



[Start Here](#) [Blog](#) [Books](#) [FAQ](#) [About](#) [Contact](#)



Need help with machine learning? [Take the FREE Crash-Course.](#)

What is the Difference Between a Parameter and a Hyperparameter?

by Jason Brownlee on July 26, 2017 in [Machine Learning Process](#)



It can be confusing when you get started in applied machine learning.

There are so many terms to use and many of the terms may not be used consistently. This is especially true if you have come from another field of study that may use some of the same terms as machine learning, but they are used differently.

For example: the terms “*model parameter*” and “*model hyperparameter*.”

Not having a clear definition for these terms is a common struggle for beginners, especially those that have come from the fields of statistics or economics.

In this post, we will take a closer look at these terms.



What is the Difference Between
Photo by Bruce Guen

Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE

What is a Model Parameter?

A model parameter is a configuration variable that is internal to the model and whose value can be estimated from data.

- They are required by the model when making predictions.
- Their values define the skill of the model on your problem.
- They are estimated or learned from data.
- They are often not set manually by the practitioner.
- They are often saved as part of the learned model.

Parameters are key to machine learning algorithms. They are the part of the model that is learned from historical training data.

In classical machine learning literature, we may think of the model as the hypothesis and the parameters as the tailoring of the hypothesis to a specific set of data.

Often model parameters are estimated using an optimization algorithm, which is a type of efficient search through possible parameter values.

- **Statistics:** In statistics, you may assume a distribution for a variable, such as a Gaussian distribution. Two parameters of the Gaussian distribution are the mean (μ) and the standard deviation (σ). This holds in machine learning, where these parameters may be estimated from data and used as part of a predictive model.

- **Programming:** In programming, you may pass a parameter to a function. In this case, a parameter is a function argument that could have one of a range of values. In machine learning, the specific model you are using is the function and requires parameters in order to make a prediction on new data.

Whether a model has a fixed or variable number of parameters determines whether it may be referred to as “*parametric*” or “*nonparametric*“.

Some examples of model parameters include:

- The weights in an artificial neural network.
- The support vectors in a support vector machine.
- The coefficients in a linear regression or logistic regression model.

What is a Model Hyperparameter?

A model hyperparameter is a configuration that is often used in processes to help estimate from data.

- They are often used in processes to help estimate from data.
- They are often specified by the practitioner.
- They can often be set using heuristics.
- They are often tuned for a given predictive model.

We cannot know the best value for a model hyperparameter, so we often use a rule of thumb, copy values used on other problems, or search online.

When a machine learning algorithm is tuned for a specific problem, such as when you are using a grid search or a random search, then you are tuning the hyperparameters of the model or order to discover the parameters of the model that result in the most skillful predictions.

“ Many models have important parameters which cannot be directly estimated from the data. For example, in the K-nearest neighbor classification model ... This type of model parameter is referred to as a tuning parameter because there is no analytical formula available to calculate an appropriate value.

— Page 64-65, [Applied Predictive Modeling](#), 2013

Model hyperparameters are often referred to as model parameters which can make things confusing. A good rule of thumb to overcome this confusion is as follows:

**If you have to specify a model parameter manually then
it is probably a model hyperparameter.**

Some examples of model hyperparameters include:

- The learning rate for training a neural network.
- The C and sigma hyperparameters for support vector machines.
- The k in k-nearest neighbors.

Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree.

Find out how in this free and practical email course.

Email Address

[START MY EMAIL COURSE](#)

of

Further Reading

- [Hyperparameter on Wikipedia](#)
- [What are hyperparameters in machine learning? on Quora](#)
- [What is the difference between model hyperparameters and model parameters? on StackExchange](#)
- [What is considered a hyperparameter? on Reddit](#)

Summary

In this post, you discovered the clear definitions and the difference between model parameters and model hyperparameters.

In summary, model parameters are estimated from data and are used in processes to help estimate values.

Model hyperparameters are often referred to as parameters of the learning that must be set manually and tuned.

