

# Chapter-3 Bias and Variance Trade-off in Machine Learning



Ashish Patel

Following

Jun 15, 2018 · 4 min read

Machine Learning Series!!!

Hello Folks, Once Again I am posting amazing and controversial topic of machine learning for Beginner data scientist. So, I made this clear in this article what is bias , what is variance. after reading this you are able to understand what is bias and variance and where it is useful. So, let's begin..

## Outline

1. What is Bias and Variance Definition?

2. Bias Error With Example

3. Variance Error with Example

4. Over-fitting and Under-fitting

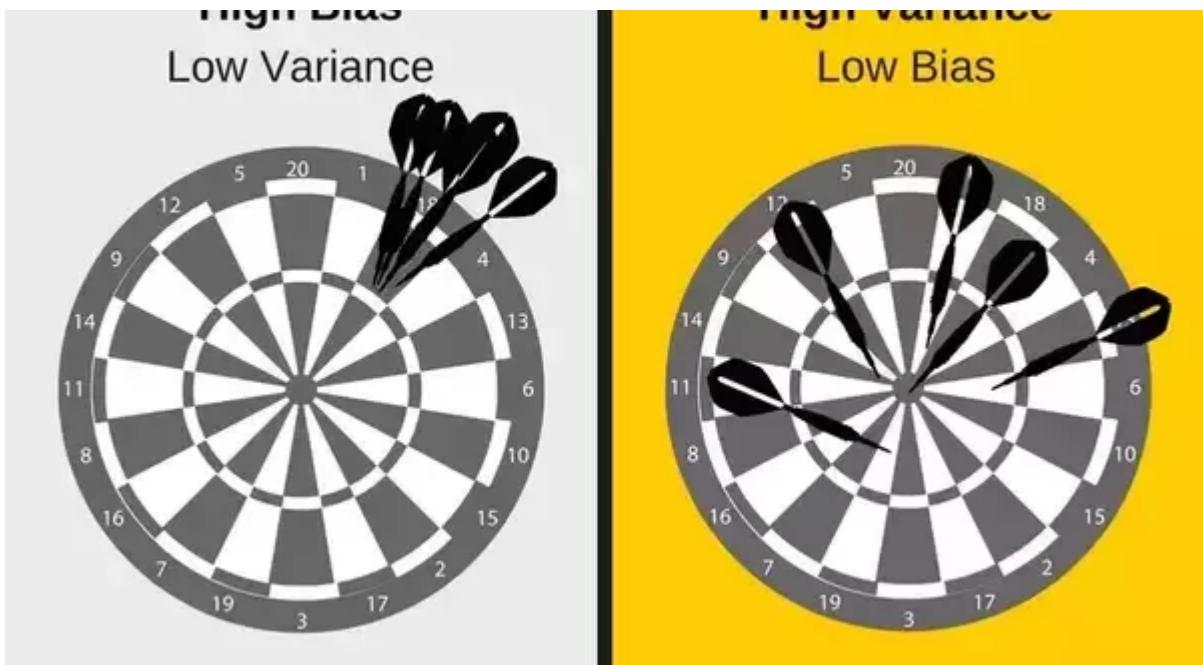
5. Big Picture of Bias and Variance.

## 1.What is Bias and Variance Definition?

In Statistics, Bias and variance is property of predictive model and machine learning it's calculated for supervise machine learning algorithm.

This picture should be a great way to explain bias-variance to a 5-year-old.





For the group of smart people with a basic understanding of modeling, statistics, and ML, let's take a slightly deeper look

$$\text{error}(X) = \text{noise}(X) + \text{bias}(X)^2 + \text{variance}(X)$$

1. **Bias:** High error due to assumption.
2. **Variance :** Error due to an overly-complex model that tries to fit the training data as closely as possible.
3. **Trade-off:** a balance achieved between two desirable but incompatible features; a compromise.

## 2.Bias Error With Example

*Note: Bias results in under-fitting the data. A high bias means our learning algorithm is missing important trends among-st the features.*

Simply Bias is difference between predicted value and actual value difference from training data, on which we are trained the machine learning model. Bias are the simplifying assumptions made by a model to make the target function easier to learn.

$$\text{Bias}(X) = E[\hat{f}(x)] - f(x)$$

High bias algorithms are easier to learn but less flexible, due to this they have lower predictive performance on complex problems. Linear algorithms and oversimplified model lead to high bias in the model. Let's see below table of bias so we make more clear picture.



