Error in the Training Classifier

What is Error in the training classifier and how do we calculate it? Is there a relationship between our parameter "A" and the error "E"? All these questions will be answered in this lesson.

Let's look at the first training example: the width is 3.0 and length is 1.0 for a ladybird. If we tested the $y = Ax$ function with this example where x is 3.0, we would get,

$$y = (0.25) * (3.0) = 0.75$$

The function, with the parameter, **A** set to the initial randomly chosen value of 0.25, is suggesting that for a bug of width 3.0, the length should be 0.75. We know that's too small because the training data example tells us it must be a length of 1.0. So we have a difference, an *error*. Just as before, with the kilometers to miles predictor, we can use this error to inform how we adjust the parameter **A**.

But before we do, let's think about what y should be again. If y was 1.0, then the line goes right through the point where the ladybird sits at $(x, y) = (3.0, 1.0)$. It's a subtle point, but we don't actually want that. We want the line to go above that point. Why? Because we want all the ladybird points to be below the line, not on it. The line needs to be a dividing line between ladybirds and caterpillars, not a predictor of a bug's length given its width.

So let's try to aim for $y = 1.1$ when $x = 3.0$. It's just a small number above 1.0, We could have chosen 1.2, or even 1.3, but we don't want a larger number like 10 or 100 because that would make it more likely that the line goes above both ladybirds and caterpillars, resulting in a separator that wasn't useful at all. So the desired target is 1.1, and the error **E** is

$$Error = (desired\ target - actual\ output)$$

which is,

$$E = 1.1 - 0.75 = 0.35$$

Let's pause and have a remind ourselves what the error, the desired target, and the calculated value mean visually.