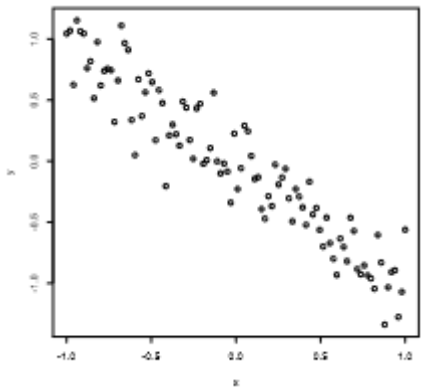
✗ 18. Calculate the correlation coefficient. 0/1

- ○ r = 0.490
- ○ r = 0.985
- ◉ r = 0.971 ✗
- ○ r = 0.240

Correct answer

- ◉ r = 0.985

Below is the scatter plot showing relationship between 'x' independent variable and 'y' dependent variable



✓ 19. The data in the scatterplot above would have a correlation coefficient that is close to: 1/1

- (●) -1.0 ✓
- ( ) -0.5
- ( ) +1.0
- ( ) +0.5

✓ 20. The intercept in linear regression represents: 1/1

- ( ) the strength of the relationship between x and y
- ( ) the expected mean x value when y is zero
- (●) the expected mean y value when x is zero ✓
- ( ) a population parameter