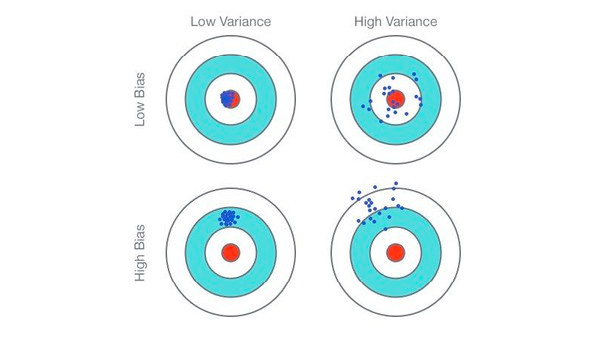
Bias and Variance

**General Definition**



**Definition in Terms of Machine Learning**

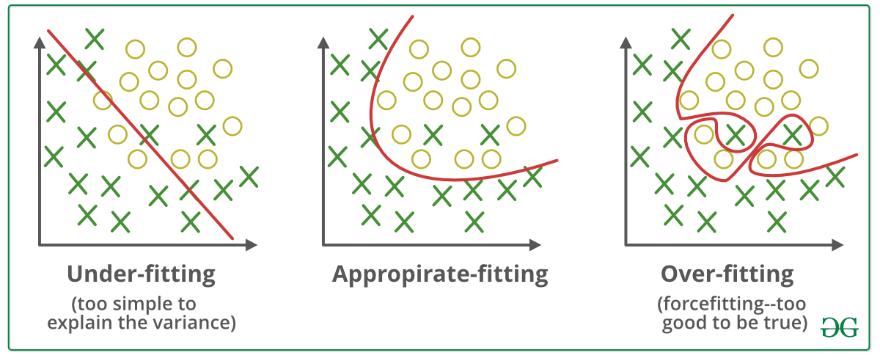
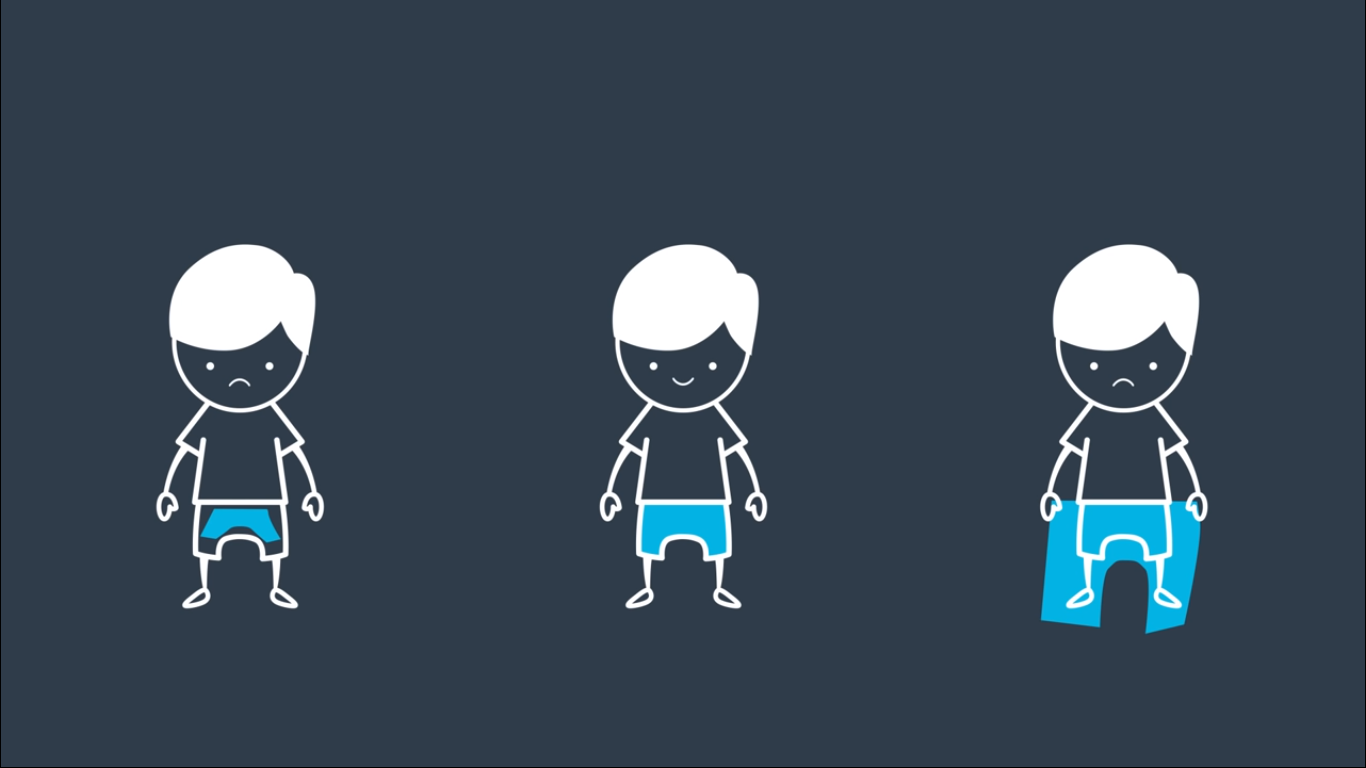
**Bias** 🡪 Amount of error on training data.

