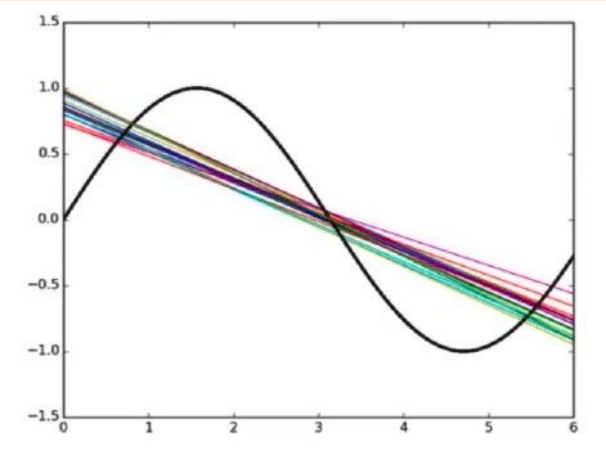
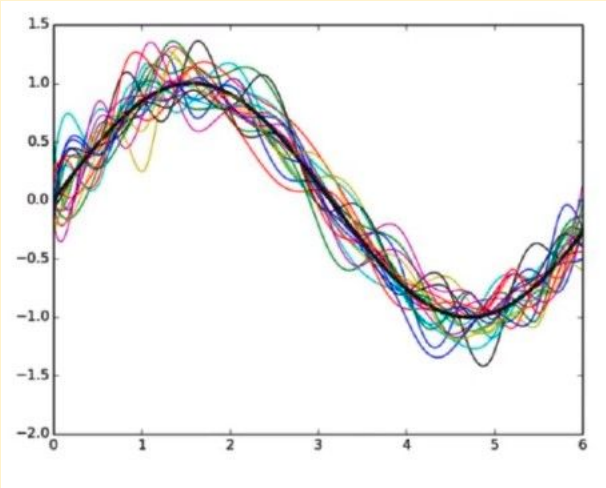
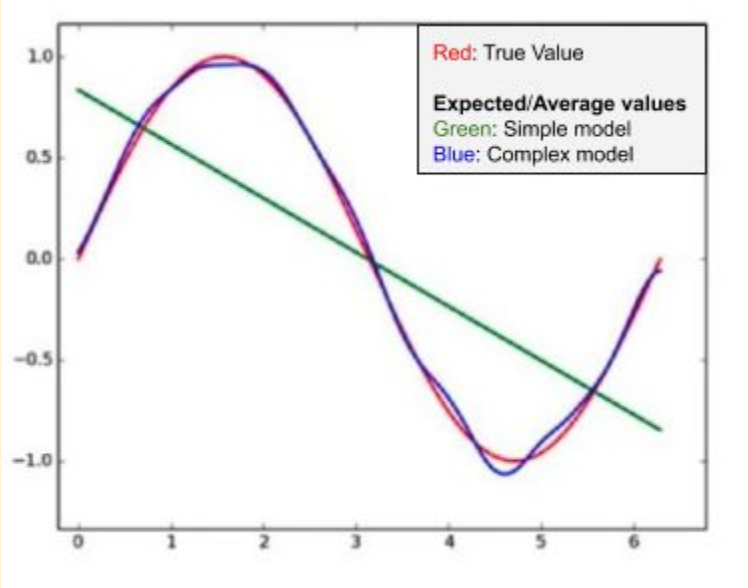
PadhAI: Regularization

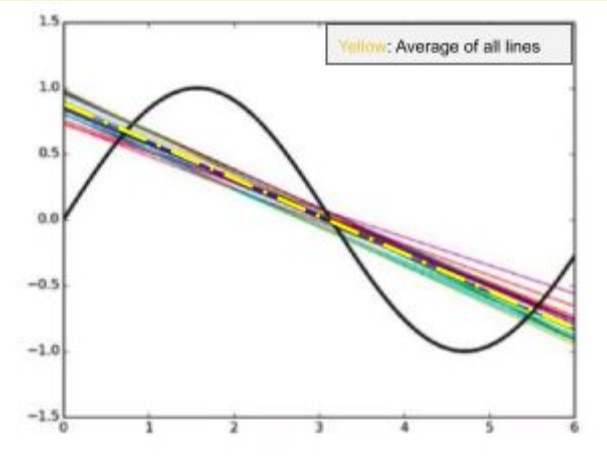
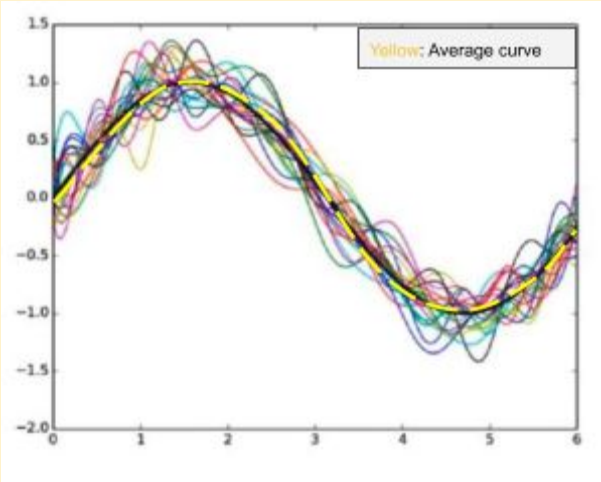
One Fourth Labs

Bias and Variance

Let's define some terms based on our observations.

1. Here is the same experiment as conducted above, except for 25 subsets instead of 3.

Simple Model	Complex Model
	
Let us define the term **Bias**: Bias $(\hat{f}(x)) = E[\hat{f}(x)] - f(x)$ Here E stands for expected value of the predictions (The average of predictions) Bias is the difference between the expectation of the predicted values and the true value	
	
In the **simple model**, the Expected value is very different from the true value, leading to a **high bias**.	In the **complex model**, the Expected value is very similar to the true value, leading to a **low bias**.
Let's define another term, **Variance**: Variance $(\hat{f}(x)) = E[(\hat{f}(x) - E[\hat{f}(x)])^2]$ As before, E stands for expected value, which is nothing but the average value of the points	

First, we calculate the square error between the predicted points and the prediction's average Then we take the average/expected value of the square error term	
	
In the **simple model**, the average line is very similar to the other lines. The lines all predict very similar values. Thus, the square error between the lines and the average line will be small, thereby its expected value will also be small. This corresponds to a **low variance.**	In the **complex model**, the average curve is quite different from the other curves. The curves predict noticeably different values. Thus, the square error between the curves and the average lincurvee will be large, thereby its expected value will also be large. This corresponds to a **high variance.**

2. The following observations can be made
   a. **Simple Model**: high bias, low variance
   b. **Complex Model**: low bias, high variance
   c. **Ideal Model**: low bias, low variance.