


# Evaluation Metrics for Regression

---

Dr. JASMEET SINGH  
ASSISTANT PROFESSOR, CSED  
TIET, PATIALA

A solid orange horizontal bar spanning the width of the slide, located at the bottom.

# Regression Evaluation Metrics

---

- The performance of the regression model is generally measured in terms of error in prediction i.e., the difference between the actual values and the predicted values for all the instances in the test set.
- The various error metrics used in regression analysis are:
  1. Mean Absolute Error
  2. Mean Squared Error
  3. Root Mean Squared Error
  4.  $R^2$  Score (Coefficient of Determination)
  5. Adjusted  $R^2$  Score

# Mean Absolute Error

---

- The Mean Absolute Error(MAE) is the average of all absolute errors where absolute error is the absolute value of the difference between the measured value (predicted) and “true” value (actual).

$$\text{Mean Absolute Error (MAE)} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- where  $y_i$  is the actual value,  $\hat{y}_i$  is the predicted value of  $i^{\text{th}}$  input of test set and  $n$  are the total number of test samples.

# Mean Squared Error

---

- The **mean squared error (MSE)** or **mean squared deviation (MSD)** of an estimator (of a procedure for estimating an unobserved quantity) measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value

$$\text{Mean Squared Error (MSE)} = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^{\wedge})^2$$

- where  $y_i$  is the actual value,  $y_i^{\wedge}$  is the predicted value of  $i^{\text{th}}$  input of test set and  $n$  are the total number of test samples.

# Root Mean Squared Error

---

- The **root-mean-square deviation (RMSD)** or **root-mean-square error (RMSE)** is a frequently used measure of the differences between values (sample or population values) predicted by a model.
- It is the square root of the mean squared error.

$$\text{Root Mean Squared Error (RMSE)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^{\wedge})^2}$$

where  $y_i$  is the actual value,  $y_i^{\wedge}$  is the predicted value of  $i^{\text{th}}$  input of test set and  $n$  are the total number of test samples.

# R<sup>2</sup> Score/ Coefficient of Determination

---

- It measures the proportion of the variation independent variable explained by all the independent variables in the model.
- It assumes that every independent variable in the model helps to explain variation in the dependent variable.
- It is measured as the ratio of the explained variance of the model is to the total variance of the data.

$$R^2 = \frac{\text{Explained Variance of the model}}{\text{Total Variance of the data}} = \frac{\sum_{i=1}^n (y_i^{\wedge} - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

- where  $y_i$  is the actual value,  $y_i^{\wedge}$  is the predicted value of  $i^{\text{th}}$  input of test set,  $\bar{y}$  is the mean of actual values of  $y$  and  $n$  are the total number of test samples.

# R<sup>2</sup> Score

---

- Alternately, R<sup>2</sup> Score is measured from the unexplained variance as follows:

$$R^2 = 1 - \frac{\text{Unexplained Variance of the model}}{\text{Total Variance of the data}} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

- where SSE denote sum square error and SST denote sum square total.
- The value of R<sup>2</sup> lies in between -1 and 1. **R<sup>2</sup> is negative** only when the chosen model does not follow the trend of the data, so fits worse than the regression line.
- **Mathematically**, it is possible when error sum-of-squares from the model is larger than the total sum-of-squares from the horizontal line.

# Significance of $R^2$ Score

---

- R-squared is a statistical measure of how close the data are to the fitted regression line.
- 0% indicates that the model explains none of the variability of the response data around its mean.
- 100% indicates that the model explains all the variability of the response data around its mean.
- *Higher the R-squared, the better the model fits your data.*



# Evaluation Metrics- Numerical Example

---

Consider that the number lectures per day (x) affects the number of hours spent at university per day (y).

The equation of the regression line is

$$\hat{y} = 0.143 + 1.229x$$

Find

(i) MAE

(ii) MSE

(iii) RMSE

(iv)  $R^2$  Score

For the test set shown in Table

S.No	x	y
1	2	2
2	3	4
3	4	6
4	6	7

# Evaluation Metrics- Numerical Example

S.No	x	y	$\hat{y} = 0.143 + 1.229x$	Error = $y - \hat{y}$	Ab. Error = $ y - \hat{y} $	Sq. Error	$y - \text{mean}(y)$	SST
1	2	2	2.601	-0.601	0.601	0.361201	-2.75	7.5625
2	3	4	3.83	0.17	0.17	0.0289	-0.75	0.5625
3	4	6	5.059	0.941	0.941	0.885481	1.25	1.5625
4	6	7	7.517	-0.517	0.517	0.267289	2.25	5.0625
Total		19			2.229	1.542871	0	14.75

# Evaluation Metrics- Numerical Example

---

$$\text{Mean Absolute Error} = \frac{1}{n} \sum_{i=1}^n |y_i - y_i^{\wedge}| = \frac{2.29}{4} = 0.5725$$

$$\text{Mean Square Error} = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^{\wedge})^2 = \frac{1.542871}{4} = 0.3857$$

$$\text{Root Mean Square Error} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^{\wedge})^2} = 0.6210$$

$$R^2 \text{ score} = 1 - \frac{SSE}{SST} = 1 - \frac{1.542871}{14.75} = 0.895$$

# Adjusted $R^2$ Score

---

- It measures the proportion of variation explained by only those independent variables that really affect the dependent variable.
- It penalizes you for adding independent variable that do not affect the dependent variable.
- Every time you add a independent variable to a model, the R-squared increases, even if the independent variable is insignificant. It never declines.
- Adjusted R-squared increases only when independent variable is significant and affects dependent variable.

# Adjusted R<sup>2</sup> Score

---

- Adjusted R<sup>2</sup> is computed as follows:

$$\text{Adjusted } R^2 = 1 - \frac{(1 - R^2)(N - 1)}{(N - k - 1)}$$

- where R<sup>2</sup> is the sample R-square; k is the number of predictors (independent variables) and N is the total sample size.
- Adjusted R<sup>2</sup> score must be used to compare different regression models with different number of predictors and in case we want to decide the important predictors in our training set.