1. Draw a function (not necessarily a straight line!) y = f(x) that "fits" the given data points for each of the plots below.
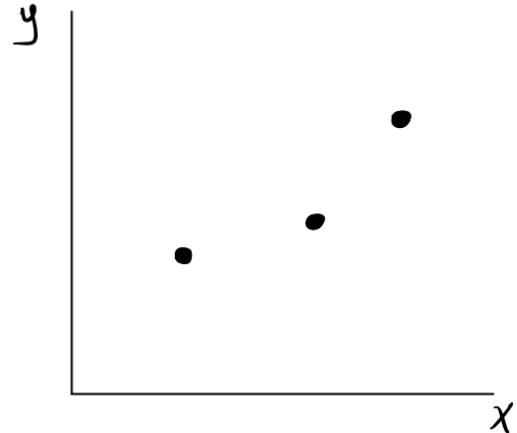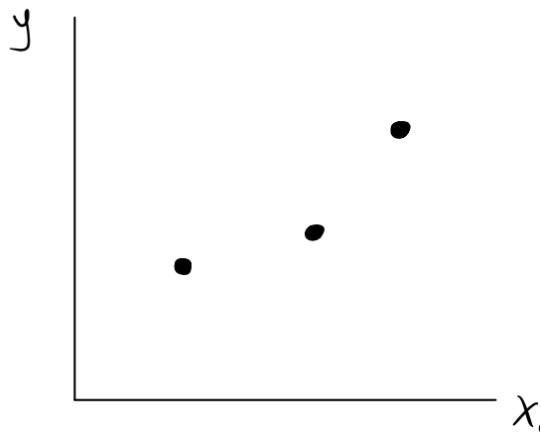
2. Draw a function y = f(x) on the plot below (again, not necessarily a line) that **overfits** the training data depicted below.

• training point

x unseen point
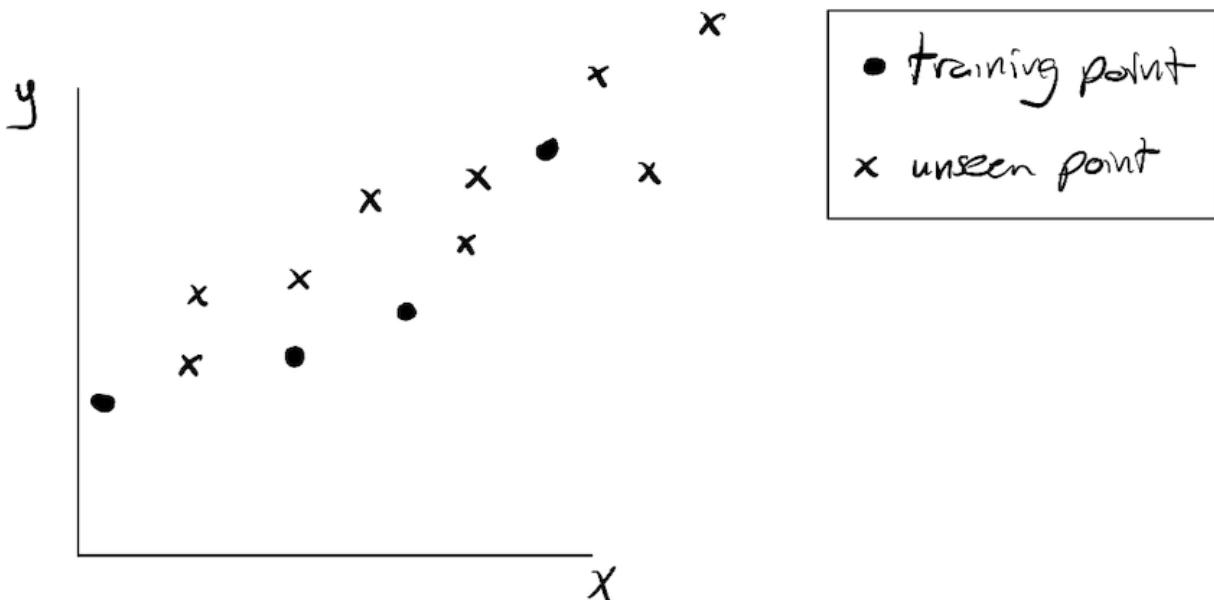
3. Using a dashed line in the plot above, draw a function y = f(x) that **underfits** the training data.

4. Consider the "your model" (dotted) lines in the plots below. In each plot, determine whether the model is "bad" because it is high-bias or because it is high-variance.

...... your model

a better model, presumably
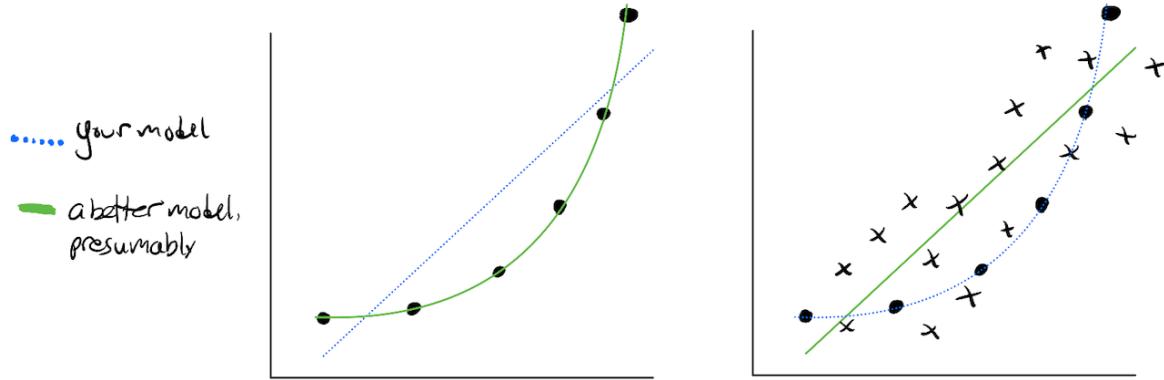
5. Suppose you split your data into training and validation sets, train a model, and evaluate its accuracy on both the training set and on the validation set. How would you describe each of the four possible scenarios? Focus on the left column first.

		Validation Accuracy	
		Bad	Good
Training Accuracy	Bad		
	Good		