Name: Enyia Esther

# $R^2$ Score Explanation And Meaning

The $R^2$ Score is a statistical measure used to access the goodness of fit of a regression model. It determines the proportion of variance in a dependent variable that can be predicted or explained by an independent variable. It usually ranges from 0 to 1, but it can sometimes take values like +inf or negative values.

Here is an explanation why we can get these values:

1. $R^2 = 0$ :

R-squared is 0 when the model explains 0% or no relationship between the dependent and independent variables. It shows that the model performs no better than a horizontal line(the mean of the target variable). It also means that the model doesn't predict any variability in the model.

2. $R^2 = 1$ :

R-squared is 1 when the model perfectly fits the data and predicts 100% variance. This could be because the model is overfitted or there is no noise in the data.

3. $R^2$ = +inf:

R-squared equals +inf when $R^2$ formula breaks down mathematically. It also shows the target variable has no variance. The value is evaluated on a constant y.

4. $R^2$ = negative:

R-squared equares negative when the model is worst than a horizontal line. This could be because the model is extremely poor or the model was trained on a different data. This happens when the sum of squared residuals (SSR) exceeds the total sum of squares (SST).

Conclusion

The $R^2$ score helps evaluate regression models, but extreme values indicate special cases:
- $R^2 = 0$: The model is no better than predicting the mean.
- $R^2 = 1$: The model fits perfectly (may indicate overfitting).
- $R^2 = +\infty$: The target variable has no variance (invalid for regression).
- $R^2$ = negative: The model performs worse than a horizontal line (likely misspecified).