


# Parametric and Nonparametric Machine Learning Algorithms

by **Jason Brownlee** on March 14, 2016 in **Machine Learning Algorithms**



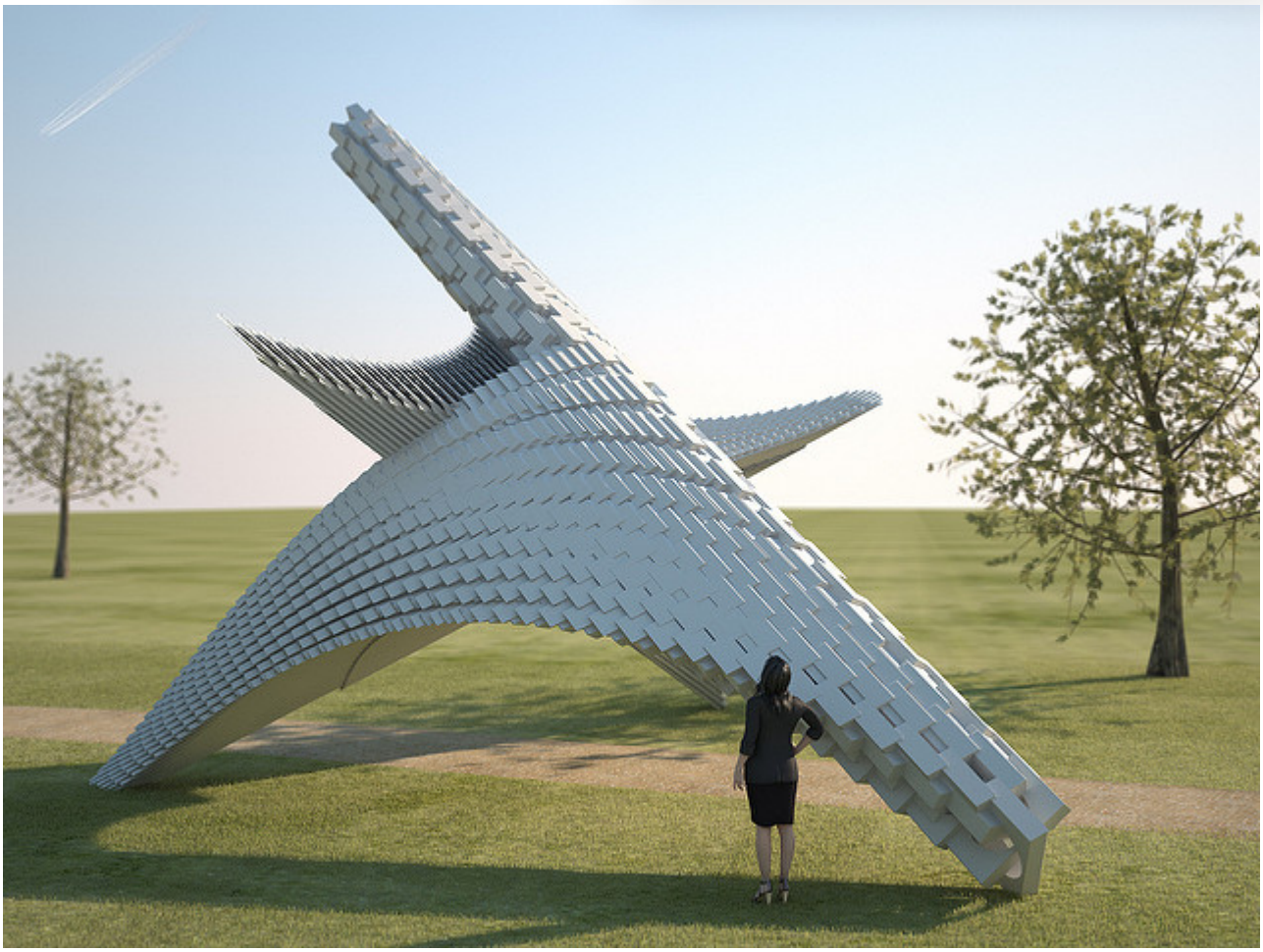
Last Updated on October 25, 2019

What is a parametric machine learning algorithm and how is it different from a nonparametric machine learning algorithm?

In this post you will discover the difference between parametric and nonparametric machine learning algorithms.

Discover how machine learning algorithms work including kNN, decision trees, naive bayes, SVM, ensembles and much more in my new book, with 22 tutorials and examples in excel.

