Regression Performance Evaluation

Tushar B. Kute,
http://tusharkute.com

Performance Evaluation

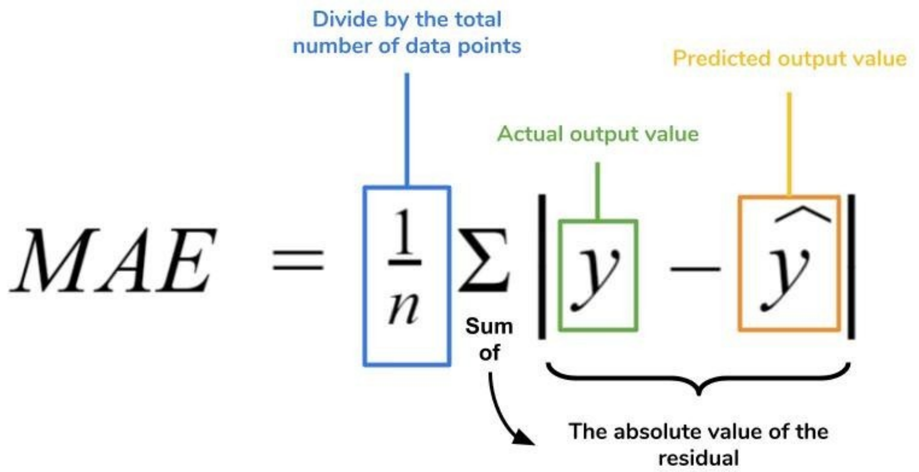
- The performance of a regression model can be understood by knowing the error rate of the predictions made by the model.
- You can also measure the performance by knowing how well your regression line fit the dataset.
- A good regression model is one where the difference between the actual or observed values and predicted values for the selected model is small and unbiased for train, validation and test data sets.

Performance Evaluation

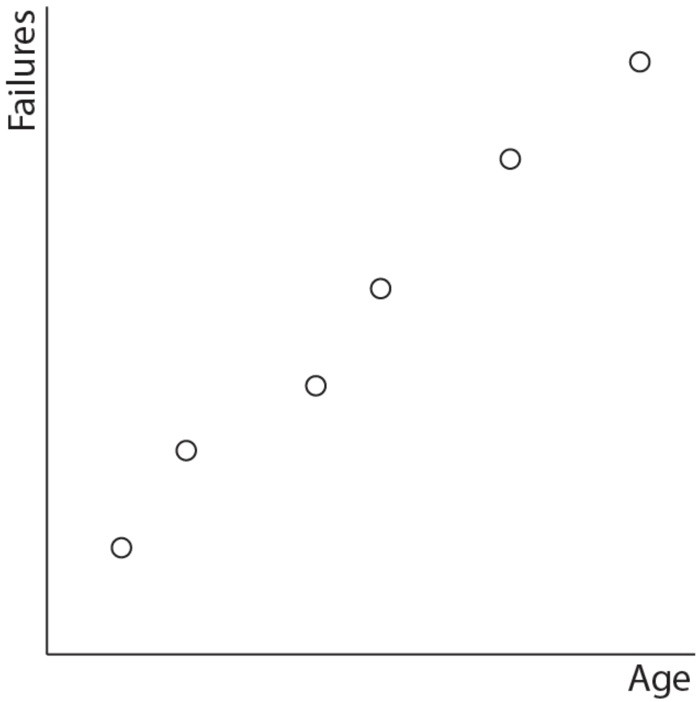
- To measure the performance of your regression model, some statistical metrics are used. Here we will discuss four of the most popular metrics. They are-
  - Mean Absolute Error(MAE)
  - Root Mean Square Error(RMSE)
  - Coefficient of determination or R2
  - Adjusted R2

Mean Absolute Error

- This is the simplest of all the metrics. It is measured by taking the average of the absolute difference between actual values and the predictions.

