The bias variance trade off

In machine learning, the bias–variance tradeoff is **the property of a model that the variance of the parameter estimated across samples can be reduced by increasing the bias in the estimated parameters**.

There is a tradeoff between a model's ability to minimize bias and variance which is referred to as the best solution for selecting a value of **Regularization** constant.

Proper understanding of these errors would help to avoid the over fitting and under fitting of a data set while training the algorithm.

Bias

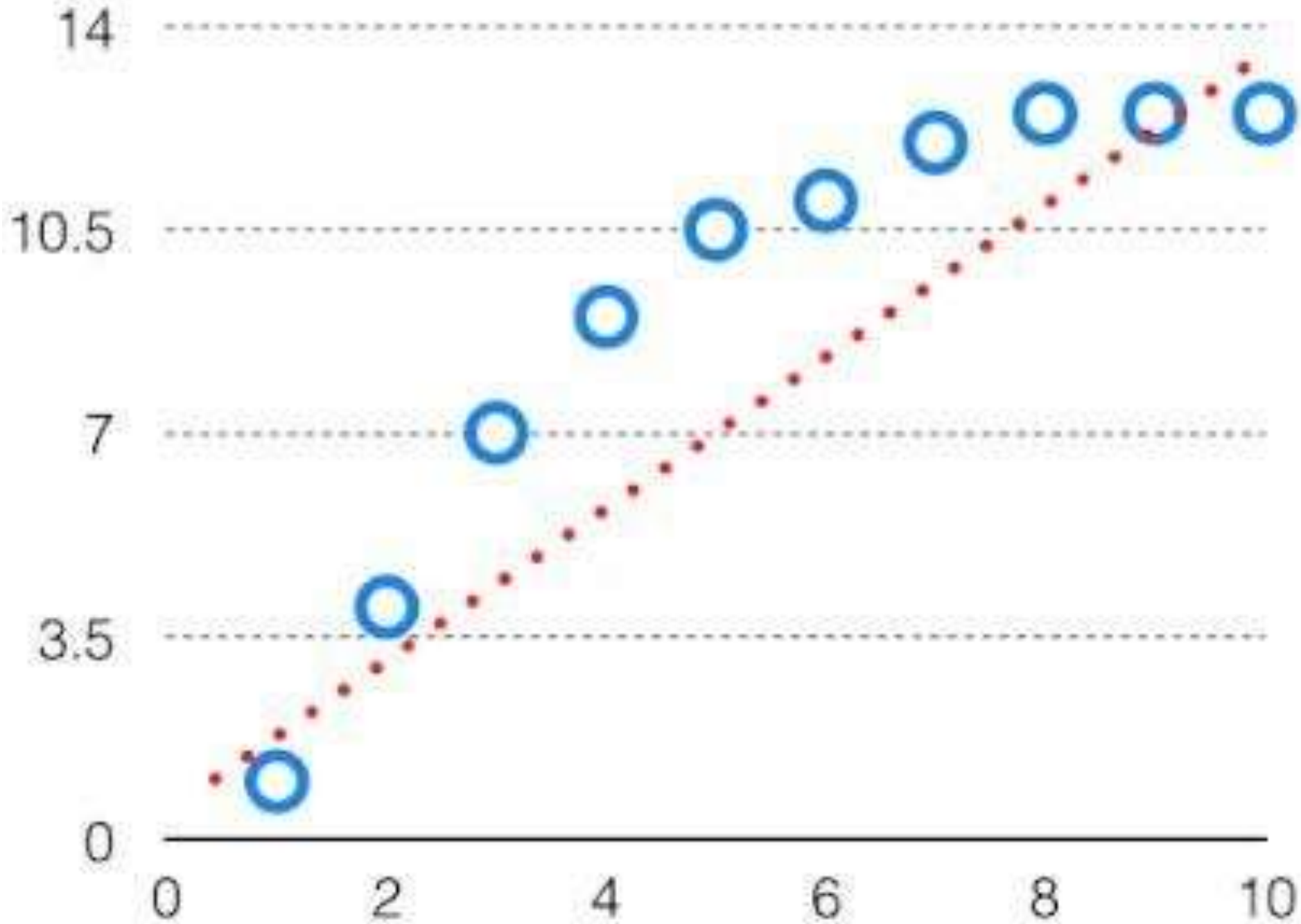
The bias is known as the difference between the prediction of the values by the ML model and the correct value.

