**BIAS VARIANCE TRADEOFF**

The prediction error for any machine learning algorithm can be broken down into three parts

Total Error=Reducible Error + Irreducible Error

Reducible Error=Bias^2+Variance

The irreducible error is also known as noise which is usually caused by unknown variables that may be having an influence on the output variable

Bias

Bias are the simplifying assumptions made by a model to make the target function easier to learn.

Bias is average of difference between predicted and Actual result. High Bias means We are getting low performance..

High bias is due to a simple model since model has learnt less and made more assumptions. High bias means a lot of assumptions made by the model and causes algorithm to miss relevant relationship between 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 under fitting the data. If the average predicted values are far off from the actual values, then the bias is high i.e. if model accuracy is low on a training dataset, the model is said to be under-fitting or that the model has high bias.

While low bias means less assumptions made by the model and thus model is able to capture relevant relationship between input and output variable. If the average predicted values are not far off from the actual values, then the bias is low i.e. if model accuracy is high on a training dataset, then the model has low bias.

Variance

Variance is the amount that the estimate of the target function will change if different training data was used.

High variance suggests large changes to the estimate of the target function with changes to the training dataset. Whenever a model has a high variance then the model becomes very flexible and tune itself to the data points of the training set and when this model encounters a different data point that it has not learnt then it cannot make accurate prediction. This is known as over-fitting.

Low Variance suggests small changes to the estimate of the target function with changes to the training dataset. Whenever a model has a low variance then the model doesn't become flexible and doesn't tune itself to the data points of the training set and is able to predict unseen data point accurately.

How to achieve the optimal bias variance tradeoff?

High Bias Low Variance: Models are consistent but inaccurate on average

High Bias High Variance: Models are inconsistent and also inaccurate on average

Low Bias Low Variance: Models are accurate and consistent on averages. We strive for this in our model

Low Bias High variance: Models are somewhat accurate but inconsistent on averages. A small change in the data can cause a large error.
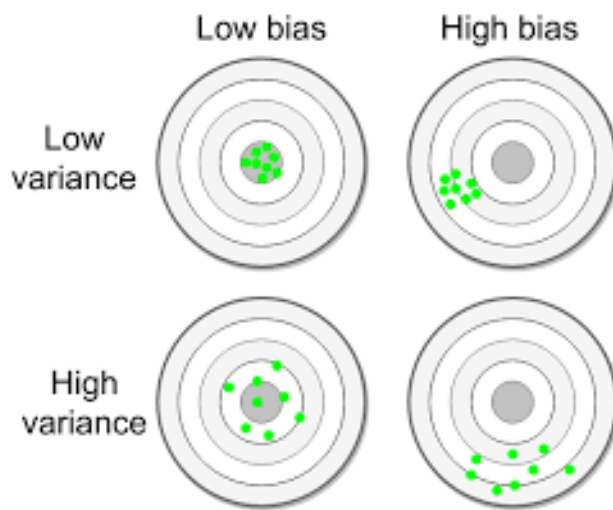


The goal of any supervised machine learning algorithm is to achieve low bias and low variance. Increasing the bias will decrease the variance. Increasing the variance will decrease the bias. So we need optimal tradeoff of bias and variance. An optimal balance of bias and variance would never over fit or under fit the model. The parameterization of machine learning algorithms is often a battle to balance out bias and variance.

High bias can be identified when we have

- High training error and Test error is same or higher than training error

High bias can be fixed by

- Add more meaningful input features

- Decrease Regularization term


High variance can be identified when

- Low training error and High test error


High variance can be fixed by

- Reduce meaningless input features

- Increase Regularization term

- Dimension reduction can eliminate noisy features, in turn reducing the model variance.

- Brining more data points to make training dataset large will also reduce variance.