Bias Categories

**Examples of low-bias machine learning algorithms include:** Decision Trees, k-Nearest Neighbors and Support Vector Machines.

**Examples of high-bias machine learning algorithms include:** Linear Regression, Linear Discriminant Analysis and Logistic Regression.

### 3. Variance Error with Example

Simply Variance is when your training data is change your model give the different result and this different result from the first result have variation. The estimate of the target function will change if different training data was used.

$$\text{Var}(X) = E[(\hat{f}(x) - E[\hat{f}(x)])^2]$$

Generally, non-parametric machine learning algorithms that have a lot of flexibility have a high variance. For example, decision trees have a high variance, that is even higher if the trees are not pruned before use.



Variance

**Examples of low-variance machine learning algorithms include:** Linear Regression, Linear Discriminant Analysis and Logistic Regression.

**Examples of high-variance machine learning algorithms include:** Decision Trees, k-Nearest Neighbors and Support Vector Machines.

## 4. Simple Definition Over-fitting and under-fitting

**Overfitting:** Good performance on the training data, poor result while giving other data.

**Underfitting:** Poor performance on the training data and poor result while giving the other data.

In detail on overfitting and underfitting I gone write article further.now understand this basic definition.

## 4.1 Nature of the problem

When the nature of the problem is changing the trade-off is changing.

- **the truth is wiggly and the noise is high, so the quadratic do the best**

- **the truth is smoother, so the linear model do really well**

- **the truth is wiggly and the noise is low, so the more flexible do the best**

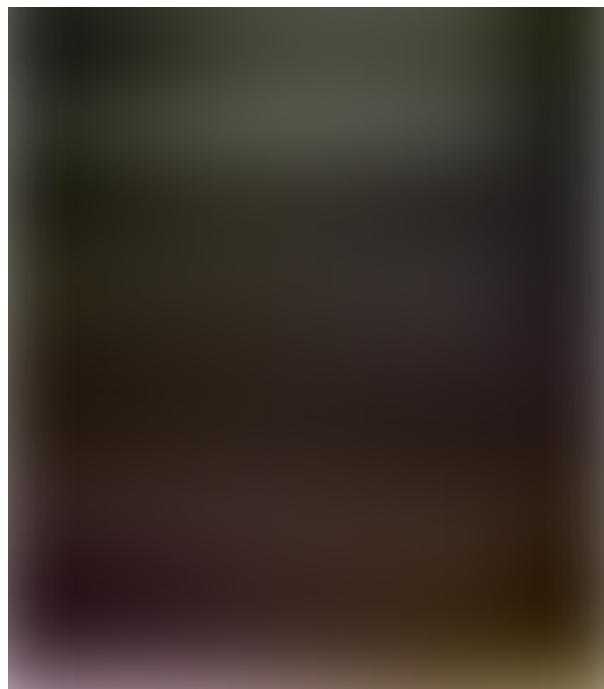


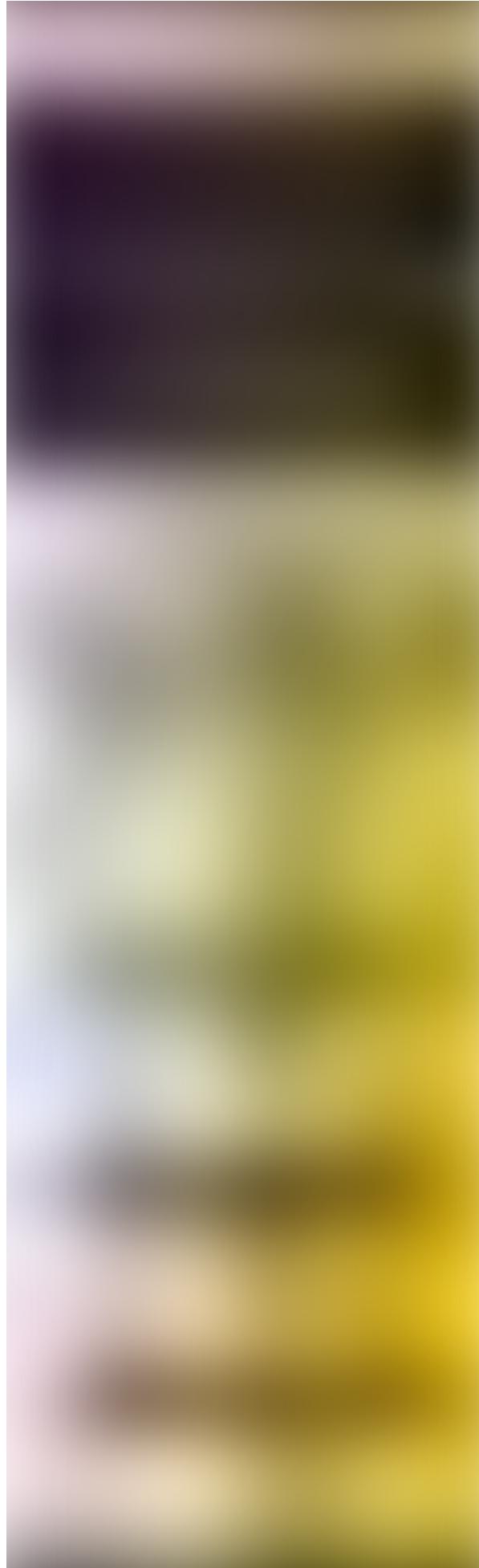
## 5. Big Picture of Bias and Variance.

