

MACHINE LEARNING WORKSHEET_SET_1

Ques1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

ANS B (maximum likelihood)

Ques2. Which of the following statement is true about outliers in linear regression?

ANS A (linear regression is sensitive to outliers)

Ques3. A line falls from left to right if a slope is _____?

ANS D (Zero)

Ques4. Which of the following will have symmetric relation between dependent variable and independent variable?

ANS A (Regression)

Ques5. Which of the following is the reason for over fitting condition?

ANS C (low bias and high variance)

Ques6. If output involves label then that model is called as:

ANS B (predictive modal)

Ques7. Lasso and Ridge regression techniques belong to _____?

ANS D (Regularization)

Ques8. To overcome with imbalance dataset which technique can be used?

ANS D (SMOTE)

Ques9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

ANS A (TPR and FPR)

Ques10. . In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

ANS A (TRUE)

Ques11. Pick the feature extraction from below:

ANS A (construction bag of words from a email)

Ques12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

ANS A and B

Ques13. Explain the term regularization?

ANS Regularization in machine learning terms is to make things acceptable or regular. The process involves the shrinking of data coefficients to tend to zero values. In other words, the process of regularization of the regularization methods in machine learning will discourage overfitting the model, which then learns to be more flexible in a complex environment.

REGULARIZATION TECHNIQUES

- 1) LASSO REGRESSION**
- 2) RIDGE REGRESSION**

Ques14. Which particular algorithms are used for regularization?

ANS REGULARIZATION ALGORITHM:-

- 1) LASSO REGRESSION**
- 2) RIDGE REGRESSION**

RIDGE REGRESSION:-

Also known as Ridge Regression, it modifies the over-fitted or under fitted models by adding the penalty equivalent to the sum of the squares of the magnitude of coefficients.

This means that the mathematical function representing our machine learning model is minimized and coefficients are calculated. The magnitude of coefficients is squared and added. Ridge Regression performs regularization by shrinking the coefficients present. The function depicted below shows the cost function of ridge regression.

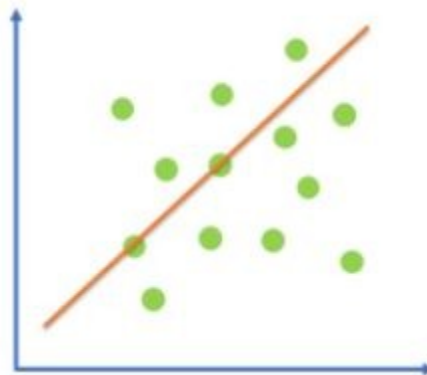
$$\text{Cost function} = \text{Loss} + \lambda \times \sum \|w\|^2$$

Here,

Loss = Sum of the squared residuals

λ = Penalty for the errors

W = slope of the curve/ line



LASSO REGRESSION

It modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the absolute values of coefficients.

Lasso regression also performs coefficient minimization, but instead of squaring the magnitudes of the coefficients, it takes the true values of coefficients. This means that the coefficient sum can also be 0, because of the presence of negative coefficients. Consider the cost function for Lasso regression

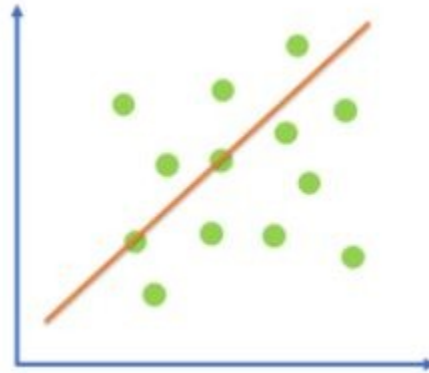
Cost function = Loss + $\lambda \times \sum \|w\|$

Here,

Loss = Sum of the squared residuals

λ = Penalty for the errors

w = slope of the curve/ line



Ques15. Explain the term error present in linear regression equation?

ANS The linear regression model contains an error term that is represented by ϵ . The error term is used to account for the variability in y that cannot be explained by the linear relationship between x and y. If ϵ were not present, that would mean that knowing x would provide enough information to determine the value of y.