Did this post help you clear up the confusion?
Let me know in the comments below.

Are there model parameters or hyperparameters that I missed?
Post them in the comments and I'll do my best to help.



Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



About Jason Brownlee

Jason Brownlee, Ph.D. is a machine learning specialist who teaches developers how to get results with modern machine learning methods via hands-on tutorials.

[View all posts by Jason Brownlee →](#)

◀ [How Much Training Data is Required for Machine Learning?](#)

[Why One-Hot Encode Data in Machine Learning? >](#)

64 Responses to *What is the Difference Between a Parameter and a Hyperparameter?*



Kiki July 26, 2017 at 7:32 am #

Awesome article! This was a big point of contention for me when I first got into machine learning. It's so important to keep these two concepts straight when it comes to disposal to tune my model — there are a lot of them!

REPLY ↗

[Your Start in Machine Learning](#)

car. 😊 Thank you for making this clear!



Jason Brownlee July 26, 2017 at 8:04 am #

REPLY ↩

Thanks. I'm glad it helped!



Dr Alan Beckles July 26, 2017 at 7:57 am #

REPLY ↩

Excellent post, Jason. Thanks!



Jason Brownlee July 26, 2017 at 8:04 am #

You're welcome Alan.



ujjawal sinha July 26, 2017 at 8:02 am #

Thanks Jason , Excellent



Jason Brownlee July 26, 2017 at 8:04 am #

REPLY ↩

I'm glad it helped.



Wesley July 26, 2017 at 8:04 am #

REPLY ↩

Great explanation...



Jason Brownlee July 26, 2017 at 8:05 am #

REPLY ↩

Thanks Wesley.



Deepak Sharma July 27, 2017 at 3:57 am #

REPLY ↩

Superb explanation Jason....love reading your articles!!!

Your Start in Machine Learning



Jason Brownlee July 27, 2017 at 8:11 am #

REPLY ↗

Thanks Deepak.



Jie July 27, 2017 at 6:06 pm #

REPLY ↗

In part model para, you give this example “The support vectors in a support vector machine.” I am a little confusing, why not the coefficients in SVM?



Jason Brownlee July 28, 2017 at 8:29 am #

We call the instances found by SVM “coefficients”.



Luis July 28, 2017 at 6:15 am #

Great post, Jason. Thanks!

One question: k-nearest neighbourhood is considered a parameter. Shouldn't k be considered as a hyperparameter then?

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Jason Brownlee July 28, 2017 at 8:36 am #

REPLY ↗

The “k” in kNN is a hyperparameter. I say exactly this Luis.



Luis July 28, 2017 at 6:22 am #

REPLY ↗

The confounding part was the use of “parameter” in:

“Many models have important parameters which cannot be directly estimated from the data. For example, in the K-nearest neighbor classification model ... This type of model parameter is referred to as a tuning parameter because there is no analytical formula available to calculate an appropriate value.”



Jason Brownlee July 28, 2017 at 8:37 am #

REPLY ↗

Why is this confounding Luis?

Your Start in Machine Learning



Tommy July 31, 2017 at 7:39 pm #

REPLY ↗

The book Applied Predictive Modeling does not contain the word hyperparameter. The article above states that many experts mix up the terms parameter and hyperparameter.

So what's the point of including the quote? Here are some potential answers:

1. The authors used the term “tuning parameter” incorrectly, and should have used the term hyperparameter. This understanding is supported by including the quote in the section on hyperparameters, Furthermore my understanding is that using a threshold for statistical significance as a tuning parameter may be called a hyperparameter because it

However, I believe that “tuning parameter” is not an incorrect description.

Also, you linked to the Wikipedia page for hyperparameters in Machine learning [https://en.wikipedia.org/wiki/Hyperparameter_\(machine_learning\)](https://en.wikipedia.org/wiki/Hyperparameter_(machine_learning))

The Wikipedia page gives the straightforward definition: hyperparameters are parameters whose values are not learned by the machine learning process. By contrast, the value of



Tommy July 31, 2017 at 7:56 pm #

Correct me if I'm wrong, but a hyperparameter is a type of parameter.