Let's get started.



Parametric and Nonparametric Machine Learning Algorithms
Photo by John M., some rights reserved.

## Learning a Function

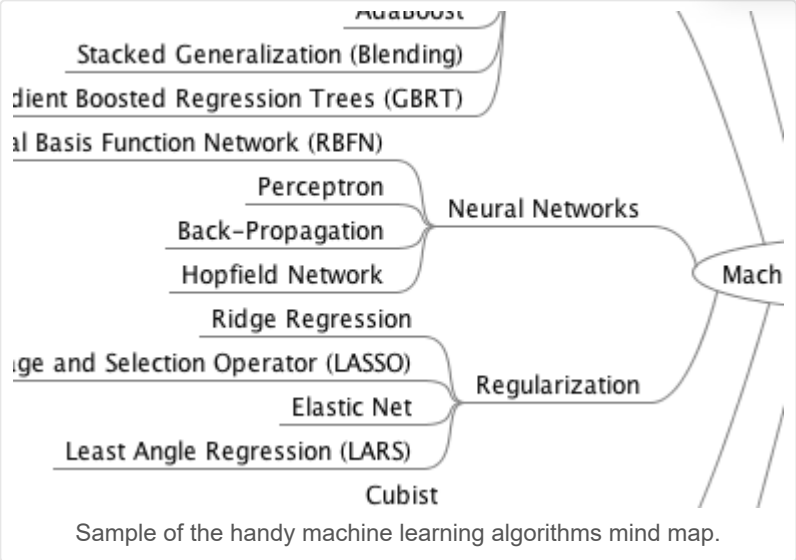
Machine learning can be summarized as learning a function (f) that maps input variables (X) to output variables (Y).

$$Y = f(x)$$

An algorithm learns this target mapping function from training data.

The form of the function is unknown, so our job as ma
machine learning algorithms and see which is better a

Different algorithms make different assumptions or bia
learned.



Sample of the handy machine learning algorithms mind map.



# Parametric Machine Learning Algorithms

Assumptions can greatly simplify the learning process, but can also limit what can be learned. Algorithms that simplify the function to a known form are called parametric machine learning algorithms.

> *A learning model that summarizes data with a set of parameters of fixed size (independent of the number of training examples) is called a parametric model. No matter how much data you throw at a parametric model, it won't change its mind about how many parameters it needs.*

— Artificial Intelligence: A Modern Approach, page 737

The algorithms involve two steps:

1. Select a form for the function.
2. Learn the coefficients for the function from the training data.

An easy to understand functional form for the mapping function is a line, as is used in linear regression:

$$b0 + b1*x1 + b2*x2 = 0$$

Where b0, b1 and b2 are the coefficients of the line that control the intercept and slope, and x1 and x2 are two input variables.

Assuming the functional form of a line greatly simplifie
estimate the coefficients of the line equation and we h

Often the assumed functional form is a linear combina
machine learning algorithms are often also called "*line*

The problem is, the actual unknown underlying functic
almost a line and require some minor transformation c
like a line in which case the assumption is wrong and

Some more examples of parametric machine learning algorithms include:

- Logistic Regression
- Linear Discriminant Analysis
- Perceptron
- Naive Bayes
- Simple Neural Networks

Benefits of Parametric Machine Learning Algorithms:

- **Simpler**: These methods are easier to understand and interpret results.
- **Speed**: Parametric models are very fast to learn from data.
- **Less Data**: They do not require as much training data and can work well even if the fit to the data is not perfect.

Limitations of Parametric Machine Learning Algorithms:

- **Constrained**: By choosing a functional form these methods are highly constrained to the specified form.
- **Limited Complexity**: The methods are more suited to simpler problems.
- **Poor Fit**: In practice the methods are unlikely to match the underlying mapping function.

# Nonparametric Machine Learning Algorithms

Algorithms that do not make strong assumptions about the form of the mapping function are called nonparametric machine learning algorithms. By not making assumptions, they are free to learn any functional form from the training data.

> *Nonparametric methods are good when you have a lot of data and no prior knowledge, and when you don't want to worry too much about choosing just the right features.*

— Artificial Intelligence: A Modern Approach, page 757

Nonparametric methods seek to best fit the training data in constructing the mapping function, whilst maintaining some ability to generalize to unseen data. As such, they are able to fit a large number of functional forms.

