**Evaluation of Linear Regression**

- Soham Patel (46)

Practical - [click here]("Mean)

Model evaluation helps you to understand the performance of your model and makes it easy to present your model to other people. Here, we will discuss different metrics for regression model and difference between them. Using different evaluation model one can know **how close your prediction is against the real value**.

There are 3 main metrics for model evaluation in regression:

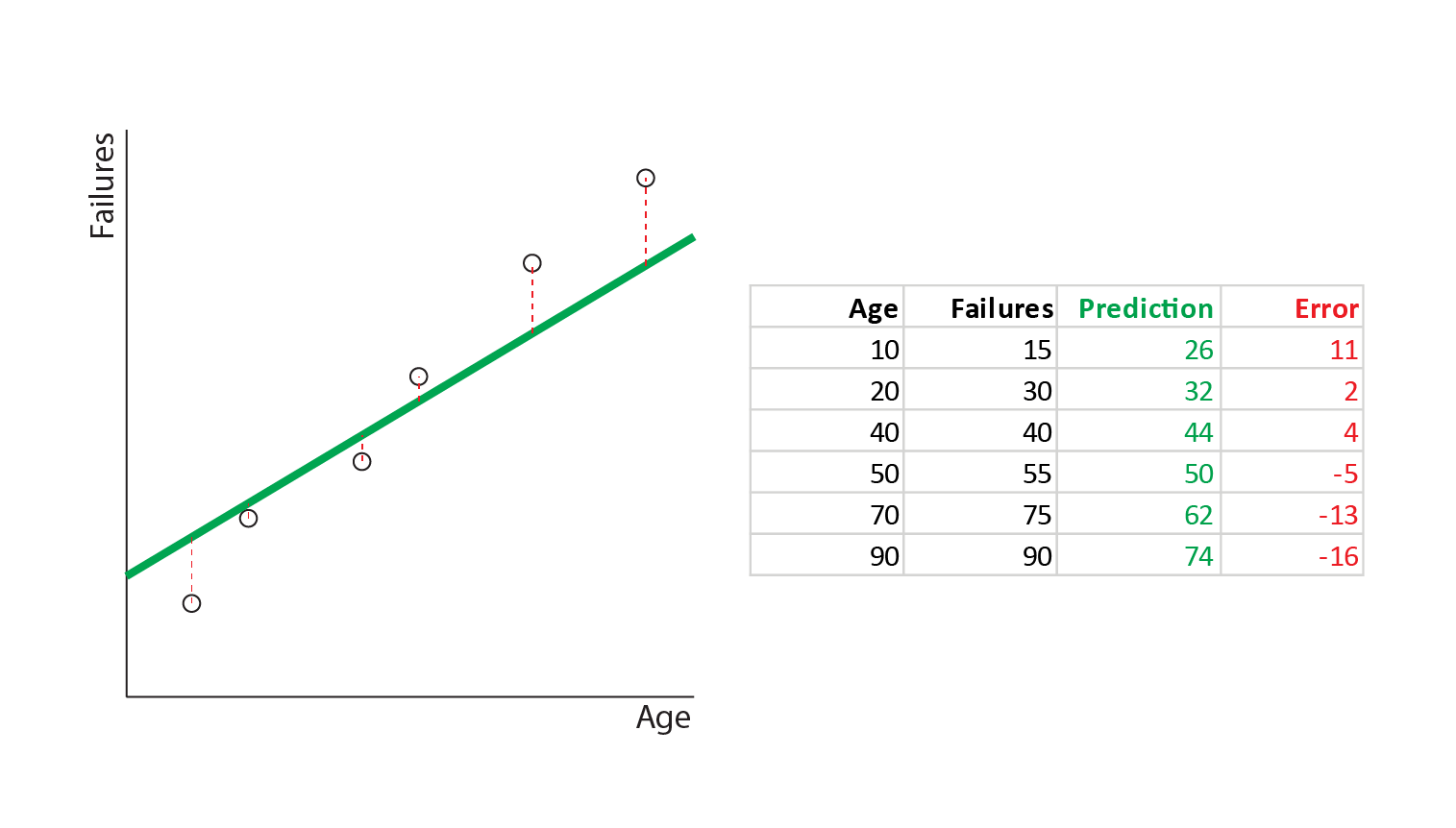
1. **Mean Absolute Error(MAE)**

We can define the MAE Formula as, **MAE = Σ|y−y^|/N**

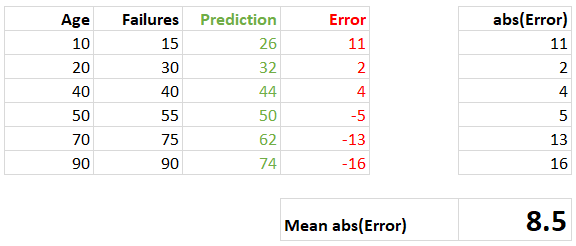
where y is the actual value ŷ is the predicted value and |y−y^| is the absolute value of the difference between the actual and predicted value. N is the number of sample points.

We take absolute value to remove the sign on the error value. If we don't, the positive and negative errors will tend to cancel each other out, giving a misleadingly small value for our evaluation metric.

In order to understand how well this line represents the actual data, we need to measure how good a fit it is. We can do this by measuring the distance from the actual data points to the line

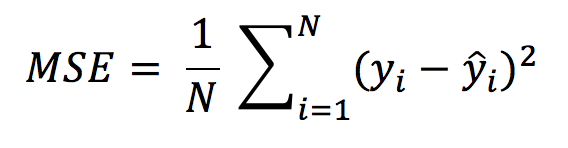


These red colored distances are called residuals or errors. The mean size of these errors is the MAE. We can calculate it as follows:



1. **Mean Square Error (MSE)**

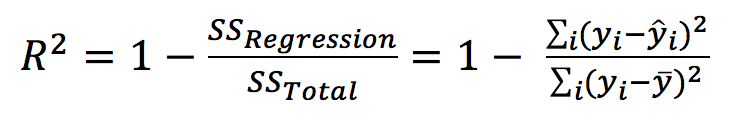
It indicates the average error in units of y, the predicted feature, but penalizes larger errors more severely than MAE.



Compared to MAE, MSE 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.

1. **R Square**

R Square is a good measure to determine how well the model fits the dependent variables. Here is the formula for R2



R Square is calculated by the sum of squared of prediction error divided by the total sum of square which replace the calculated prediction with mean. R Square value is between 0 to 1 and bigger value indicates a better fit between prediction and actual value.

**My Conclusion :**

MSE, RMSE or MAE are used to compare performance between different regression models while R Square is used to explain the model to other people because you can explain the number as a percentage of the output variability.

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 (R2 ) 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.