Synonyms for hyperparameters: tuning

Since hyperparameters are a type of parameter, the two terms are interchangeable when discussing hyperparameters. However, not all parameters are hyperparameters.

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Jason Brownlee August 1, 2017 at 7:58 am #

Nice perspective, thanks Tommy.

I cannot disagree generally, but the distinction is important, especially if you are a beginner trying to figure out what to “configure” or “tune”.



Jason Brownlee August 1, 2017 at 7:56 am #

REPLY ↗

Hi Tommy, I provided the quote to help clarify the definitions, not as an example of misuse. Sorry for the confusion.

Nice, your definition matches with the “estimated from data vs not” approach used in the post.

Sasikanth July 28, 2017 at 11:58 am #

Your Start in Machine Learning



Crystal clear. Thanks Jason



Jason Brownlee July 29, 2017 at 8:01 am #

REPLY ↗

I'm glad it helped.



Bharath Bhushan July 28, 2017 at 4:12 pm #

REPLY ↗

thanks. I was thinking both of them refer to the same thing. Thanks for clarification.



Jason Brownlee July 29, 2017 at 8:06 am #

I'm glad it help helped.



Ravindra July 28, 2017 at 4:34 pm #

Awesome! It was really confusing(parameters). Your post made it very clear.
Thank You!!

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Jason Brownlee July 29, 2017 at 8:06 am #

REPLY ↗

Happy it helped!



Abkul July 28, 2017 at 4:44 pm #

REPLY ↗

superbly explained.Thanks for the always handy post.



Jason Brownlee July 29, 2017 at 8:06 am #

REPLY ↗

Thanks!



Tim July 29, 2017 at 5:31 pm #

REPLY ↗

clf = svm.SVC(C = 0.01, kernel = 'rbf', rand

Your Start in Machine Learning

random_state is parameter or hyperparameter?



Jason Brownlee July 30, 2017 at 7:45 am #

REPLY ↗

Deep Tim... great question!

A gut check says “hyperparameter”, but we do not optimize it, we control for it. This feels wrong though. Perhaps it is neither.

What I mean is, it impacts the skill of the model, or most models that are stochastic, but we do not “tune” the value for a specific model/dataset. That’s the sense. Instead, we would re-run the experiment many times. We would create an ensemble of n final models.

Does that help? Am I making sense?



Vinícius July 30, 2017 at 1:28 am #

Excellent post! I am currently studying an application of machine learning for sonar classification and your posts have been very useful. I am trying to understand the advantage, do you have any material on this topic?

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Jason Brownlee July 30, 2017 at 7:48 am #

REPLY ↗

THanks.

Sorry, I don't have posts on these topics, I hope to get to them sometime.



Siva August 2, 2017 at 5:40 pm #

REPLY ↗

Good clarification and explanation. Thanks!



Jason Brownlee August 3, 2017 at 6:47 am #

REPLY ↗

Thanks Siva.



Gunjan August 31, 2017 at 9:09 pm #

REPLY ↗

Hi Jason, good explanation. I have one doubt that. if we have some hyperparameter for a given data sequence. Can we predict new set of hyperparameters?

Your Start in Machine Learning



Jason Brownlee September 1, 2017 at 6:45 am #

REPLY ↗

Parameters and hyperparameters refer to the model, not the data.



Antonio September 13, 2017 at 8:29 am #

REPLY ↗

To me, a model is fully specified by its family (linear, NN etc) and its parameters. The hyperparameters are used prior to the prediction phase and have an impact on the parameters, but are no longer needed. So coefficients in a linear model are not hyperparameters. The learning rate for a descent procedure is a hyperparameter. Structural parameters like the number of hidden units are somewhere in between. Some hyperparameters are implicit in the parameters themselves. Whether they are or not is a subtle distinction. On the other hand, I don't see that as a very helpful distinction. I think Tukey's quote is spot on. Two years ago. Tukey cites drawing something like a super-



Jason Brownlee September 13, 2017 at 10:05 pm #

Great note Antonio, thanks.

