Bias-Variance Tradeoff

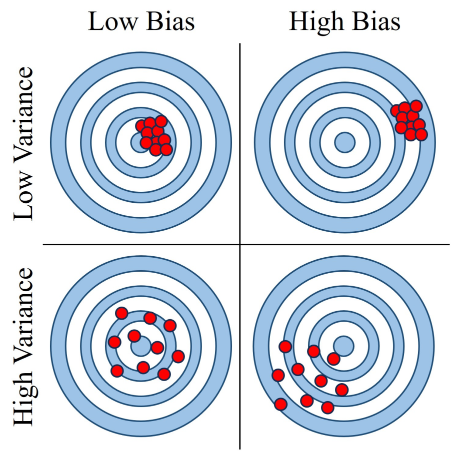
What is Bias: $bias = \mathbb{E}[f'(x)] - f(x)$

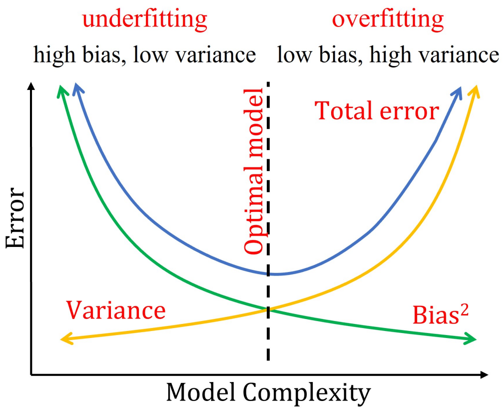
The difference between average model prediction and ground truth. The bias of the estimated function tells us the capacity of the model for correct predictions.

What is Variance: $variance = \mathbb{E}\left[(f'(x) - \mathbb{E}[f'(x)])^2\right]$

The variation in the model predictions for the given dataset.
The variance of the expected function tells the variation of the model predictions for new examples of the same distribution.

$$Error = bias^2 + variance + irreducible\,error$$





Complexity of the model and dataset

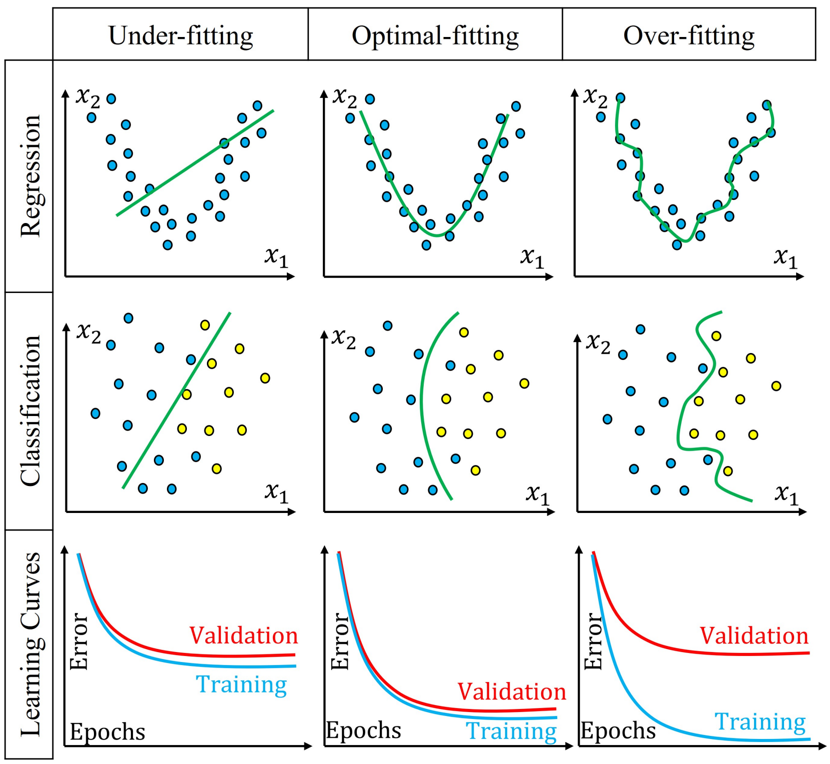
$$\text{data}^{\alpha} \Longleftrightarrow \text{model}^{\beta}$$

Underfit (high bias, low variance):
→ The model is over-simplified as compared to data complexity.
→ Large training and validation error in the learning curve. Both training and validation error curves are close to each other.

Overfit (low bias, high variance):
→ The model is over-complex as compared to data.
→ Small training error and large validation error in the learning curve and a large gap between training and validation error curves.

How to address under-fitting:
1. Use polynomial features in the dataset.
2. Remove/decrease the regularization factor. $\lambda \downarrow$.
3. Reduce noise in the data.
4. For support vector machine, use kernel trick.
5. Use deep neural networks.

How to address over-fitting:
1. Use feature selection
2. More instances in the dataset.
3. Data augmentation in the dataset.
4. Add noise to the dataset.
5. Early stop technique in training.
6. K-Fold cross-validation technique for training.
7. For decision trees, use random forest or pruning.
8. Dropout or pruning in neural networks.
9. Fewer layers in the neural network.
10. Add/Increase the regularization factor. $\lambda \uparrow$.
11. For support vector machine, (C $\downarrow$).
12. Use ensemble methods.



For any mistake/query: alishafique3@gmail.com For references and updates: Website, Github