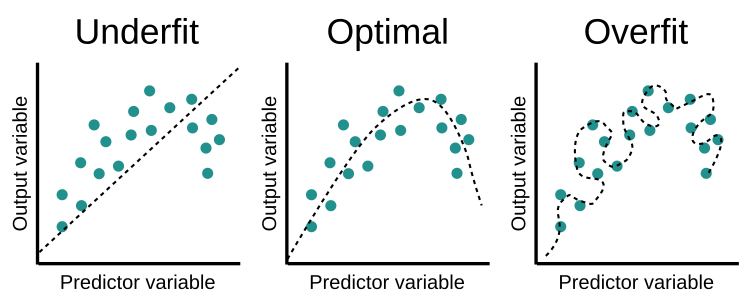
**Variance** 🡪 Amount of error on testing/run-time data.

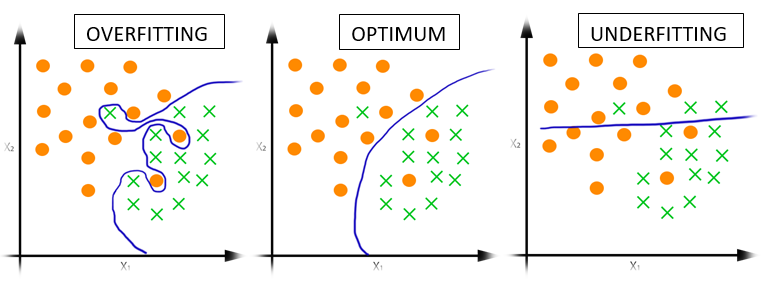
* **What we need for a better prediction model?**

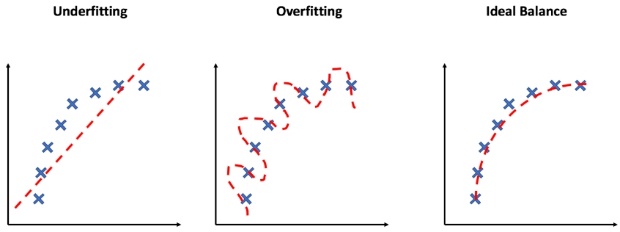
**Answer 🡪 “Low bias and low variance”**

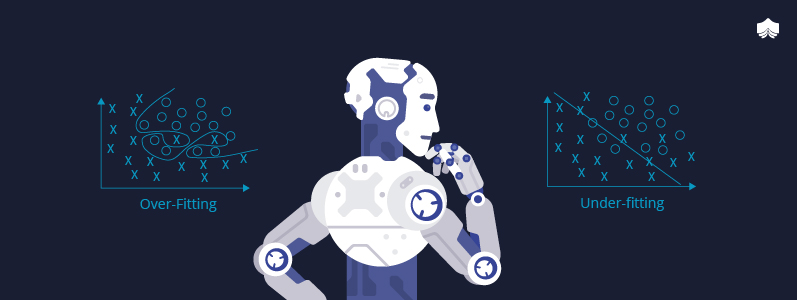
**Underfitting Model Vs Optimal Model Vs Overfitting Model**



