Training Day-14 Report:

R-squared in Regression Analysis in Machine Learning:-

The most important thing we do after making any model is evaluating the model. We have different evaluation matrices for evaluating the model. However, the choice of evaluation matrix to use for evaluating the model depends upon the type of problem we are solving whether it’s a regression, classification, or any other type of problem. In this article, we will explain R-Square for regression analysis problems.

What is R-Squared?

R-squared is a statistical measure that represents the goodness of fit of a regression model. The value of R-square lies between 0 to 1. Where we get R-square equals 1 when the model perfectly fits the data and there is no difference between the predicted value and actual value. However, we get R-square equals 0 when the model does not predict any variability in the model and it does not learn any relationship between the dependent and independent variables.

We calculate R-Square in the following steps:-

1. First, calculate the mean of the target/dependent variable y and we denote it by $\bar{y}$
2. Calculate the total sum of squares by subtracting each observation y_i from $\bar{y}$, then squaring it and summing these square differences across all the values. It is denoted by

   $$SStot = \sum_{i=1}^{n} (y_i - \bar{y})^2$$
3. We estimate the model parameter using a suitable regression model such as Linear Regression or SVM Regressor
4. We calculate the Sum of squares due to regression which is denoted by SSR. This is calculated by subtracting each predicted value of y denoted by y_pred_i from y_i squaring these differences and then summing all the n terms. SSR = \sum_{i=1}^{n} (\hat{ypred}_i – \bar{y})^2
5. We calculate the sum of squares *(SS_{res})*. It explains unaccounted variability in the dependent y after predicting these values from an

independent variable in the model. SSres = \sum_{i=1}^n (y_i – {ypred}_i)^2

6. we can then use either

$$R^2 = \frac{SSR}{SStot} or R^2 = 1 - \frac{SSres}{SStot}$$