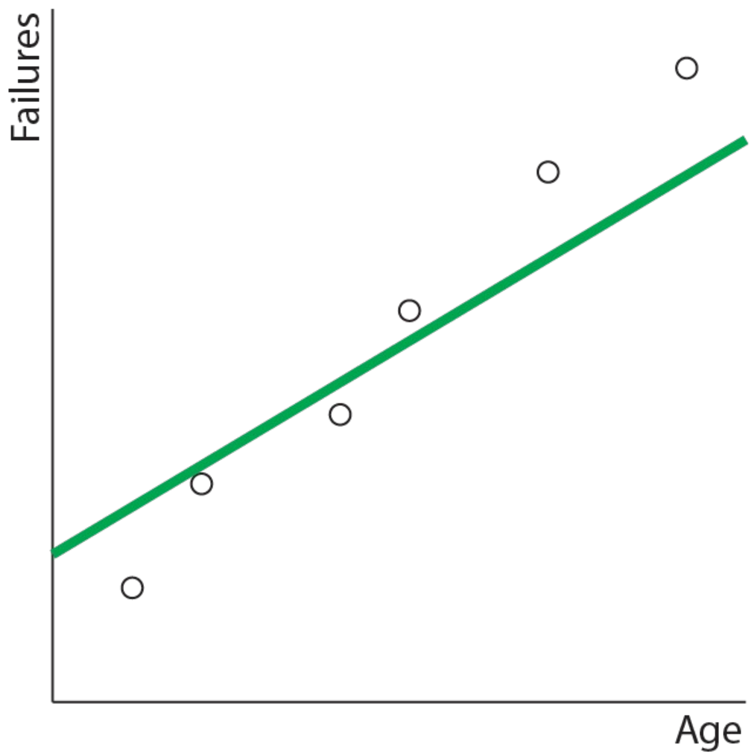


Example:



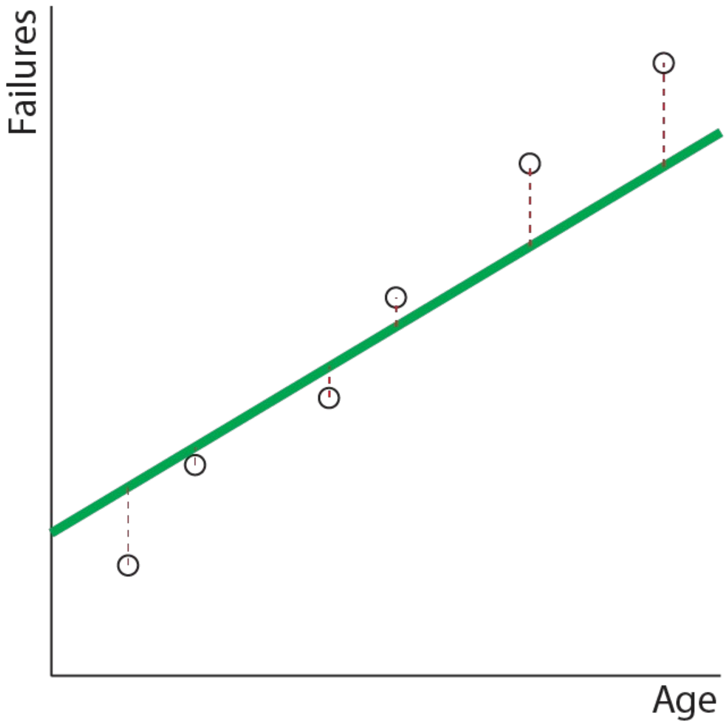
Age	Failures
10	15
20	30
40	40
50	55
70	75
90	90

Example:



Age	Failures	Prediction
10	15	26
20	30	32
40	40	44
50	55	50
70	75	62
90	90	74

Example:



Age	Failures	Prediction	Error
10	15	26	11
20	30	32	2
40	40	44	4
50	55	50	-5
70	75	62	-13
90	90	74	-16

Mean Absolute Error

Age	Failures	Prediction	Error
10	15	26	11
20	30	32	2
40	40	44	4
50	55	50	-5
70	75	62	-13
90	90	74	-16

abs(Error)
11
2
4
5
13
16

Mean abs(Error)	8.5

Mean Absolute Error

	y	$\hat{y}$	$y-\hat{y}$
Age	Failures	Prediction	Error
10	15	26	11
20	30	32	2
40	40	44	4
50	55	50	-5
70	75	62	-13
90	90	74	-16