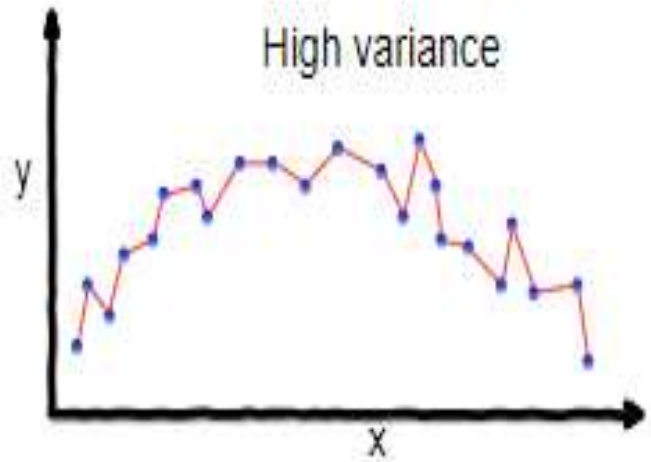
Being high in biasing gives a large error in training as well as testing data. Its recommended that an algorithm should always be low biased to avoid the problem of underfitting.

By high bias, the data predicted is in a straight line format, thus not fitting accurately in the data in the data set. Such fitting is known as **Underfitting of Data**.

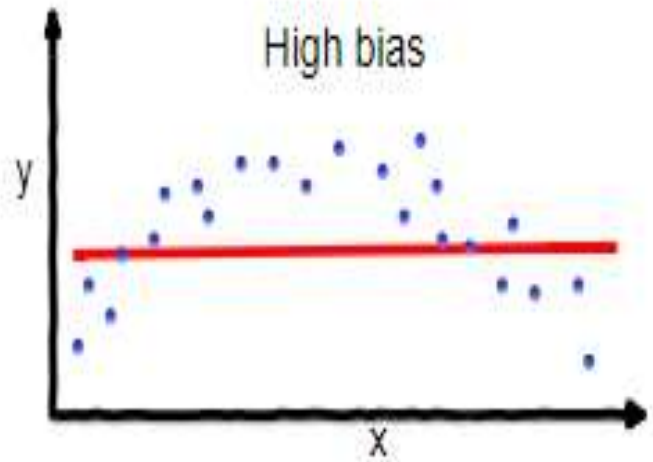
This happens when the hypothesis is too simple or linear in nature. Refer to the graph given below for an example of such a situation.

In such a problem, a hypothesis looks like follows.

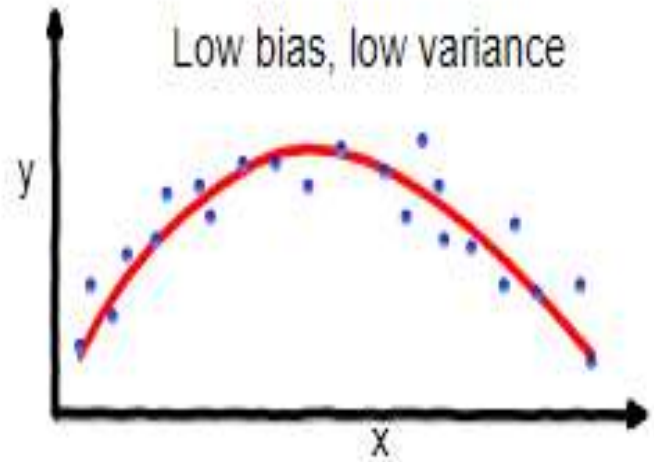

14
10.5
7
3.5
0
0
2
4
6
8
10



overfitting



underfitting



Good balance

Variance

The variability of model prediction for a given data point which tells us spread of our data is called the variance of the model.