An easy to understand nonparametric model is the k-nearest neighbors algorithm that makes predictions based on the k most similar training patterns for a new data instance. The method does not make any assumptions anything about the form of the mapping function other than patterns that are close are likely to have a similar output variable.

Some more examples of popular nonparametric machine learning algorithms are:

- k-Nearest Neighbors
- Decision Trees like CART and C4.5

- Support Vector Machines

Benefits of Nonparametric Machine Learning Algorithms:

- **Flexibility**: Capable of fitting a large number of functional forms.
- **Power**: No assumptions (or weak assumptions) about the underlying function.
- **Performance**: Can result in higher performance models for prediction.

Limitations of Nonparametric Machine Learning Algorithms:

- **More data**: Require a lot more training data to estimate the mapping function.
- **Slower**: A lot slower to train as they often have far more parameters to train.
- **Overfitting**: More of a risk to overfit the training data and it is harder to explain why specific predictions are made.

# Further Reading

This section lists some resources if you are looking to learn more about the difference between parametric and non-parametric machine learning algorithms.

## Books

- An Introduction to Statistical Learning: with Applications in R, chapter 2
- Artificial Intelligence: A Modern Approach, chapter 18

## Posts

- What are the advantages of using non-parametric methods in machine learning? on Quora
- What are the disadvantages of non-parametric methods in machine learning? on Quora
- Nonparametric statistics on Wikipedia
- Parametric statistics on Wikipedia
- Parametric vs. Nonparametric on Stack Exchange

# Summary

In this post you have discovered the difference between parametric and nonparametric machine learning algorithms.

You learned that parametric methods make large assu to the output variable and in turn are faster to train, requi

You also learned that nonparametric methods make fe d in turn require a lot more data, are slower to train and ha re powerful models.

If you have any questions about parametric or nonpar leave a comment and I will do my best to answer them

**Update**: I originally had some algorithms listed under the wrong sections like neural nets and naive bayes, which made things confusing. All fixed now.





### About Jason Brownlee

Jason Brownlee, PhD 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 Machine Learning Algorithms Work (they learn a mapping

ms ›



## 43 Responses to *Parametric and Nonpa*

**confused beginner** March 14, 2016 at 6:02 pm #

hi jason

thanks for taking your time to summarize these topics so that even a novice like me can understand. love your posts

i have a problem with this article though, according to the small amount of knowledge i have on parametric/non parametric models, non parametric models are models that need to keep the whole data set around to make future predictions. and it looks like Artificial Intelligence: A Modern Approach, chapter 18 agrees with me on this fact stating neural nets are parametric and once the weights w are learnt we can get rid of the training set. i would say its the same case with trees/naive bays as well.

so what was your thinking behind in categorizing these methods as non-parametric?

thanks,
a confused beginner

**Jason Brownlee** July 17, 2016 at 6:57 am #

REPLY

Indeed simple multilayer perceptron neural nets are parametric models.

Non-parametric models do not need to keep the whole dataset around, but one example of a non-parametric algorithm is kNN that does keep the whole dataset. Instead, non-parametric models can vary the number of parameters, like the number of nodes in a decision tree or the number of support vectors, etc.

**mlvi** July 27, 2017 at 1:49 am #

REPLY

Isn't number of nodes in the decision tree a hyper parameter?

One more question is, How do you deploy non parametric machine learning models in production as there parameters are not fixed?

**Jason Brownlee** July 27, 2017 at 8:10 am #

REPLY

No, but the max depth of the tree is

You can finalize your model, save it to file

See this post:
http://machinelearningmastery.com/train-fi

Does that help?

**Another confused beginner** March 15, 2016 a

I am also interesting to know why Naive Baye

**Jason Brownlee** July 17, 2016 at 7:06 am #

REPLY

Yes, Naive bayes is generally a parametric method as we choose a distribution (Gaussian) for the input variables, although there are non-parametric formulations of the method that use a kernel estimator. In fact, these may be more common in practice.

**Ecolss** March 15, 2016 at 5:41 pm #

REPLY

Confused here too.

AFAIK, parametric models have fixed parameter set, i.e. the amount of parameters won't change once you have designed the model, whereas the amount of parameters of non-parametric models varies, for example, Gaussian Process and matrix factorization for collaborative filtering etc.

Correct me if I'm wrong ☺

**Jason Brownlee** July 17, 2016 at 7:06 am #

REPLY

This is correct.

