Bias and Variance

FIXING FUNDAMENTALS
BY ROBOFIED
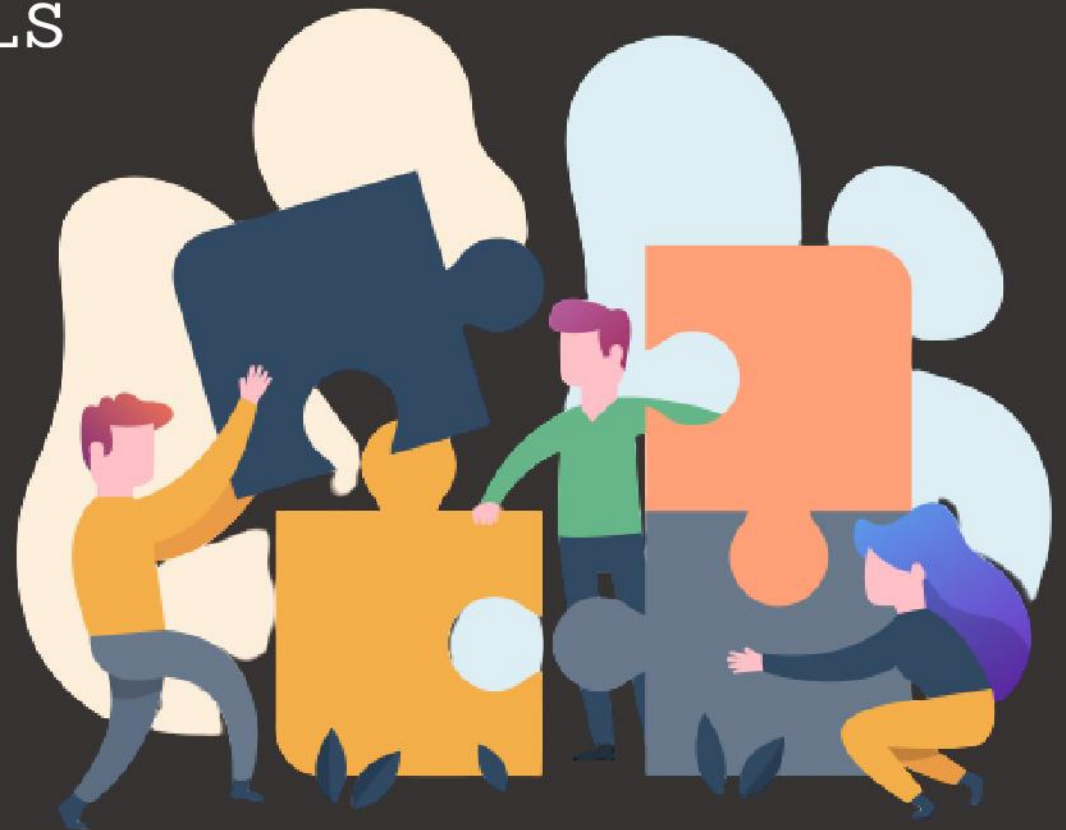
What is Bias?