$\lvert y-\hat{y}\rvert$
abs(Error)
11
2
4
5
13
16

Mean abs(Error)	$\frac{\Sigma\lvert y-\hat{y}\rvert}{N}$	**8.5**

Mean Absolute Error

- Mean Absolute Error (MAE) tells us the average error in units of y, the predicted feature. A value of 0 indicates a perfect fit, i.e. all our predictions are spot on.
- The MAE has a big advantage in that the units of the MAE are the same as the units of y, the feature we want to predict.
- In the example above, we have an MAE of 8.5, so it means that on average our predictions of the number of machine failures are incorrect by 8.5 machine failures.
- This makes MAE very intuitive and the results are easily conveyed to a non-machine learning expert!

Root Mean Square Error

- The Root Mean Square Error is measured by taking the square root of the average of the squared difference between the prediction and the actual value.
- It represents the sample standard deviation of the differences between predicted values and observed values(also called residuals).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

Root Mean Square Error

Age	y Failures	$\hat{y}$ Prediction	$y-\hat{y}$ Error
10	15	26	11
20	30	32	2
40	40	44	4
50	55	50	-5
70	75	62	-13
90	90	74	-16

$(y-\hat{y})^2$ Error^2
121
4
16
25
169
256

Mean of Error^2	$\frac{\Sigma(y-\hat{y})^2}{N}$	98.5
Square root of Mean of Error^2	$\sqrt{\frac{\Sigma(y-\hat{y})^2}{N}}$	**9.9**

RMSE

- As with MAE, we can think of RMSE as being measured in the y units.
- So the above error can be read as an error of 9.9 machine failures on average per observation.

MAE vs. RMSE

- Compared to MAE, RMSE gives a higher total error and the gap increases as the errors become larger. It penalizes a few large errors more than a lot of small errors. If you want your model to avoid large errors, use RMSE over MAE.
- Root Mean Square Error (RMSE) indicates the average error in units of y, the predicted feature, but penalizes larger errors more severely than MAE. A value of 0 indicates a perfect fit.
- You should also be aware that as the sample size increases, the accumulation of slightly higher RMSEs than MAEs means that the gap between these two measures also increases as the sample size increases.

R^2 Error

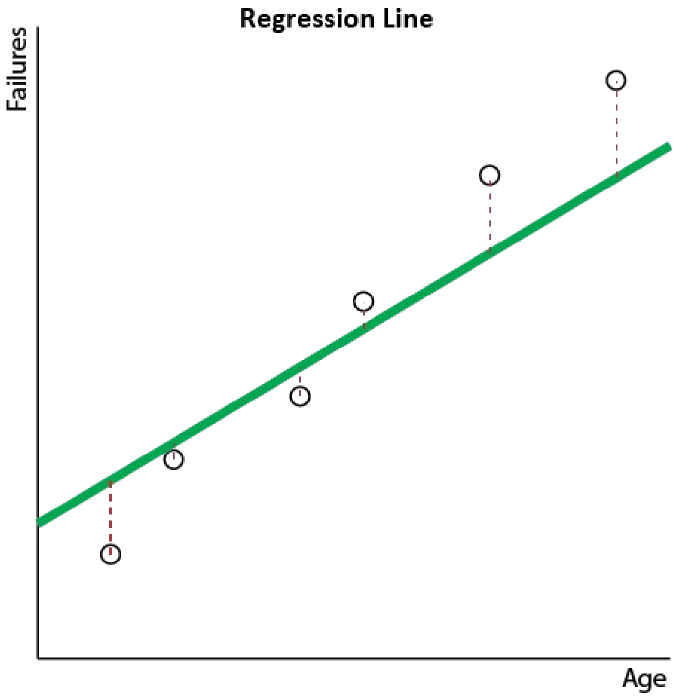
- It measures how well the actual outcomes are replicated by the regression line.
- It helps you to understand how well the independent variable adjusted with the variance in your model.
- That means how good is your model for a dataset. The mathematical representation for R^2 is-

