

Kirsten Barkved  
Head of Content

March 9, 2022

# How To Know if Your Machine Learning Model Has Good Performance



**Good accuracy in machine learning is subjective. But there's a sweet spot when it comes to model performance.**

Your [machine learning models](#) can be trained all day long, with many parameters and new techniques, but if you aren't evaluating it, you won't know if it's any good.

So, how do you know if the performance of your model is good? What does good accuracy look like? And what's the sweet spot for models?

We'll answer all your questions and show you what good accuracy looks like, as well how model performance in Obviously AI is evaluated!

Before we get to all that, we need to set the stage.

## Why Evaluate Model Performance?

It's incredibly important that your models produce high levels of performance. High-performing models means accurate and trustworthy predictions for your respective use cases.

After all, the best-run businesses are those that make informed decisions. And you can't make informed decisions if your predictions are inaccurate or faulty.

For classification problems, a very common way to evaluate performance is to measure its

**obviously.ai** Product

Plans

Case  
Studies

Blog

Sign  
in

Sign  
up

Accuracy is an evaluation metric particularly used for classification tasks. It represents the percentage of accurate predictions. We calculate it as a ratio of the total number of correct predictions to the total number of predictions generated by the model.

Since accuracy is easy to understand and implement, it is one of the most popular, and best-known, machine learning model validation methods.

For binary classification,  
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

In this instance, TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

For simple cases, it is usually best to access the performance of a machine learning algorithm/model by checking its accuracy.



Become the Data  
Scientist your team  
always needed.

**Get Started**

# Other Ways to Evaluate Machine Learning Models

While accuracy is the best known evaluation metric for classification, it might not always be enough while working with real life datasets.

Other important evaluation metrics for classification includes:

- Precision
- Recall
- AUC/ROC curve
- F-score

In the end, it's up to you to select the right metric(s) for your use case so you can effectively leverage your model and predictions.

## So, What Exactly Does Good Accuracy Look Like?

Good accuracy in machine learning is subjective. But in our opinion, **anything greater than 70%** is a great model performance. In fact, an accuracy measure of anything between 70%-90% is not only ideal, it's realistic. This is also

consistent with industry standards.

Anything below this range and it may be worth talking to a data scientist to understand what's going on. Our team of data scientists are always on hand to chat through your model's performance. They'll see if your dataset can be optimized to achieve better accuracy.

In fact, this is one of the reasons why [Obviously AI](#) is one of the industry's best: Our data science team is always there to help you out. You won't be alone to determine accuracy. Our data scientists can help optimize your model or just answer any questions you might have about the platform, your dataset, or even just machine learning in general.

## How Models Are Evaluated in Obviously AI

In our platform, you'll never have to worry about creating and efficiently evaluating your models—we do all of that for you! In fact, the entire process (training and testing) is conducted in a matter of seconds, so you don't have to

worry about fine-tuning.

However, we always believe that it's always good to know what's happening behind the scenes so it's not a black box. So let's take some time to explore what this whole process of evaluating models is like behind the scenes.

## Training and Testing Data

Machine learning uses algorithms to learn from data. They find patterns, develop understanding, make decisions, and evaluate those decisions.

In machine learning, datasets are split into two subsets: training data and testing data. And it's important to know the [difference between training and testing data](#).

The first subset is known as the **training data** – it's a portion of our actual dataset that is fed into the machine learning model to discover and learn patterns. In this way, it trains our model.

The other subset is known as the **testing data**. Once your machine learning model is built, i.e. trained, you need unseen

data to test your model. This data is called testing data, and you can use it to evaluate the performance and progress of your algorithms' training and adjust or optimize it for improved results.

Trained enough, an algorithm will essentially memorize all of the inputs and outputs in a training dataset – this becomes a problem when it needs to consider data from other sources, such as real-world customers.

In data science, it's typical to see your data split into 80% for training and 20% for testing.

The process of building a machine learning model is comprised of three steps:

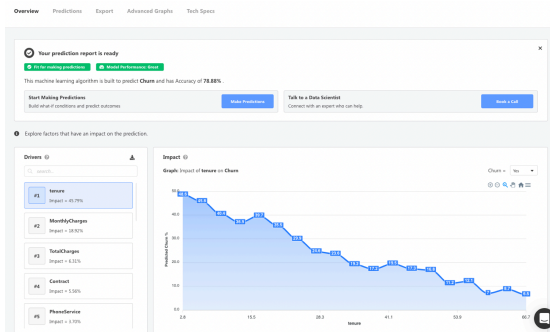
1. Feed - Feeding a model with training data
2. Define - The model learns patterns from the data
3. Test - Finally, the trained model is tested with previously unseen data, i.e. test data.

