Bias-Variance Trade-Of

- A machine learning model's performance is considered good based on its prediction and how well it generalizes on an independent test dataset (different from training dataset). Based on the performance of different models we choose the model which ranks highest in performance.
- Let's understand this with an example, let's say we want to predict who will do well in the Lok Sabha elections(in India) of 2019, will it BJP or Congress?
- We go to a neighbourhood and start asking people if they would vote for a BJP or a Congress. we interview 1000 people, 440 say they will vote for BJP, 400 say they will vote for Congress and 120 are undecided. Based on this data we can make a prediction that chances of BJP winning are higher than Congress (note: I'm not promoting BJP). So will these ratios be the same when I ask everyone in each state? No, because this might change when I go to other states.
- We will observe inconsistencies in the prediction. This means our model is not performing well as it cannot be used reliably to make predictions.

Bias-Variance Trade-Of

- One of the reasons for our model to performance is due to the small sample size and not having enough variation in the data. This introduces error in our prediction. Error is when the predicted value is different from the actual value.
- When we have an input x and we apply a function f on the input x to predict an output y. Difference between the actual output and predicted output is the error. Our goal with a machine learning algorithm is to generate a model which minimizes the error of the test dataset.
- Error in our model is a summation of reducible and irreducible error.

$$\text{Error} = \text{Reducible Error} + \text{Irreducible Error}$$

$$\text{Reducible Error} = \text{Bias}^2 + \text{Variance}$$

Bias-Variance Trade-Of

Irreducible Error

- Errors that cannot be reduced no matter what algorithm you apply is called an irreducible error. It is usually caused by unknown variables that may be having an influence on the output variable.

Reducible Error

- Reducible Error has two components—bias and variance.
- Presence of bias or variance causes overfitting or underfitting of data.

$$\text{Error} = \text{Reducible Error} + \text{Irreducible Error}$$

$$\text{Reducible Error} = \text{Bias}^2 + \text{Variance}$$

Bias-Variance Trade-Of

Bias Error

- It occurs when there is limited flexibility to learn from the dataset.
- Bias is the simplifying assumptions made by a model to make the target function easier to learn. Generally, parametric algorithms have a high bias making them fast to learn and easier to understand but generally less flexible. In turn, they have a lower predictive performance on complex problems that fail to meet the simplifying assumptions of the algorithms bias.
- High bias causes the algorithm to miss the relevant relationship between the input and output variable. When a model has a high bias then it implies that the model is too simple and does not capture the complexity of data thus underfitting the data.
- We can have 2 kinds of Bias errors.
  1. High Bias
  2. Low Bias

Bias-Variance Trade-Of

High Bias Error

- Where the algorithm makes more assumptions and makes the algorithm less flexible. where it gives the same output even we change the dataset a little bit.
- Example Algorithms: Linear Regression, Linear Discriminant Analysis and Logistic Regression.

Low Bias Error

- Where the algorithm makes fewer assumptions and makes the algorithm more flexible.
- Example Algorithms: Decision Trees, k-Nearest Neighbors and Support Vector Machines.

Bias-Variance Trade-Of

Variance Error

- Variance is the amount that the estimate of the target function will change if different training data was used.
- It refers to the sensitivity of the algorithm to specific sets of training data.
- A variance occurs when the model performs well on the trained dataset but does not do well on a dataset that it is not trained on, like a test dataset or validation dataset. Variance tells us how scattered is the predicted value from the actual value.
- The target function is estimated from the training data by a machine learning algorithm, so we should expect the algorithm to have some variance. Ideally, it should not change too much from one training dataset to the next, meaning that the algorithm is good at picking out the hidden underlying mapping between the inputs and the output variables.

Bias-Variance Trade-Of

Variance Error

- There are two variance errors.
  1. High variance
  2. Low Variance

High Variance

- High variance algorithms train models that are accurate on average but inconsistent. If there is a little bit change in the data the models don't work well.
- Example algorithms: Decision Trees, k-Nearest Neighbors and Support Vector Machines.

Low Variance

- Low Variance algorithms are less flexible like high bias
- Example algorithms: Linear Regression, Linear Discriminant Analysis and Logistic Regression.

Bias-Variance Trade-Of

Why is Bias Variance Tradeoff?