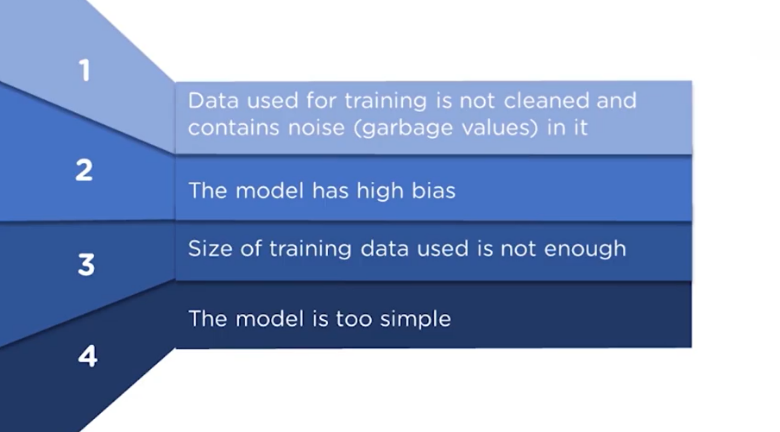




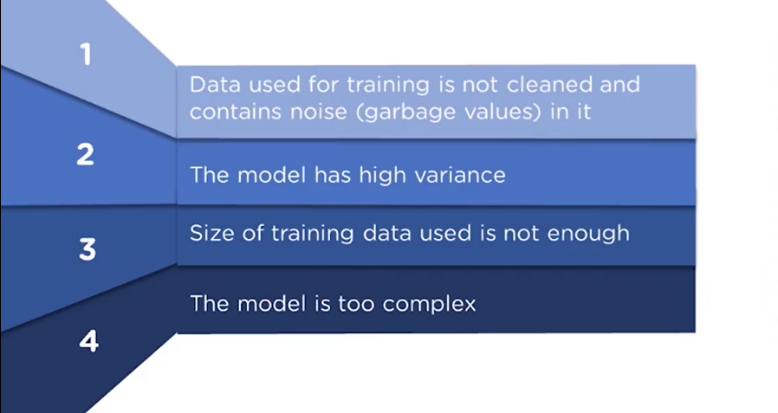


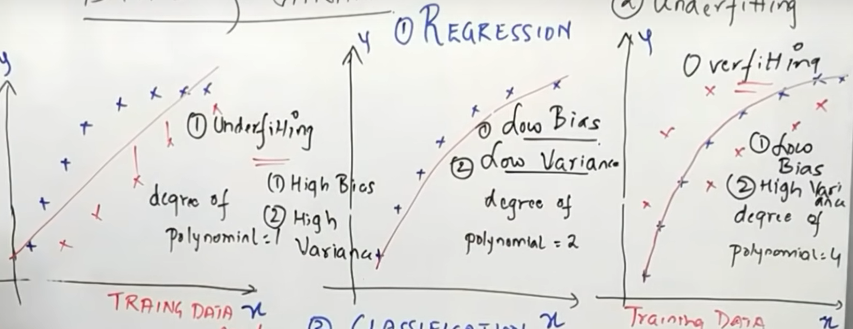


Reasons of underfitting



Reasons of Overfitting





Lower polynomial degree 🡪 low training error (high bias) + High testing error (high variance).

Higher polynomial degree 🡪 low training error (low bias) + High testing error (high variance).

Balanced polynomial degree 🡪 low training error (low bias) + low testing error (low variance).

In the above image, in first graph, we can see that when the polynomial degree was ‘1’ the model has higher training error (high bias) and it will be having higher testing/run-time error (high variance) as well. The accuracy for training as well as testing data is maximum in underfitting.

Similarly, see the third graph, when the degree was higher that is when the polynomial degree was ‘4’, it has lower training error (low bias) which means higher accuracy on training data but still this graph of overfitting shows that it will be having higher testing/run-time error(variance). Consider the new red crosses as the testing/run-time data we see that model has higher testing error (high variance) if we consider the new or testing/run-time data as red crosses. The accuracy for training data is maximum in overfitting but the accuracy on training data is still lower.

Lastly, when we see the graph in the middle, we see the polynomial degree was neither that much higher nor lower it was a balanced polynomial degree. We see that we get low training and low testing error (low bias and low variance) which means high accuracy on both training and testing data.

**Conclusion** 🡪 For a better model, always choose a balanced polynomial degree because it will give us low bias as well as low variance which can help us make our model more accurate.

Optimal Model