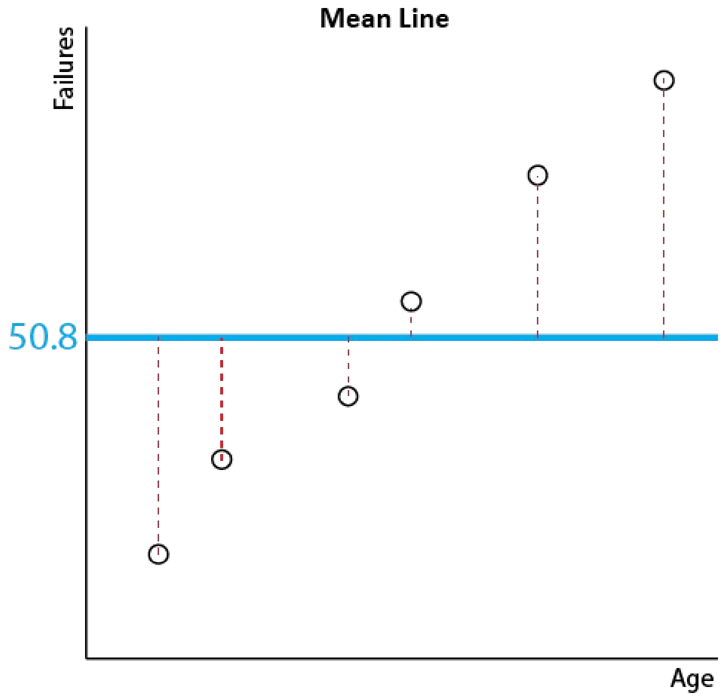
$$R^2 = \frac{SSR}{SST} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} \qquad R^2 = \frac{var(mean) - var(line)}{var(mean)}$$

R^2 Error

- Here,
  - SSR = Sum Square of Residuals(the squared difference between the predicted and the average value)
  - SST = Sum Square of Total(the squared difference between the actual and average value)

Example:





You can see that the regression line fits the data better than the mean line, which is what we expected (the mean line is a pretty simplistic model, after all). But can you say how much better it is? That's exactly what R2 does! Here is the calculation.

Example:

		Regression Line	Mean Line	Regression Line	Mean Line
	y	$\hat{y}$	$y-\hat{y}$	$y-\bar{y}$	$(y-\hat{y})^2$	$(y-\bar{y})^2$
Age	Failures	Prediction	Error	Error	Error^2	Error^2
10	15	26	11	-35.8	121	1281.6
20	30	32	2	-20.8	4	432.6
40	40	44	4	-10.8	16	116.6
50	55	50	-5	4.2	25	17.6
70	75	62	-13	24.2	169	585.6
90	90	74	-16	39.2	256	1536.6

Mean of Error^2	$\frac{\Sigma(y-\hat{y})^2}{N}$ 98.5	$\frac{\Sigma(y-\bar{y})^2}{N}$ 661.8

R^2	$\frac{\Sigma(y-\bar{y})^2-\Sigma(y-\hat{y})^2}{\Sigma(y-\bar{y})^2}$	**0.85**

R^2 Error

- The additional parts to the calculation are the column on the far right (in blue) and the final calculation row, computing R^2
- So we have an R-squared of 0.85. Without even worrying about the units of y we can say this is a decent model. Why? Because the model explains 85% of the variation in the data. That's exactly what an R-squared of 0.85 tells us!
- R-squared (R^2) tells us the degree to which the model explains the variance in the data. In other words, how much better it is than just predicting the mean.

Example:

- Here's another example. What if our data points and regression line looked like this?

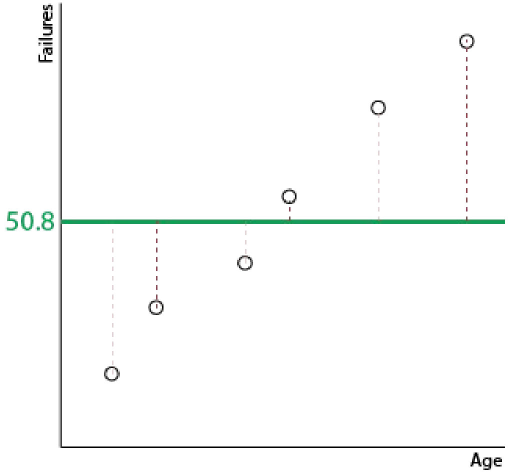
- The variance around the regression line is 0. In other words, var(line) is 0. There are no errors.