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Nicolas Marx October 19, 2017 at 10:50 pm #

REPLY ↗

Hi! great post, i was looking for this clarification. I wonder why it is not possible to tune the hyper-parameters from data, using another data partition as "hyperparameter test set"



Jason Brownlee October 20, 2017 at 5:35 am #

REPLY ↗

You can and this is what people do in a grid search or random search across hyperparameters.



john October 30, 2017 at 11:30 am #

REPLY ↗

a model is more strongly influenced whether by model parameters or hyperparameters?



Jason Brownlee October 30, 2017 at 3:50 pm #

REPLY ↗

A model is defined by its parameters.

The hyperparameters influence the training pro-

Your Start in Machine Learning

Does that help John?



Mark December 14, 2017 at 5:56 am #

REPLY ↗

A great article.

I have a question as to the mathematical meaning of a hyperparameter.

If one had to view a machine learning process as a function f that maps some input (from a domain) to some other type of output (codomain).

Is setting a value to a hyper-parameter the same as what is mathematically known as the 'restriction of a function'?

Thanks.



Jason Brownlee December 14, 2017 at 4:44 pm #

Perhaps. I'd recommend casting this



ahmed December 21, 2017 at 3:04 pm #

wow!!

that is awesome , it's finally clear

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Jason Brownlee December 21, 2017 at 3:35 pm #

REPLY ↗

I'm glad to hear that!



Sean December 23, 2017 at 12:44 am #

REPLY ↗

Are the hyperparameters the implementation of the learning bias or inductive bias??



Jason Brownlee December 23, 2017 at 5:20 am #

REPLY ↗

The hyperparameters can influence how biased the model is. Depends on the model as to which hyperparameters have an influence.



Ram February 27, 2018 at 1:06 am #

REPLY ↗

Your Start in Machine Learning

Let us consider in the case of XGBoost, the option: n_estimators, Max_depth etc., are called as hyperparameter ? Where as the number associated with the n_estimators =100 or Max_depth =6, let's 100 is called as parameter? Please clarify my confusion



Jason Brownlee February 27, 2018 at 6:34 am #

REPLY ↗

Yes, they are hyperparameters to the model and you specify values for those hyperparameters.

Parameters in the model may be specific split point values within a given tree within the xgboost model (e.g. not exposed).



Archit Rao March 4, 2018 at 1:38 pm #

Really great article. Infact all your articles clear many concepts.



Jason Brownlee March 5, 2018 at 6:21 am #

Thanks!



Manu March 20, 2018 at 8:17 pm #

REPLY ↗

Clear explanation. Thank you so much!!



Jason Brownlee March 21, 2018 at 6:31 am #

REPLY ↗

Thanks.



Nisa March 22, 2018 at 5:26 pm #

REPLY ↗

Hi Jason, do we need hyperparameter tuning while using clustering algorithm such as K-Means / Gaussian Mixture Model?

Thanks in advance 😊



Jason Brownlee March 23, 2018 at 6:03 am #

REPLY ↗

Your Start in Machine Learning

Sorry, I don't have material on clustering. I don't want to give you ad hoc misleading advice.



Anisah April 2, 2018 at 2:55 am #

REPLY ↗

Hi Jason, nice article. I am beginner in machine learning. Your article helps me so much. But, I got confuse about how to choose the best parameter? And we should regularly update the parameter right?



Jason Brownlee April 2, 2018 at 5:25 am #

REPLY ↗

The algorithm chooses the parameters for you. I recommend running experiments to see which chosen model on your specific dataset.

Model parameters may need to be updated if they are not correctly set. See this here:

<https://machinelearningmastery.com/gentle-introduction-machine-learning/>



NeuroMorphing May 1, 2018 at 9:06 am #

Dear Jason,

first of all thank you very much for the great article and the clarification regarding both types of parameters.

