

Testing

What is the best metric to judge model performance?

Here I will compile a list of different means for testing the machine learning and evaluate them before coming to a conclusion on which is best to choose. For these tests, we would use data from 20 years ago till 5 years ago and use that to forecast the next 5 years where we can then determine the effectiveness by comparing them.

Mean square error (MSE)¹

Mean Squared Error (MSE) testing refers to evaluating how accurate a model's predictions are by measuring the average of the squared differences between predicted values and actual values.²

This is effectively saying it measures the performance off of the average distance the model is from the actual prediction.

Example

Actual values: [3, 5, 7]

Predicted values: [2, 5, 8]

Errors:

- $(3-2)^2 = 1$
- $(5-5)^2 = 0$
- $(7-8)^2 = 1$

$$\text{MSE} = (1 + 0 + 1) / 3 = \mathbf{0.67}$$

So with this you would therefore want the mse to be closer to 0 to work effectively.

Mean absolute error

It is the average of the absolute differences between the predicted values and the actual values.

Simple Example

Actual values: 3, 5, 7

Predicted values: 2, 6, 8

¹ <https://neptune.ai/blog/performance-metrics-in-machine-learning-complete-guide>

² Chatgpt.com:prompt "what is mean square error testing"

Absolute errors:

- $|3 - 2| = 1$
- $|5 - 6| = 1$
- $|7 - 8| = 1$

$$\text{MAE} = \frac{1 + 1 + 1}{3} = 1$$

Root mean square error

RMSE is the square root of the average of the squared differences between predicted values and actual values.

Simple Example

Actual values: 3, 5, 7

Predicted values: 2, 6, 8

Squared errors:

- $(3 - 2)^2 = 1$
- $(5 - 6)^2 = 1$
- $(7 - 8)^2 = 1$

$$\text{RMSE} = \sqrt{\frac{1 + 1 + 1}{3}} = 1$$

mean accuracy percentage error (MAPE)

MAPE measures forecast accuracy as a percentage. It tells you, on average, how far off your predictions are relative to the actual values.

Simple Example

Actual values: 100, 200, 300

Predicted values: 90, 220, 270

Percentage errors:

- $|100 - 90| / 100 = 10\%$
- $|200 - 220| / 200 = 10\%$
- $|300 - 270| / 300 = 10\%$

$$\text{MAPE} = \frac{10 + 10 + 10}{3} = 10\%$$

this will probably be the best

Accuracy

Accuracy testing measures **how often a model's predictions are correct**. It's one of the most common evaluation methods, especially for **classification problems**.³

Example

Actual labels: [Cat, Dog, Dog, Cat]

Predicted labels: [Cat, Dog, Cat, Cat]

Correct predictions = 3

Total predictions = 4

Accuracy = → **75% accuracy**

This wouldn't be a very effective way of testing as this requires absolute accuracy to obtain a higher percentage but that just isn't likely to happen.

Ways of testing front end systems.

- Key ideas
 - o Break it.
 - Test every possible solution, eliminating any chance that the end user has of breaking it.
 - Input validation
 - In this instance, making sure that a product is there for data to be read would be a key feature
 - If it's not there, a suitable error message displaying clearly which would guide the user to be correct.
 - This feature may also need something of a spelling checker or auto correct – this would further reduce the user error rate and ultimately improve user experience and usability.
 - Test whether the backend is displaying accurate data in line with the backend
 - o Looks
 - Not necessarily a key component of the testing but would look unprofessional if not aligned properly or something of that calibre
 - Check spelling and grammar.

References

<https://neptune.ai/blog/performance-metrics-in-machine-learning-complete-guide>

chatgpt.com

³ Chatgpt.com;prompt "what is accuracy testing"