Now,

- Now, remember that the formula for R-squared is:
- $$R^2 = \frac{var(mean) - var(line)}{var(mean)}$$
- So, with var(line) = 0 the above calculation for R-squared is $$R^2 = \frac{var(mean) - 0}{var(mean)} = \frac{var(mean)}{var(mean)} = 1$$
- So, if we have a perfect regression line, with no errors, we get an R-squared of 1.

R^2 Error

- Let's look at another example. What if our data points and regression line looked like this, with the regression line equal to the mean line?
- Data points where the regression line is equal to the mean line

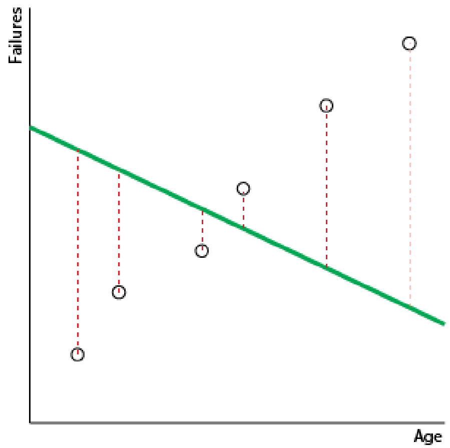


$$R^2 = \frac{var(mean) - var(mean)}{var(mean)} = \frac{0}{var(mean)} = 0$$

- In this case, var(line) and var(mean) are the same. So the above calculation will yield an R-squared of 0:

R^2 Error

- What if our regression line was really bad, worse than the mean line?



- It's unlikely to get this bad! But if it does, var(mean)-var(line) will be negative, so R-squared will be negative.
- An R-squared of 1 indicates a perfect fit. An R-squared of 0 indicates a model no better or worse than the mean. An R-squared of less than 0 indicates a model worse than just predicting the mean.

Adjusted R^2 Error

- Adjusted R-Squared is a modified form of R-Squared whose value increases if new predictors tend to improve models performance and decreases if new predictors does not improve performance as expected.
- R-squared is a comparison of Residual sum of squares (SSres) with total sum of squares(SStot).
- It is calculated by dividing sum of squares of residuals from the regression model by total sum of squares of errors from the average model and then subtract it from 1.

Adjusted R^2 Error

- Unlike R-squared, the Adjusted R-squared would penalize you for adding features which are not useful for predicting the target.
- It takes into account the number of independent variables used for predicting the target variable.

Adjusted R^2 Error

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

- where,
  - N = number of records in the data set.
  - p = number of independent variables.

Adjusted R^2 Error

- For a simple representation, you can rewrite the above formula like the following:

  Adjusted R-squared = 1 — (x * y)

  where,

  x = 1 — R Squared

  y = (N-1) / (n-p-1)

- Adjusted R-squared can be negative when R-squared is close to zero.
- Adjusted R-squared value always be less than or equal to R-squared value.

Mean Absolute Percentage Error (MAPE)

- The mean absolute percentage error (MAPE) is a measure of how accurate a forecast system is.
- It measures this accuracy as a percentage, and can be calculated as the average absolute percent error for each time period minus actual values divided by actual values.

Mean Absolute Percentage Error (MAPE)

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

- Where:
  - n is the number of fitted points,
  - A_t is the actual value,
  - F_t is the forecast value.
  - Σ is summation notation (the absolute value is summed for every forecasted point in time).

Summary

- Mean Absolute Error (MAE) tells us the average error in units of y, the predicted feature. A value of 0 indicates a perfect fit.
- Root Mean Square Error (RMSE) indicates the average error in units of y, the predicted feature, but penalizes larger errors more severely than MAE. A value of 0 indicates a perfect fit.
- R-squared (R^2) tells us the degree to which the model explains the variance in the data. In other words how much better it is than just predicting the mean.
  - A value of 1 indicates a perfect fit.
  - A value of 0 indicates a model no better than the mean.
  - A value less than 0 indicates a model worse than just predicting the mean.

Thank you

@mitu_skillologies

/mITuSkillologies

@mitu_group

/company/mitu-skillologies

MITUSkillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

contact@mitu.co.in

tushar@tusharkute.com