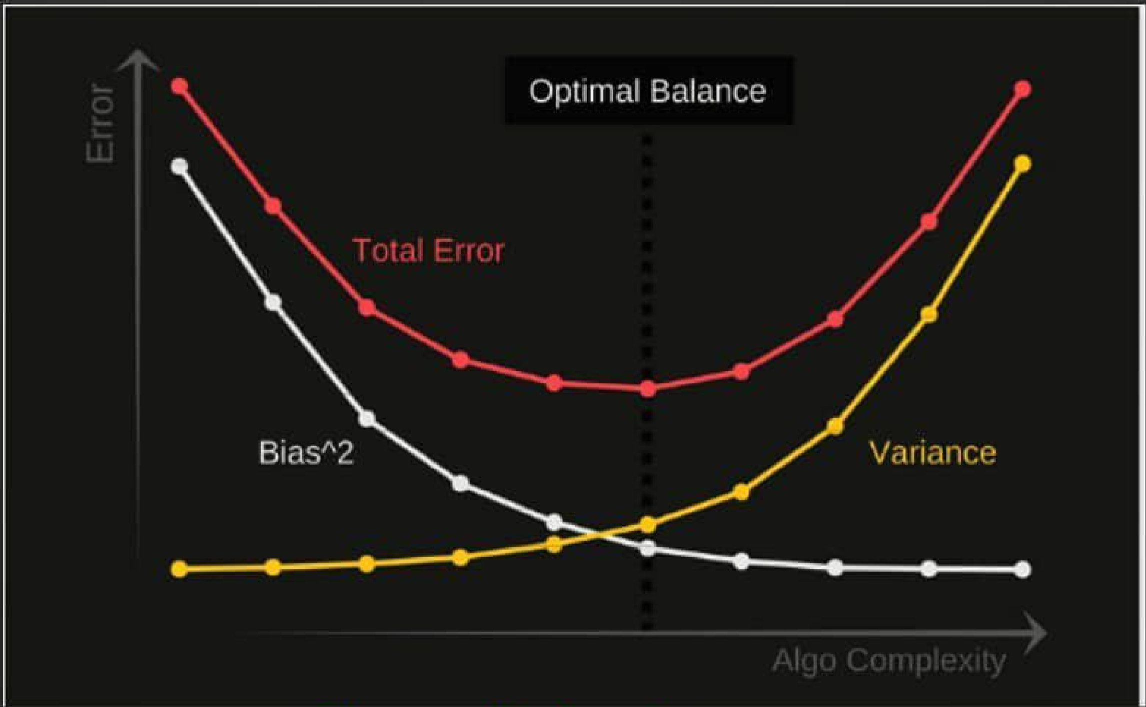
- The goal of any supervised machine learning algorithm is to achieve low bias and low variance. In turn, the algorithm should achieve good prediction performance.
- Commonly Parametric or linear machine learning algorithms often have a high bias but a low variance.
- Non-parametric or non-linear machine learning algorithms often have low bias but high variance.
- Simply The algorithms which are not complex enough(parametric models) produce underfit models that can't learn the pattern from the data.
- Where the algorithms that are too complex (non-parametric models) produce overfit models that memorize the noise instead of the pattern.
- The parameterization of machine learning algorithms is often a battle to balance out bias and variance.

Total error = Bias^2 + Variance + Irreducible error

Bias-Variance Trade-Of

Why is Bias Variance Tradeoff?

- There is no escaping the relationship between bias and variance in machine learning.
  1. Increasing the bias will decrease the variance.
  2. Increasing the variance will decrease bias.
- In reality, we cannot calculate the real bias and variance error terms because we do not know the actual underlying target function. Nevertheless, as a framework, bias and variance provide the tools to understand the behaviour of machine learning algorithms in the pursuit of predictive performance.