**Simon Tse** July 16, 2016 at 10:21 pm #

REPLY

I think the classification does not really depend on what 'parameters' are. It's about the assumption you have made when you try to construct a model or function. Parametric models usually has a probability model (i.e. pdf) behind to support the function-finding process such as normal distribution or other distribution model.

On the other hand, non-parametric model just depends on the error minimisation search process to identify the set of 'parameters' which has nothing to do with a pdf.

So, parameters are still there for both parametric and n
layer of assumption to govern the nature of pdf of whic

**Jason Brownlee** July 17, 2016 at 7:10 am #

Hi Simon, the statistical definition of parar

The crux of the definition is whether the number of

It might be more helpful for us to consider linear ar

**Kevin** August 11, 2016 at 1:11 pm #

REPLY

Is there a relation between parametric/nonparametric models and lazy/eager learning?

**ANUDEEP VANJAVAKAM** September 24, 2016 at 11:29 am #

REPLY

In machine learning literature, nonparametric methods are also call instance-based or memory-based learning algorithms.
-Store the training instances in a lookup table and interpolate from these for prediction.
-Lazy learning algorithm, as opposed to the eager parametric methods, which have simple model and a small number of parameters, and once parameters are learned we no longer keep the training set.

**Jianye** September 27, 2016 at 11:18 am #

REPLY

I have questions of distinguishing between parametric and non parametric algorithms: 1) for linear regression, we can also introducing x^2, x^3 … to make the boundary we learned nonlinear, does it mean that it becomes non parametric in this case?

2) The main difference between them is that SVM puts additional constraints on how do we select the hyperplane . Why perception is considered as parametric while svm is not?

**Jason Brownlee** September 28, 2016 at 7:35 am #

REPLY

Hi Jianye,

When it comes down to it, parametric means a fixed number of model parameters to define the modeled decision.

Adding more inputs makes the linear regression ec

SVM can choose the number of support vectors ba
it non-parametric.

I hope that is clearer.

**Pramit Choudhary** January 23, 2017 at 1:09 pm #

Hi Jason,
Nice content here. Had some suggestions,
1. Do you think, it would be a good idea to include histogram: as a simple non-parametric model for estimation probability distribution ? Some beginners might be able to related to histograms.
2. Also, may be mentioning SVM(RBF kernel) as non-parametric to be precise.
What do you think ?

**Jason Brownlee** January 24, 2017 at 10:54 am #

REPLY

Hi Pramit,

1. nice suggestion.
2. perhaps, there is debate about where SVM sits. I do think it is nonparametric as the number of support vectors is chosen based on the data and the interaction with the argument-defined margin.

**Manish Barnwal** March 30, 2017 at 8:50 pm #

REPLY

Jason, as always, an excellent post.

**Jason Brownlee** March 31, 2017 at 5:54 am #

REPLY

Thanks Manish.

**amr gamal** April 12, 2017 at 1:40 am #

REPLY

jason ,it is a good post about parametric and non parametric model
but i still confused
did deep learning supposed to be parametric or non parametric and why
Best Regards

**Jason Brownlee** April 12, 2017 at 7:55 am #

There is not a hard line between paramet

I think of neural nets as non-parametric myself.

See this:
https://www.quora.com/Are-Neural-Networks-para

**anubhav** November 25, 2018 at 5:47 am #

REPLY

Hello Jason, thanks for discussing this topic. your consideration of NN as non-parametric doesn't make sense to me as per your post & suggestions above.! Please [read here](https://stats.stackexchange.com/questions/322049/are-deep-learning-models-parametric-or-non-parametric) & [hear](r/MachineLearning – Is artificial neural network a parametric or non-parametric method? (https://www.reddit.com/r/MachineLearning/comments/kvkud/is_artificial_neural_network_a_parametric_or/c2noexo)) for more clarity and correct me where I am wrong.!

**Jason Brownlee** November 25, 2018 at 7:01 am #

REPLY

Perhaps you can summarize the links for me?

**Aishwarya** May 4, 2017 at 8:10 am #

REPLY

Hi
The answer is very convincing, i just have a small question, for pressure distribution plots which ML algorithm should we consider?

**Jason Brownlee** May 5, 2017 at 7:26 am #

REPLY

Sorry, I don't know what pressure distribution plots are.