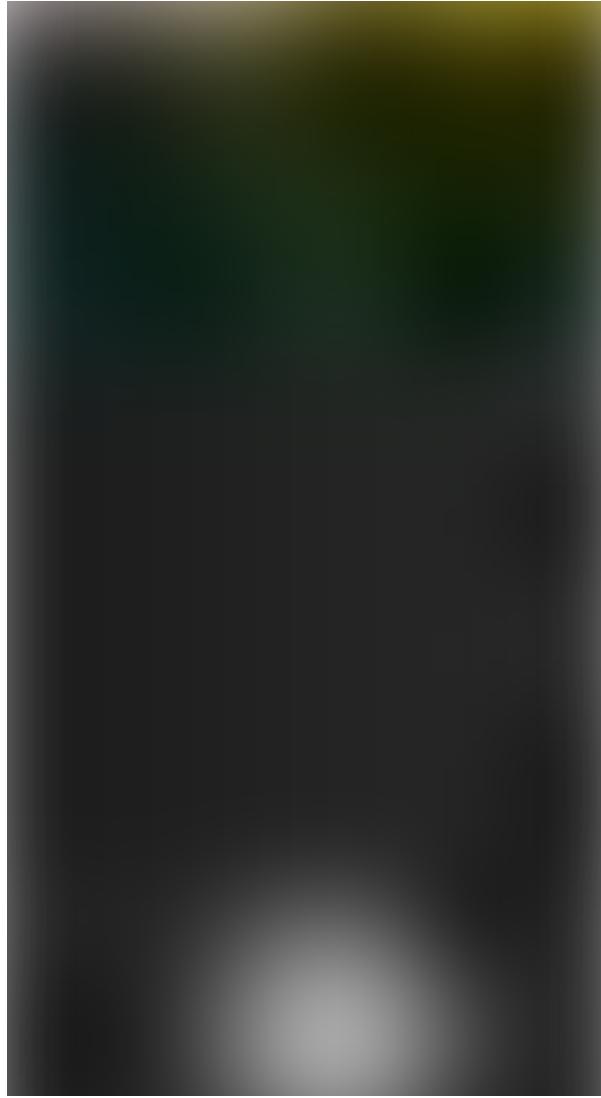
Picture said more than thousand words see here below picture and you understand what the big picture.good picture made by elite data science here attached this with this article.



Big Picture







## References:

1. <https://www.quora.com/What-is-bias-and-variance>
2. <https://www.quora.com/What-is-the-simplest-way-to-understand-bias-variance-trade-off-in-machine-learning>
3. <https://machinelearningmastery.com/gentle-introduction-to-the-bias-variance-trade-off-in-machine-learning/>
4. <https://machinelearningmastery.com/gentle-introduction-to-the-bias-variance-trade-off-in-machine-learning/>
5. [https://en.wikipedia.org/wiki/Bias%E2%80%93variance\\_tradeoff](https://en.wikipedia.org/wiki/Bias%E2%80%93variance_tradeoff)
6. [https://gerardnico.com/data\\_mining/bias\\_trade-off](https://gerardnico.com/data_mining/bias_trade-off)

[Machine Learning](#)[Statistics](#)[Bias](#)[Variables](#)[MI Research Lab](#)[About](#)   [Help](#)   [Legal](#)