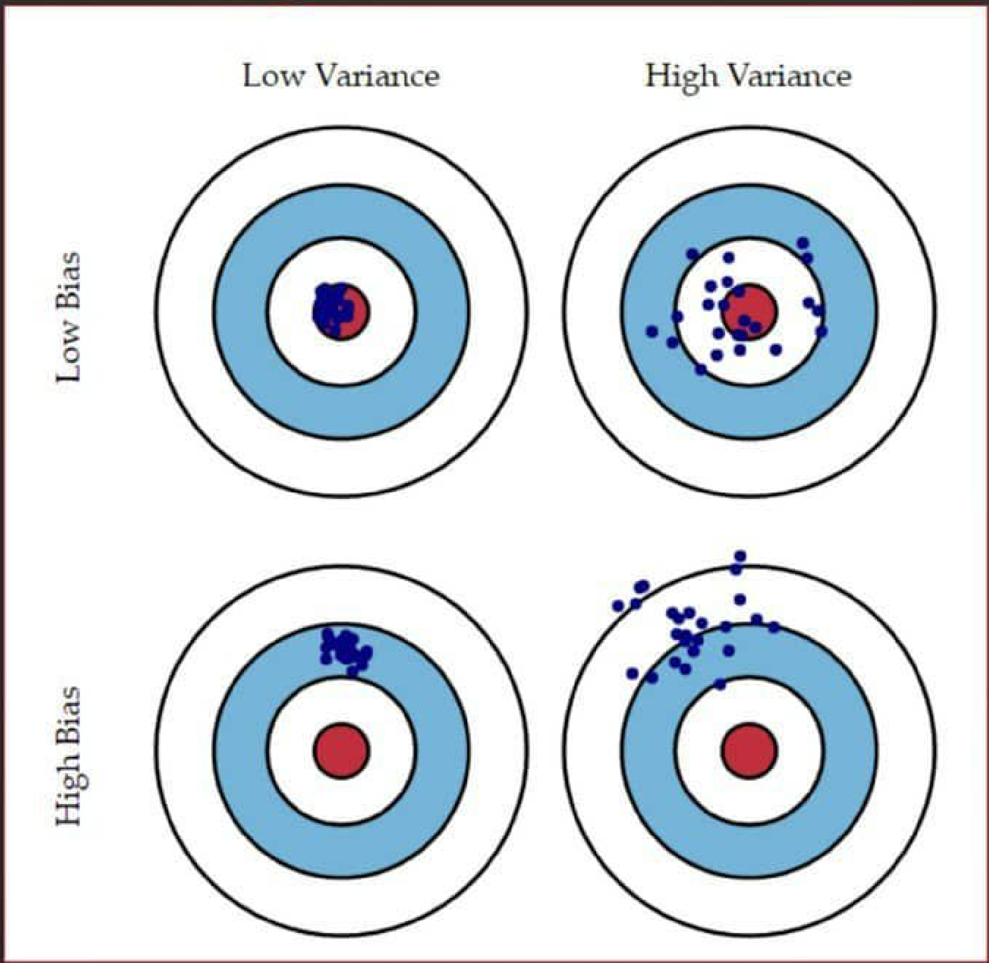


@learn.machinelearning

Bias-Variance Trade-Of

Why is Bias Variance Tradeoff?

- There are 4 types of tradeoffs
  1. Low Bias and Low Variance
  2. Low Bias and High Variance
  3. High Bias and Low Variance
  4. High Bias and High Variance



@learn.machinelearning

Bias-Variance Trade-Of

Low Bias and Low Variance

- Models are accurate and consistent on averages. We try to get to this position in our models.
- Let's take an example, You ask a few questions to a person and that person who listens to you very carefully and gives you good answers pretty much all the time.

High Bias High Variance

- Models are inaccurate and also inconsistent on average
- Let's take an example, You ask a few questions to a person and he takes wild guesses, all of which are sort of wrong.

High Bias Low Variance:

- Models are consistent but inaccurate on average
- Let's take an example, You ask a few questions to a person and he usually gives you the same answer, no matter what you ask, and is usually wrong about it.

Bias-Variance Trade-Of

Low Bias and High Variance

- Models are somewhat accurate but inconsistent on averages. A small change in the data can cause a large error.
- Let's take an example, You ask a few questions to a person and that person who listens to you and tries to answer the best they can, but that daydreams a lot and may say something totally crazy.

Bias-Variance Trade-Of

How to find when we have high bias or high variance?

- High Bias can be identified when we have
  1. High training error
  2. Validation error or test error is the same as training error
- High Variance can be identified when
  1. Low training error
  2. High validation error or high test error

Bias-Variance Trade-Of

How do we fix high bias or high variance in the data set?

Fixing High Bias

- High bias is due to a simple model and we also see a high training error. To fix that we can do the following things
  1. Add more input features
  2. Add more complexity by introducing polynomial features
  3. Decrease Regularization term (will say later in the course)

Fixing High variance

- High variance is due to a model that tries to fit most of the training dataset points and hence gets more complex. To resolve the high variance issue we need to work on
  1. Getting more training data
  2. Reduce input features
  3. Increase Regularization term
  4. Cross-Validation or early stopping