Bias is the difference between the average predicted value of our model and the correct value which we are trying to predict. "Low" bias occurs when model's predicted values are closed to real values i.e. model is doing well on training set and is good enough to mimic the training data distribution. Conversely, "high" bias means the model is not fitting well on the training set.
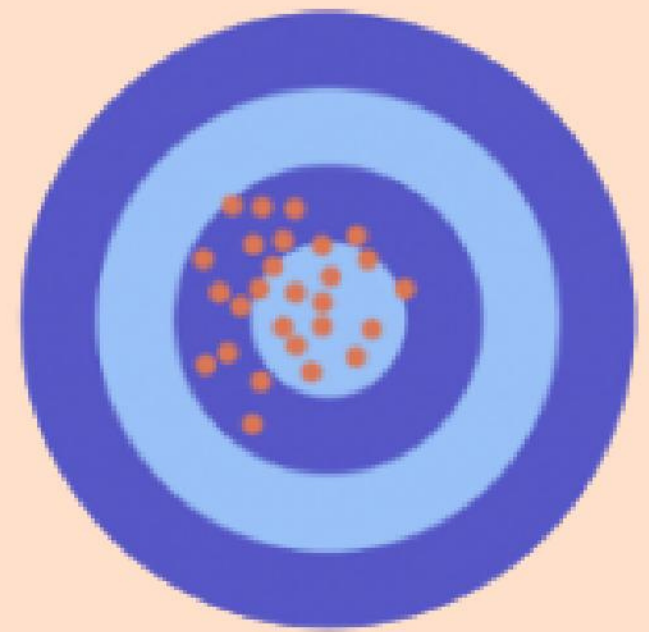
What is Variance?
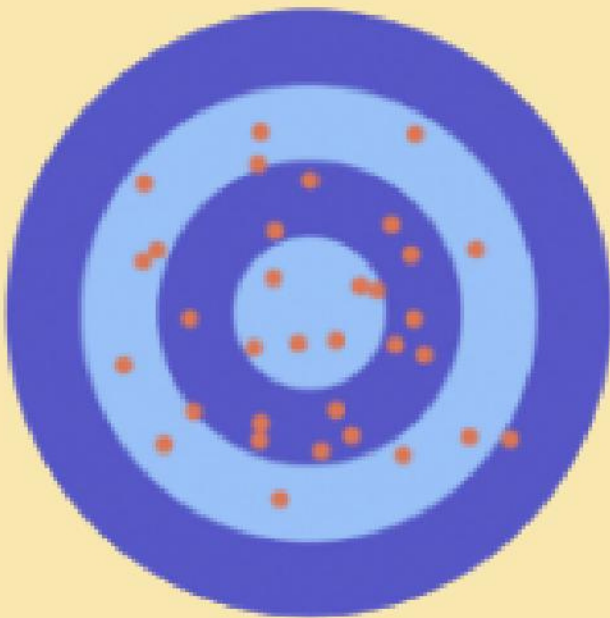
Variance is described as the variability of model prediction for a given data point/instance which tells us the spread of our data. It basically describes how much the model changes based on the changes in the inputs. If we have "low" variance, we are performing well on validation set. Conversely, if we have "high" variance, we are not performing good enough on validation set.

Bias Variance Tradeoff
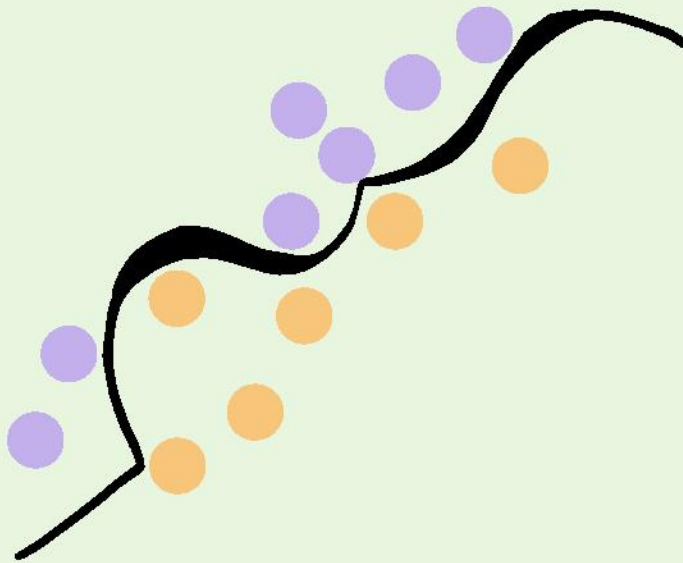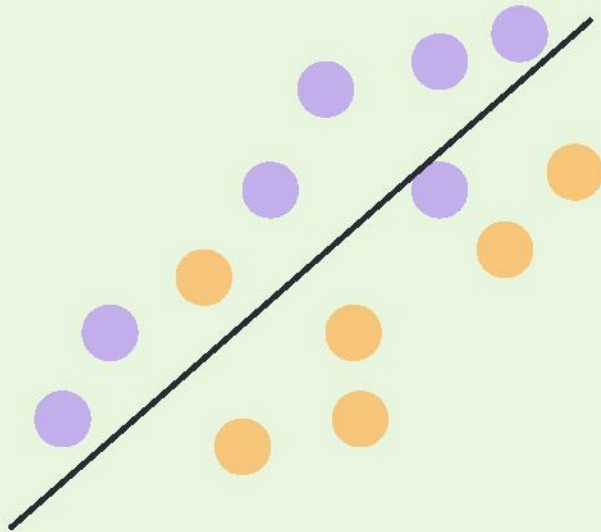
Predictive Models have a tradeoff between the bias (how well does the model fit the data) and the variance (how much does the model change based on the changes in the inputs).

Bias vs Variance Tradeoff Contd...

Simpler Models are stable (low variance) but they do not get close to the truth (high bias). Complex models are more prone to overfitting (high variance) but they are expressive enough to get close to the truth values (low bias). The best model for a given problem lies somewhere in middle (low bias and low variance).

