What is Considered a Good RMSE Value?

Authored by

stats writer

April 29, 2024

RECOMMENDED CITATION

stats writer (2024). *What is Considered a Good RMSE Value?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=141222

ARABPSYCHOLOGY.COM

A good root mean square error (RMSE) value is typically considered to be low, indicating a small difference between the predicted and actual values. This indicates a high level of accuracy in the model's predictions. The exact value of a good RMSE may vary depending on the specific context and industry, but in general, a lower RMSE value is desirable. It is important to note that the interpretation of a good RMSE value may also depend on the scale of the data being analyzed.

What is Considered a Good RMSE Value?

One way to assess how well a fits a dataset is to calculate the root mean square error, which tells us the average distance between the predicted values from the model and the actual values in the dataset.

The formula to find the root mean square error, often abbreviated RMSE, is as follows:

RMSE = $\sqrt{\Sigma}$(Pi - Oi)2 / n

where:

Σ is a fancy symbol that means "sum"
Pi is the predicted value for the ith observation in the dataset
Oi is the observed value for the ith observation in the dataset
n is the sample size

One question people often have is: What is a good

ARABPSYCHOLOGY.COM

RMSE value?

The short answer: It depends.

The lower the RMSE, the better a given model is able to "fit" a dataset. However, the range of the dataset you're working with is important in determining whether or not a given RMSE value is "low" or not.

For example, consider the following scenarios:

Scenario 1: We would like to use a regression model predict the price of homes in a certain city. Suppose the model has an RMSE value of $500. Since the typical range of houses prices is between $70,000 and $300,000, this RMSE value is extremely low. This tells us that the model is able to predict house prices accurately.

Scenario 2: Now suppose we would like to use a regression model to predict how much someone will spend per month in a certain city. Suppose the model has an RMSE value of $500. If the typical range of monthly spending is $1,500 - $4,000, this RMSE value is quite high. This tells us that the model is not able to

ARABPSYCHOLOGY.COM

predict monthly spending very accurately.

These simple examples show that there is no universally "good" RMSE value. It all depends on the range of values in the dataset you're working with.

Normalizing the RMSE Value

One way to gain a better understanding of whether a certain RMSE value is "good" is to normalize it using the following formula:

Normalized RMSE = RMSE / (max value - min value)

This produces a value between 0 and 1, where values closer to 0 represent better fitting models.

For example, suppose our RMSE value is $500 and our range of values is between $70,000 and $300,000. We would calculate the normalized RMSE value as:

Normalized RMSE = $500 / ($300,000 - $70,000) = 0.002

Conversely, suppose our RMSE value is $500 and our range of values is between $1,500 and $4,000. We would calculate the normalized RMSE value as:

ARABPSYCHOLOGY.COM

Normalized RMSE = $500 / ($4,000 - $1,500) = 0.2.

The first normalized RMSE value is much lower, which indicates that it provides a much better fit to the data compared to the second normalized RMSE value.

Comparing RMSE Across Models

Instead of picking some arbitrary number to represent a "good" RMSE value, we can simply compare RMSE values across several models.

For example, suppose we fit three different regression models to predict house prices. Suppose the three models have the following RMSE values:

RMSE of Model 1: $550
RMSE of Model 2: $480
RMSE of Model 3: $1,400

Since the RMSE value of Model 2 is lowest, we would select Model 2 as the best model for predicting house prices since the average distance between the predicted prices and the actual prices is lowest for that model.

ARABPSYCHOLOGY.COM