# WORKSHEET

MACHINE LEARNING – WORKSHEET 11 (LINEAR REGRESSION)

## In Q1 to Q8, only one option is correct, Choose the correct option:

1. What happens to R2 measure if we add a new feature?
   1. remains same B) always increases

C) may or may not increase D) always decreases

**Answer:** C) may or may not increase

1. The correct relationship between SST, SSR and SSE is given by:
   1. SSR = SST + SSE B) SST = SSR + SSE

C) SSE = SSR – SST D) None of the above

**Answer:** B) SST= SSR + SSE

1. Residuals in regression analysis can be defined as:
   1. difference between the actual value and the predicted value.
   2. difference between the actual value and the mean of predicted value.
   3. difference between the actual value and mean of dependent variable.
   4. None of the above.

**Answer:** A) difference between the actual value and the predicted value

1. In a simple linear regression model, if we change the input variable by 1 unit, then how much output variable will change?
   1. By 1 B) No Change

C) By its slope D) None of the above

**Answer:** C) By its slope

1. If the coefficient of determination is equal to 1, then correlation coefficient:
   1. must also be equal to 1 B) can be either -1 or 1

C) can be any value between -1 to 1 D) must be -1

**Answer:** C) can be any value between -1 to 1

1. Which of the following plot is best suited for the linear relationship of continuous variables?
   1. Scatter plot B) Histograms

C) Pie charts D) All of the above

**Answer:** A) Scatter plot

1. The ratio of MSR/MSE produces:
   1. t-statistics B) f-statistics

C) z-statistics D) None of the above.

**Answer:** B) f-statistics

1. Which of the following regularizations uses only L2 normalization for its penalty parameter?
   1. Lasso B) Elastic Nets

C) Ridge D) All of the above

**Answer:** C) Ridge

## In Q9 to Q11, more than one options are correct, Choose all the correct options:

1. Which of the following statement/s are true for best fitted line?
   1. It shows the causal relationship between dependent and independent variables
   2. It shows the positive or negative relation between dependent and independent variables
   3. It always goes through origin
   4. It is a straight line that is the best approximation of the given data sets

Answer:

1. Regularizations helps in:
   1. Reducing the training time B) Generalizing the test set

C) Automatic feature selection D) Grouping the data

**Answer:** A) Reducing the training time

1. Linear regression can be implemented through:
   1. Normal Equation B) Singular Value Decomposition

C) Parity checks D) nodes

**Answer:** B) Singular Value Decomposition

## Q12 to Q15 are subjective answer type questions, Answer them briefly.

1. Explain R2 and adjusted R2 metrics?

**Answer:** R2 measures the proportion of the variation in your dependent variable explained by all of your independent variables in the model. It assumes that every independent variable in the model helps to explain variation in the dependent variable.

Adjusted R-Squared can be calculated mathematically in terms of sum of squares. The only difference between R-square and Adjusted R-square equation is degree of freedom.

1. Explain the cost function of linear regression?

**Answer:** A cost function in layman’s term is a term used to represent how far the result is from the real value. A cost function in linear regression is the quantified value of error that is how deviated it is from the actual value.

1. Differentiate SSE, SSR and SST.

Answer:

SSE: The error is the difference between the observed value and the predicted value.

SSR: It is the sum of the differences between the predicted value and the mean of the dependent variable.

SST: It is the squared differences between the observed dependent variable and its mean.

1. What are the various evaluation metrics for linear regression?

Answer: The various metrics used to evaluate the results of the prediction are :

1. Mean Squared Error(MSE)
2. Root-Mean-Squared-Error(RMSE).
3. Mean-Absolute-Error(MAE).
4. R² or Coefficient of Determination.
5. Adjusted R²