Bias & Variance

What is bias?

Bias is the difference between the average prediction of our model and the correct value which we are trying to predict. Model with high bias pays very little attention to the training data and oversimplifies the model. It always leads to high error on training and test data.
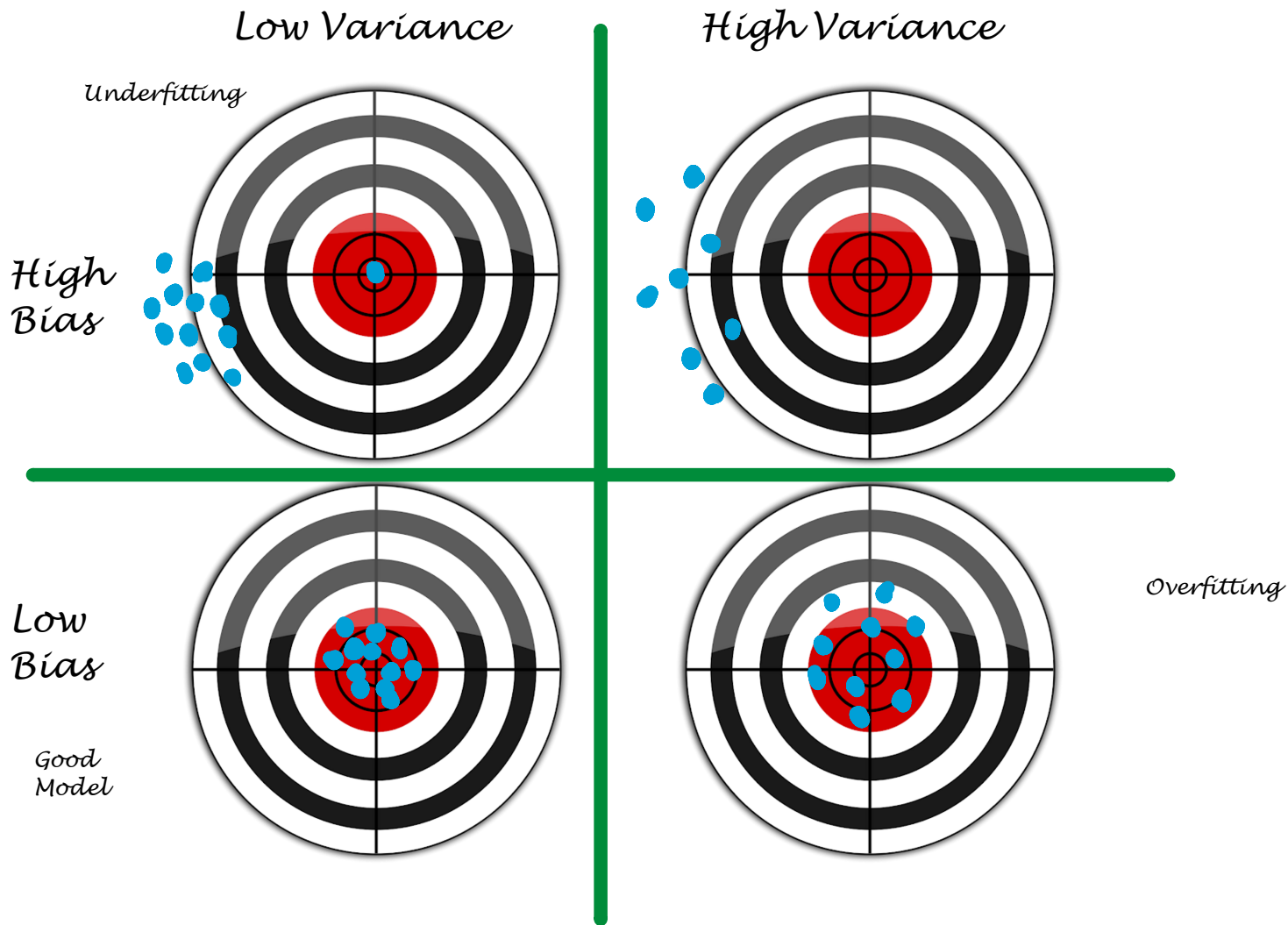
From <https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229>
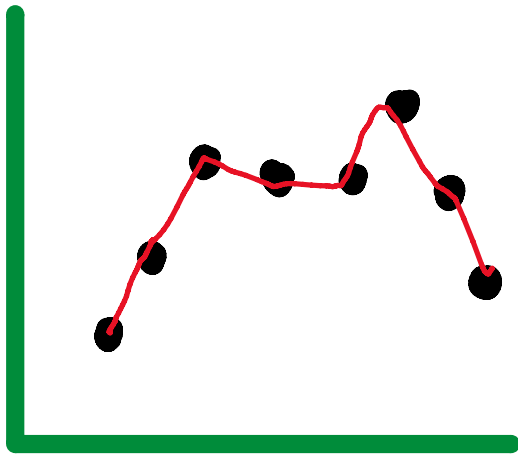
What is variance?

Variance is the variability of model prediction for a given data point or a value which tells us spread of our data. Model with high variance pays a lot of attention to training data and does not generalize on the data which it hasn't seen before. As a result, such models perform very well on training data but has high error rates on test data.

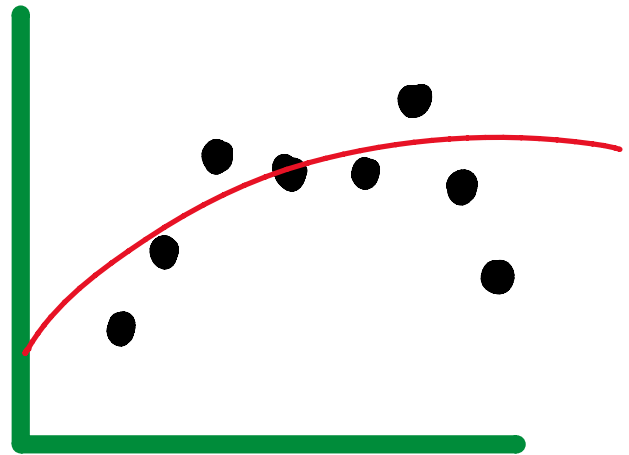From <https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229>

Bull's Eye Diagram

Low Variance

High Variance

Underfitting

High
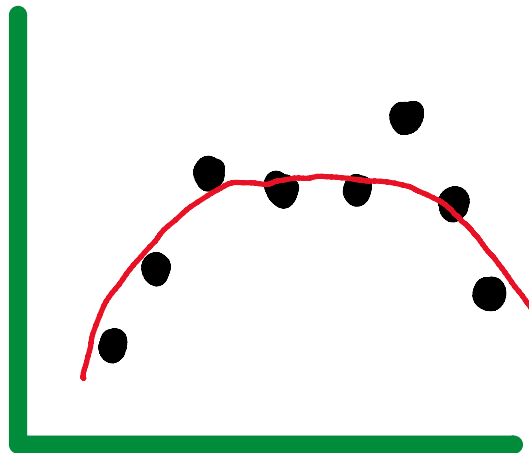Bias

Low
Bias

Good
Model

Overfitting

Overfitting & Underfitting

Overfitting

Under fitting

Good Model

What is Bias Variance Tradeoff?

If our model is too simple and has very few parameters then it may have high bias and low variance. On the other hand if our model has large number of parameters then it's going to have high variance and low bias. So we need to find the right/good balance without overfitting and underfitting the data.
This tradeoff in complexity is why there is a tradeoff between bias and variance. An algorithm can't be more complex and less complex at the same time.

Total Error = Bias^2 + Variance + Irreducible Error