**Sanket Maheshwari** May 17, 2017 at 7:45 am #

REPLY

Hi Jason,

Decision tree contains parameters like Splitting Criteria
Maximal Depth then why it is called as non-parametric

**Jason Brownlee** May 17, 2017 at 8:45 am #

They are considered hyperparameters of

The chosen split points are the parameters of the
data. Thus, the decision tree is a nonparametric al

Does that make sense?

**Sanket Maheshwari** May 18, 2017 at 7:37 pm #

REPLY

Could you please briefly tell me what are the parameters and hyperparameters in the following models:

1.Naive Baye
2.KNN
3.Decision Tree
4.Multiple Regression
5.Logistic Regression

**Jason Brownlee** May 19, 2017 at 8:16 am #

REPLY

Yes, please search the blog for posts on each of these algorithms.

**Guiferviz** November 3, 2017 at 10:45 pm #

REPLY

Hi Jason! Nice blog.

I have a doubt about the “simple neural networks”, shouldn’t it be “neural networks” in general? The number of parameters is determined a priori.

In addition, I think that linear SVM might be considered as a parametric model because, despite the number of support vector varies with the data, the final decision boundary can be expressed as a fixed number of parameters.

I know the distinction between parametric and non-parametric is a little bit ambiguous, but what I said makes sense, right?

**duribef** May 12, 2018 at 4:06 am #

Up to this question! I have the same doub

Saludos!

**Aniket Saxena** November 7, 2017 at 3:25 am #

Hi Jason, I want to know that despite having not
required much data to train, does the parametric algorithms also cause overfitting? Or can they be lead to underfitting, instead?

**Jason Brownlee** November 7, 2017 at 9:53 am #

REPLY

Both types of algorithms can over and under fit data.

It is more common that parametric underfit and non-parametric overfit.

**Aniket Saxena** November 8, 2017 at 12:20 am #

REPLY

Hi Jason, thanks for your help but there is a request by my side to also look question posted above my question because it is a nice question about distinction between parametric vs non-parametric and I am very curious to know your opinion about this question posted by Guiferviz on november 3, 2017. Please answer to this question…….

**Magnus** January 31, 2018 at 9:10 pm #

REPLY

Hi Jason, you mention that simple multilayer perceptron neural nets are parametric models. This I understand, but which neural networks are then non-parametric? I assume e.g. that neural nets with dropouts are non-parametric?

**Jason Brownlee** February 1, 2018 at 7:19 am #

REPLY

Perhaps. Categorizing algorithms gets messy.

**ali** October 1, 2018 at 11:34 pm #

if we are doing regression for decision trees
do we need to check for correlation among the feature
when we talk about nonparamertic or parametric are w
talking about the data.

and if my data are not normally distributed do I have to
want to use parametric or nonparamertic

**Jason Brownlee** October 2, 2018 at 6:25 am #

REPLY

It is a good idea to make the problem as simple as possible for the model.

Nonlinear methods do not require data with a Normal distribution.

**sindhu** October 7, 2018 at 12:37 am #

REPLY

Hi Jason,
Good post.Could u pls explain parametric and non parametric methods by an example?
Bit confused about the parameters(what are the parameters,model parameters).For example,in the script the X and y values are the parameters?

**Jason Brownlee** October 7, 2018 at 7:26 am #

REPLY

Great question, this post will help:
https://machinelearningmastery.com/difference-between-a-parameter-and-a-hyperparameter/

**Yogesh Soni** June 9, 2019 at 4:19 am #

REPLY

Hi Jason

Can you post or let me know about parameter tuning.

**Jason Brownlee** June 9, 2019 at 6:22 am #

REPLY

Yes, I have many posts, try a search for “grid search”

**Smita Bhagwat** December 4, 2019 at 1:42 am #

Hi Jason, Can you throw some light on Semi

**Jason Brownlee** December 4, 2019 at 5:40 a

I’ve not heard of them before.

Do you have an example?

## Leave a Reply

Name (required)

Email (will not be published) (required)

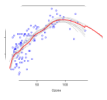
Website

SUBMIT COMMENT

**Welcome!**
My name is *Jason Brownlee* PhD, and I **help developers** get results with **machine learning**.
Read more

### Never miss a tutorial:

### Picked for you:



A Tour of Machine Learning Algorithms

Supervised and Unsupervised Machine Learning Al

Logistic Regression Tutorial for Machine Learning

Simple Linear Regression Tutorial for Machine Learning

Bagging and Random Forest Ensemble Algorithms for Machine Learning