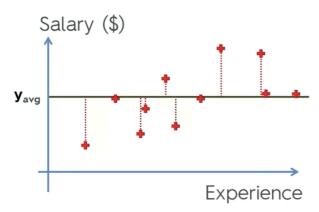
Adjusted R Squared Intuition

Machine Learning A-Z

© SuperDataScience

Adjusted R²

Simple Linear Regression:



 $SS_{res} = SUM (y_i - y_i^2)^2$

 $SS_{tot} = SUM (y_i - y_{avg})^2$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Data Science Training

© Kirill Eremenko

SS = Sum of Squares, res = residual, tot = total, avg =average

Adjusted R²

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$$y = b_0 + b_1 x_1$$

 $y = b_0 + b_1 x_1 + b_2 x_2$

Data Science Training

© Kirill Eremenko

Adjusted R²

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$$y = b_0 + b_1 x_1$$

$$y = b_0 + b_1^* x_1 + b_2^* x_2$$

Data Science Training

© Kirill Eremenko

Adjusted R²

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

R² - Goodness of fit (greater is better)

$$y = b_0 + b_1 x_1$$

Problem:

$$y = b_0 + b_1^* x_1 + b_2^* x_2 + b_3^* x_3$$

Data Science Training

© Kirill Eremenko

Adjusted R²

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

R² - Goodness of fit (greater is better)

$$y = b_0 + b_1 x_1$$

Problem:

 $y = b_0 + b_1 x_1 + b_2 x_2$

SS_{res}-> Min

R² will never decrease

Data Science Training

© Kirill Eremenko

Adjusted R²

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Adj R² = 1 -
$$(1 - R^2)\frac{n-1}{n-p-1}$$

- p number of regressors
- n sample size



Data Science Training

© Kirill Eremenko