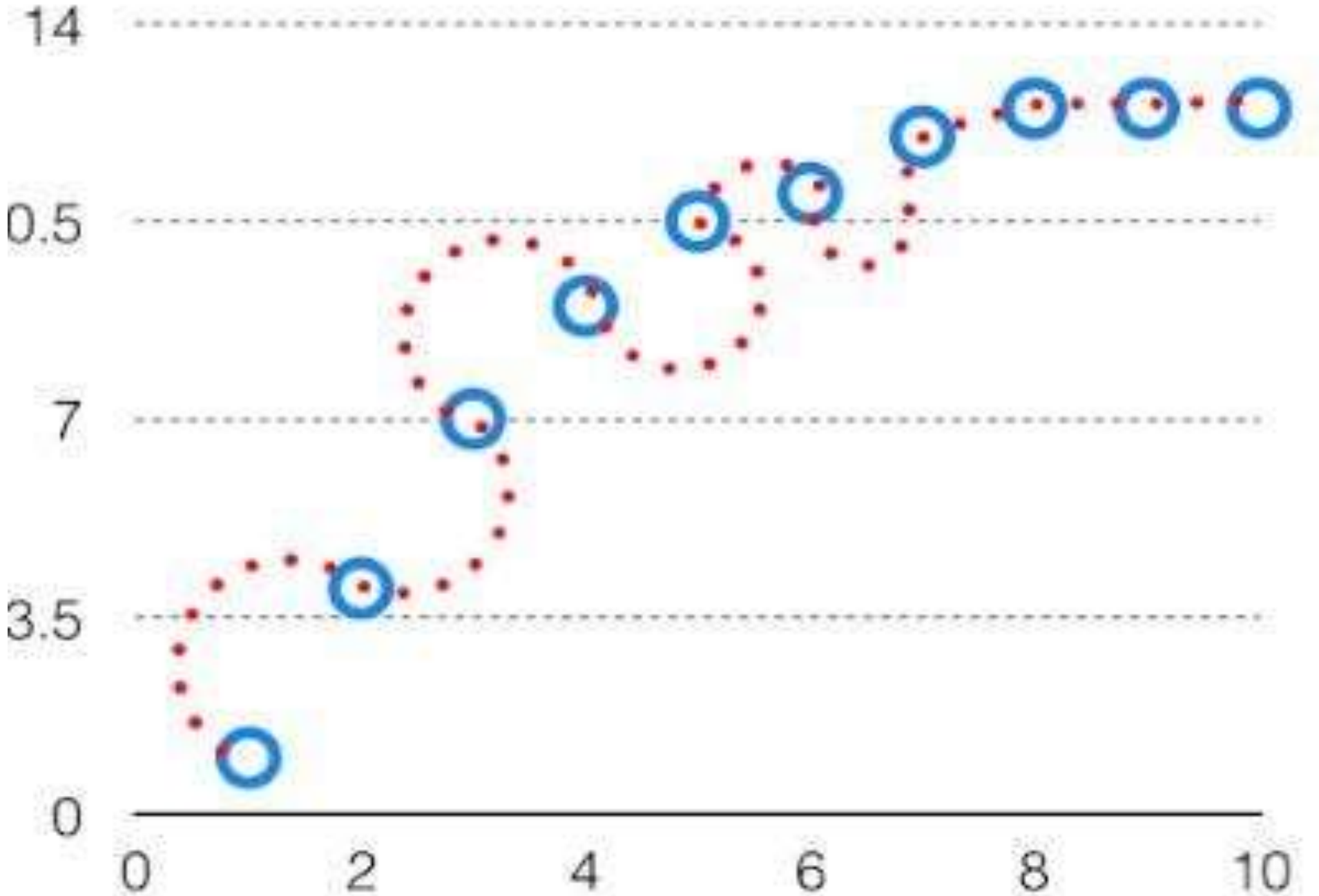
The model with high variance has a very complex fit to the training data and thus is not able to fit accurately on the data which it hasn’t seen before.

As a result, such models perform very well on training data but has high error rates on test data.

When a model is high on variance, it is then said to as **Overfitting of Data**.

Overfitting is fitting the training set accurately via complex curve and high order hypothesis but is not the solution as the error with unseen data is high.

While training a data model variance should be kept low.

