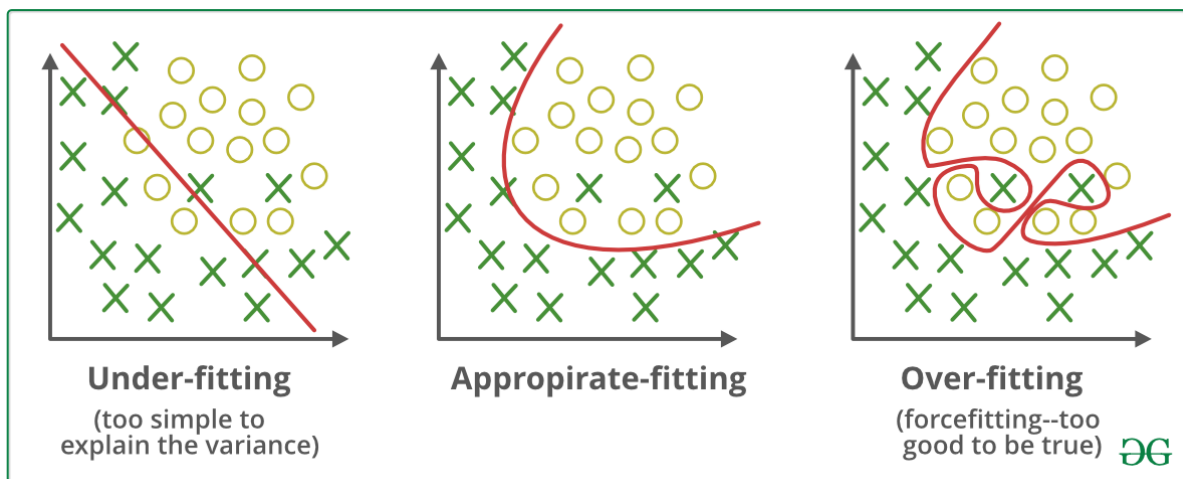


Regularization in Machine Learning

What are Overfitting and Underfitting?

Overfitting is a phenomenon that occurs when a Machine Learning model is constrained to the training set and not able to perform well on unseen data. That is when our model learns the noise in the training data as well. This is the case when our model memorizes the training data instead of learning the patterns in it.

Underfitting on the other hand is the case when our model is not able to learn even the basic patterns available in the dataset. In the case of the underfitting model is unable to perform well even on the training data hence we cannot expect it to perform well on the validation data. This is the case when we are supposed to increase the complexity of the model or add more features to the feature set.



What are Bias and Variance?

Bias refers to the errors which occur when we try to fit a statistical model on real-world data which does not fit perfectly well on some mathematical model. If we use a way too simplistic a model to fit the data then we are more probably face the situation of **High Bias** which refers to the case when the model is unable to learn the patterns in the data at hand and hence performs poorly.

Variance implies the error value that occurs when we try to make predictions by using data that is not previously seen by the model. There is a situation known as **high variance** that occurs when the model learns noise that is present in the data.

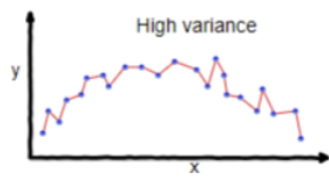
Finding a proper balance between the two that is also known as the [Bias-Variance Tradeoff](#) can help us prune the model from getting overfitted to the training data.

Regularization in Machine Learning

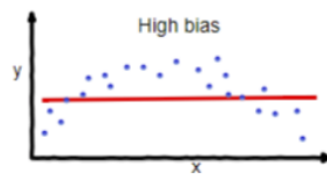
Regularization is a technique used to reduce errors by fitting the function appropriately on the given training set and avoiding overfitting. The commonly used [regularization techniques](#) are :

1. Lasso Regularization – L1 Regularization
2. Ridge Regularization – L2 Regularization

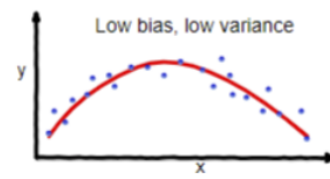
3. Elastic Net Regularization – L1 and L2 Regularization



overfitting



underfitting



Good balance