LOW VARIANCE   HIGH VARIANCE

LOW BIAS

HGH BIAS
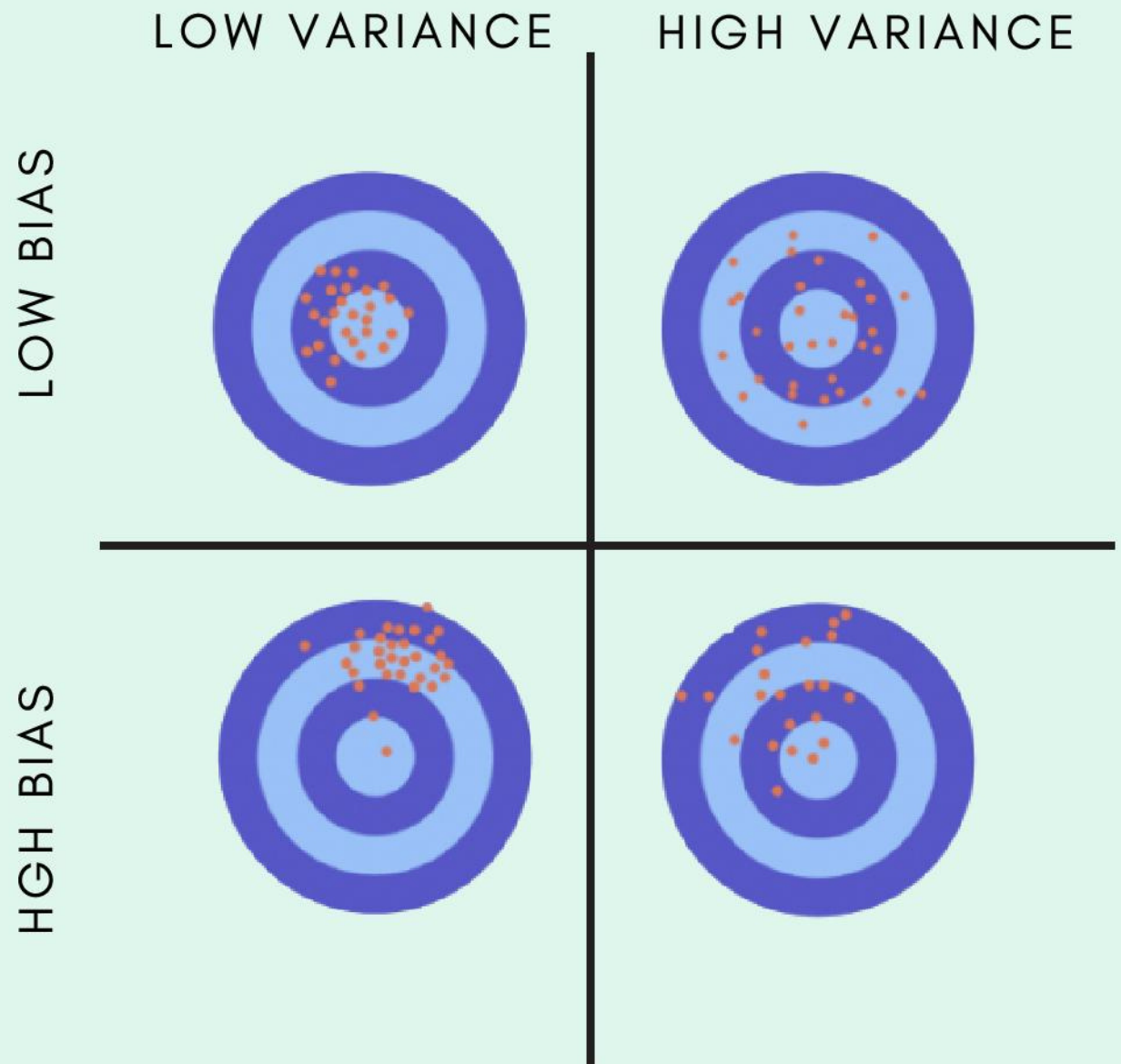
Tips from Team at Robofied

If your model is suffering from low bias and high variance, it means your model is performing poorly on unseen data. In such cases, you can try using regularization methods and can try selecting top 'n' features by looking at feature importance in predicting target variable.

Conversely, if your model is suffering from high bias and low variance problem, you should try using a more complex model so that underlying patterns from the data can be detected.

1. Machine Learning Mastery Blogs by Jason Brownlee
2. CS229 Machine Learning @Stanford by Andrew Ng
3. Hands on Machine Learning with Scikit Learn & Tensorflow 2nd Editionby Aurelien Geron