I have a question: At our organization we have the computational power to tune our hyperparameters via grid search. However, almost always when I pick the configuration of the hyperparameters that led to the maximum AUC on the training set, I'm facing pretty bad results on the test set, which implies overfitting... Which strategy would you recommend us to avoid this overfitting? Of course we could use random search (which would take much less runtime) but very likely would not lead to the best results. So what else can we do in such a case?

Thanks in advance...



Jason Brownlee May 2, 2018 at 5:36 am #

REPLY ↗

Perhaps grid search across configurations that have shown good results in the literature (e.g. use literature to define the bounding box of the search).



Parminder Kaur May 23, 2018 at 11:46 pm #

REPLY ↗

Hi...

I am using a deep convolutional neural network for

Your Start in Machine Learning

1. The variables used in training are filter weights, biases, learning rate, momentum etc...are they parameters or hyper-parameters?
2. In pooling and drop-out layer, i have defined stride factor and drop-out ratio....are they parameters or hyper-parameters?



Jason Brownlee May 24, 2018 at 8:14 am #

REPLY ↗

Weights are parameters, learning rate is a hyperparameter.

Network architecture is different again. It is more model design.



Indranil May 28, 2018 at 4:52 pm #

So since I am a machine learning practitioner
So whatever I can tune (or directly change) is hyperparameter
Whatever refuses to be in my direct control but is directly affected by me is a hyperparameter



Jason Brownlee May 29, 2018 at 6:23 am #

Nice.

Your Start in Machine Learning

You can master applied Machine Learning
without the math or fancy degree.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE

Leave a Reply

Name (required)

Email (will not be published) (required)

Website

Your Start in Machine Learning

[SUBMIT COMMENT](#)

Welcome to Machine Learning Mastery



Hi, I'm Jason Brownlee, Ph.D.

My goal is to make developers like *YOU* awesome at applied machine learning.

[Read More](#)

Need More?

Follow a top-down
approach with step-by-step
examples.

No math. No theory.

...take things slow.

[START](#)

Your Start in Machine Learning

You can master applied Machine Learning
without the math or fancy degree.

Find out how in this *free* and *practical* email course.

[START MY EMAIL COURSE](#)

POPULAR



[Multivariate Time Series Forecasting with LSTMs in Keras](#)

AUGUST 14, 2017



[How to Develop a Deep Learning Photo Caption Generator from Scratch](#)

NOVEMBER 27, 2017



[How to Use Word Embedding Layers for Deep Learning with Keras](#)

OCTOBER 4, 2017



[How to Define an Encoder-Decoder Sequence-to-Sequence Model for Neural Machine Translation in Keras](#)

OCTOBER 26, 2017



[How to Develop a Neural Machine Translation System from Scratch](#)

JANUARY 10, 2018



[How to Develop an Encoder-Decoder Model with Attention for Sequence-to-Sequence Prediction in Keras](#)

OCTOBER 17, 2017



[Why One-Hot Encode Data in Machine Learning?](#)

JULY 28, 2017

[Your Start in Machine Learning](#)



How to Reshape Input Data for Long Short-Term Memory Networks in Keras

AUGUST 30, 2017



How to Develop an Encoder-Decoder Model for Sequence-to-Sequence Prediction in Keras

NOVEMBER 2, 2017



What is the Difference Between Test and Validation Datasets?

JULY 14, 2017

You might also like...

- [Your First Machine Learning Project in Python](#)
- [Your First Neural Network in Python](#)
- [How to Setup a Python for Machine Learning](#)
- [Your First Classifier in Weka](#)
- [Your First Model for Time Series Forecasting](#)

© 2018 Machine Learning Mastery. All Rights Reserved.

[Privacy](#) | [Disclaimer](#) | [Terms of Service](#) | [Search](#) | [Newsletter](#)

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

[START MY EMAIL COURSE](#)

Your Start in Machine Learning