## Model Performance in Obviously AI

Let's take a look at a report generated in Obviously AI. We

used a sample dataset for employee attrition on our platform.

The Overview tab, pictured above, shows us our report.

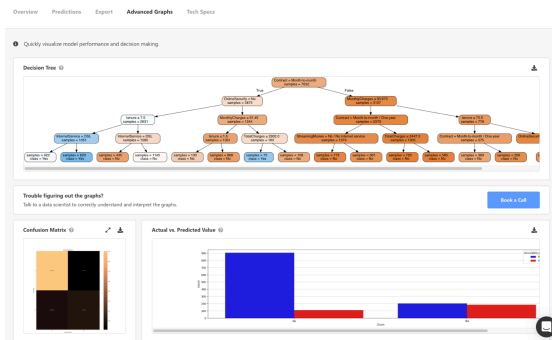


If we navigate to the **advanced graphs** tab, pictured below, we see the graphical analysis of the performance of the classification model on the test data. This section helps Obviously AI users to quickly visualize model performance and decision making. Remember: machine learning models are trained on 80% of data and model performance is evaluated on the other 20% of the data. This is the information we'll be working with in the **advanced graphs** section.

Here, we see three types of visualizations:

1. A decision tree
2. A confusion matrix
3. A bar chart depicting the Actual vs. Predicted values





The decision tree is generated on the whole dataset to give the user an idea of how a model makes decisions internally and arrives at conclusions.

A confusion matrix helps visualize the number of true positives, true negatives, false positives, and false negatives generated by the model (remember, these are the parameters for calculating accuracy).

On the Y-axis, we can see the actual label, and on the X-axis, we have the predicted labels. Ideally, the higher the values on the diagonal, the better the classification model; this tells us that the model predicted most of the values correctly.

The Actual vs. The Predicted Value bar chart represents the same information as the confusion matrix, but just with a

different and easily interpretable visualization.

The blue and red bars represent true and false labels respectively. In the graph on the x-axis, we have the classes, and on the y-axis we have the count/frequency of those classes.

For each class, we want to see a comparatively higher blue bar and a lower red bar; this tells us how many times the model correctly predicted a class.

## Accuracy in Obviously AI

Once a report is generated, the user is informed about the chosen model's accuracy score. However, if we hop over to the **Tech Specs tab**, there is also a detailed description of that classification model and its metrics.

In the model metrics window is the algorithm name and we can verify the use case is

classification. Here we can see that the platform chosen best classification algorithm is "X" with an accuracy score of "Y%."

We can always move down to the "Other Machine Learning Models that were trained" section at the end of this tab to check the accuracy measures of all the other algorithms that were parallelly trained and tested on the same data.

The Advanced Model Metrics section gives us a detailed description of the other classification metrics such as precision, recall, F score, AUC score, etc.

The best thing about our platform is that the way it chooses the best algorithm depends on multiple metrics along with the accuracy score. As mentioned before this ensures that we are not concentrating on only one metric but also taking into account other important metrics that helps in better interpretation and decision making on a model's performance.

## Summary

Like we said earlier: good

accuracy in machine learning is subjective. But in our opinion, **anything greater than 70%** is a great model performance.

Anything below this range and it may be worth talking to the Obviously AI data science team. They'll see if your dataset can be optimized to achieve better accuracy.

By doing so, you'll get help structuring your datasets to optimize your machine learning model to achieve higher accuracy. You're never left alone to figure things out.

Want to increase your model's accuracy, or maybe just ask some questions? [Book a call with us today.](#)

# Get Started Now

See how no code machine learning can transform your business and change how you make decisions.

See  
What's  
Possible

See it  
in  
Action

## Recommended Posts

No items found.

**obviously.ai**

The fastest, most precise  
No-Code AI tool in the world.

©2022 Obviously AI, Inc.

[Terms & Privacy](#)

### Case Studies

By Industry

Support

Finance

Operations

By Use Cases

Lead

Conversion

Dynamic  
Pricing

By Type

Classification

Regression



We're hiring! See [open roles](#).

Product  
Marketing  
Sales  
Time Series

Fraud  
Detection  
Property  
Prices

**Product**

Integrations  
Demo Video  
Data  
Requisites  
Algorithms  
Security

**Resources**

Help Center  
API Docs  
Status Page  
Obviously AI vs.  
AutoML  
Obviously AI vs.  
Coding

**Company**

About  
Careers  
Contact Us