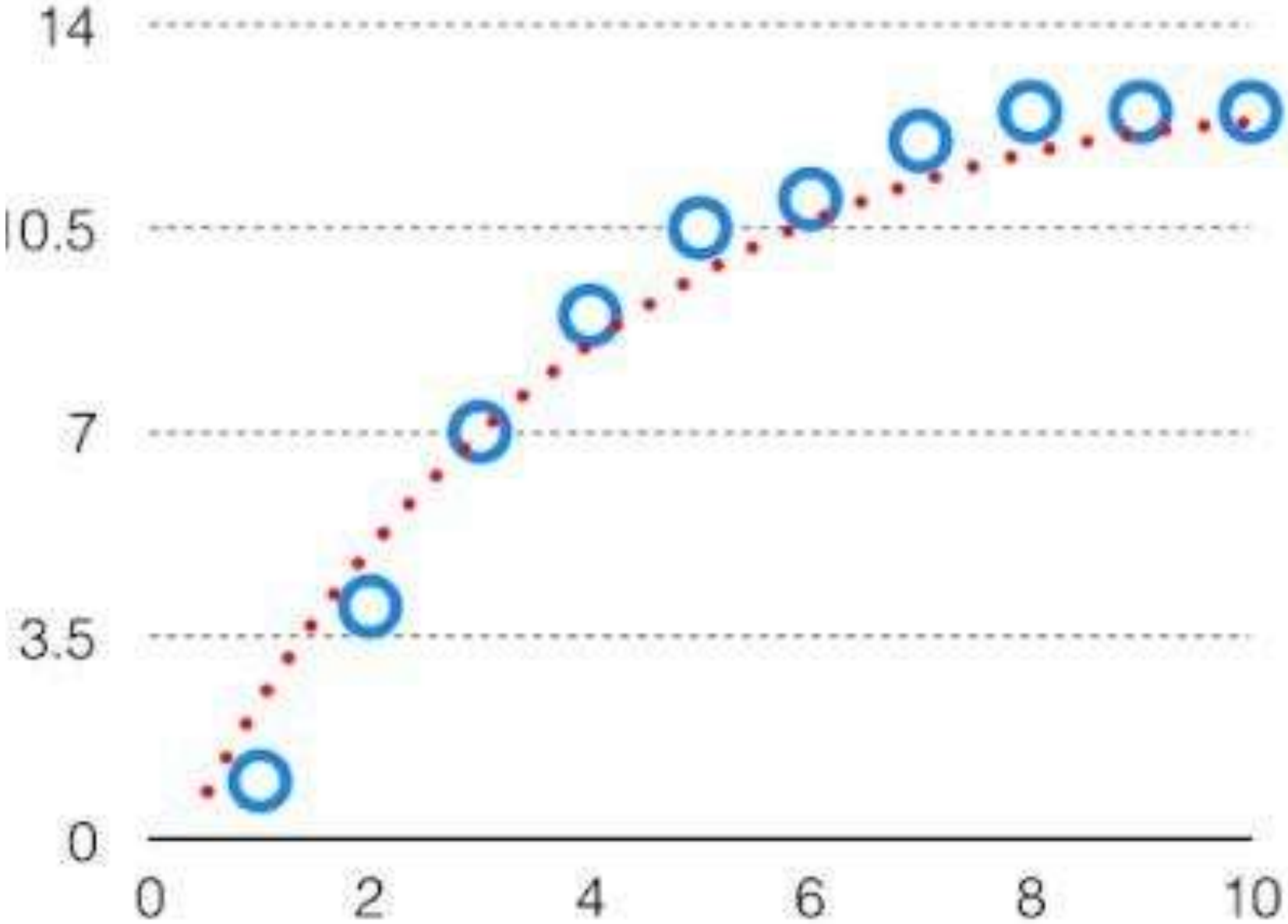
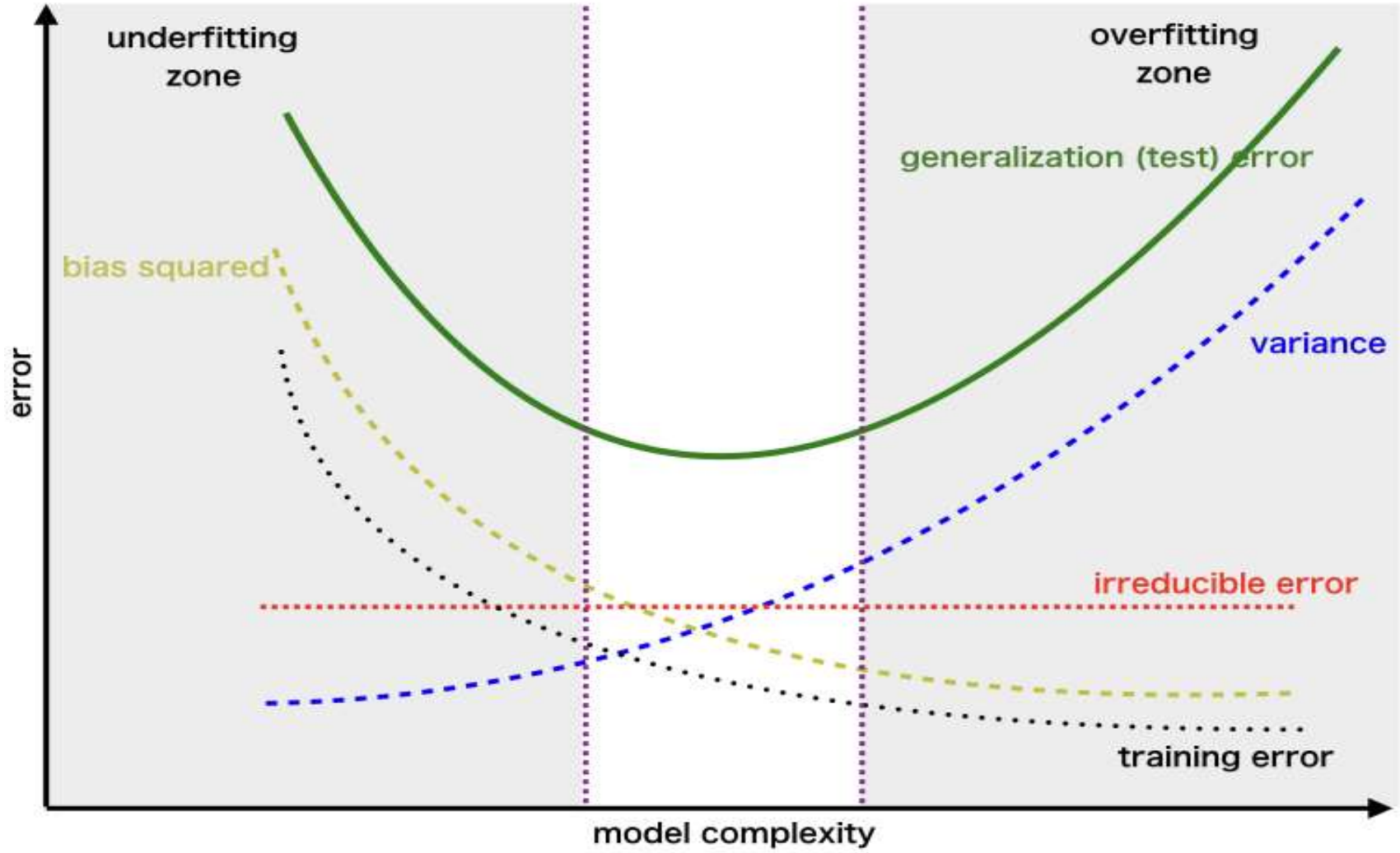

14
0.5
7
3.5
0
0
2
4
6
8
10

Bias Variance Tradeoff

If the algorithm is too simple (hypothesis with linear eq.) then it may be on high bias and low variance condition and thus is error-prone. If algorithms fit too complex (hypothesis with high degree eq.) then it may be on high variance and low bias.

In the latter condition, the new entries will not perform well. Well, there is something between both of these conditions, known as Trade-off or Bias Variance Trade-off.

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.


14
10.5
7
3.5
0
0
2
4
6
8
10

The best fit will be given by hypothesis on the tradeoff point.
The error to complexity graph to show trade-off is given as –