

[Open in app](#)

Published in Analytics Vidhya



Akshita Chugh

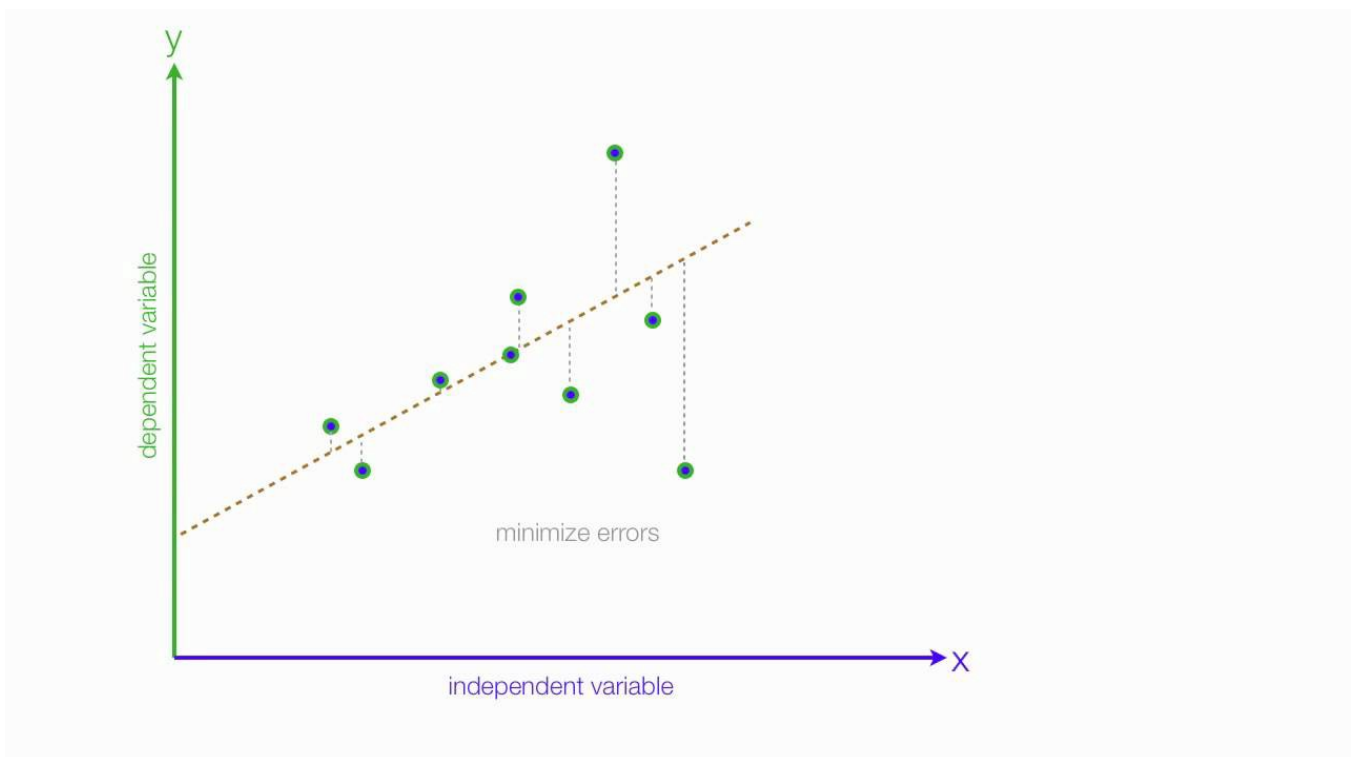
[Follow](#)Dec 8, 2020 · 4 min read · [Listen](#)

Save



# MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared — Which Metric is Better?

The objective of Linear Regression is to find a line that minimizes the prediction error of all the data points.



The essential step in any machine learning model is to evaluate the accuracy of the model. The Mean Squared Error, Mean absolute error, Root Mean Squared Error, and R-Squared or Coefficient of determination metrics are used to evaluate the performance of the model in regression analysis.



[Open in app](#)

of the residuals in the dataset.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}|$$

Where,

$\hat{y}$  — predicted value of  $y$   
 $\bar{y}$  — mean value of  $y$

- Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2$$

- Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2}$$

- The coefficient of determination or R-squared represents the proportion of the variance in the dependent variable which is explained by the linear regression




[Open in app](#)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$$

- Adjusted R squared is a modified version of R square, and it is adjusted for the number of independent variables in the model, and it will always be less than or equal to  $R^2$ . In the formula below  $n$  is the number of observations in the data and  $k$  is the number of the independent variables in the data.

$$R^2_{adj} = 1 - \left[ \frac{(1 - R^2)(n - 1)}{n - k - 1} \right]$$



301



2



### Differences among these evaluation

- Mean Squared Error(MSE) and Root Mean Square Error penalizes the large prediction errors vi-a-vis Mean Absolute Error (MAE). However, RMSE is widely used than MSE to evaluate the performance of the regression model with other random models as it has the same units as the dependent variable (Y-axis).
- MSE is a differentiable function that makes it easy to perform mathematical operations in comparison to a non-differentiable function like MAE. Therefore, in many models, RMSE is used as a default metric for calculating Loss Function despite being harder to interpret than MAE.
- The lower value of MAE, MSE, and RMSE implies higher accuracy of a regression model. However, a higher value of R square is considered desirable.
- R Squared & Adjusted R Squared are used for explaining how well the independent variables in the linear regression model explains the variability in the dependent variable. R Squared value always increases with the addition of the independent variables which might lead to the addition of the redundant variables in our model. However, the adjusted R squared solves this problem.



[Open in app](#)

value of Adjusted R squared decreases if the increase in the R square by the additional variable isn't significant enough.

- For comparing the accuracy among different linear regression models, RMSE is a better choice than R Squared.

## Conclusion

Both RMSE and R- Squared quantifies how well a linear regression model fits a dataset. The RMSE tells how well a regression model can predict the value of a response variable in absolute terms while R- Squared tells how well the predictor variables can explain the variation in the response variable.

---

## Sign up for Analytics Vidhya News Bytes

By Analytics Vidhya

Latest news from Analytics Vidhya on our Hackathons and some of our best articles! [Take a look.](#)

Emails will be sent to shivakmuddam25@gmail.com. [Not you?](#)



Get this newsletter

