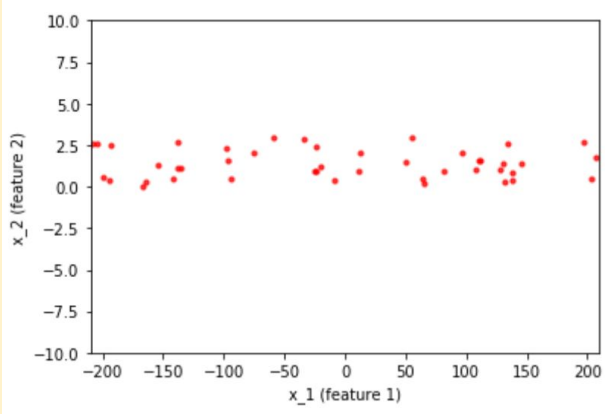
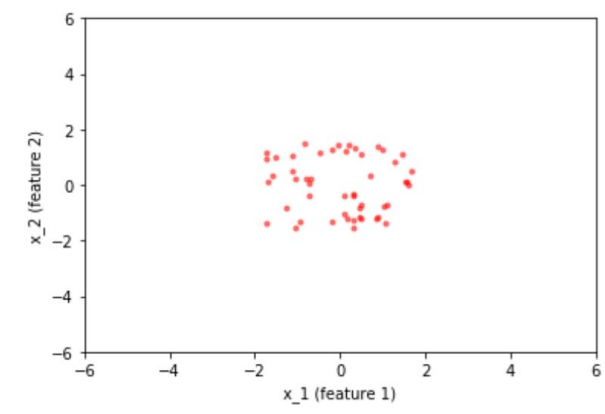


Why should we normalize inputs

Why does normalizing inputs help?

1. Let's compare the before-after of normalization of inputs

Before Normalization	After Normalization
	
Here, we observe that x_1 has a much larger range of values than x_2	Here, x_1 and x_2 have both been brought into a comparable range of values.
<p>When optimising the Loss function, one set of <u>weights corresponding to the smaller feature</u> x_2 will end up being <u>very large</u>.</p> <p>This leads to oscillations about the local minima during gradient descent as much larger updates are made due to large weights.</p>	Here, since weights end up being smaller, we see a good reduction in the oscillation about the local minima during optimization.
<p>Another point to note is that in Gradient Descent, our updates would be very biased to the <u>feature corresponding to the larger weights</u>.</p> <p>Thus we may see the model becoming very sensitive to that particular feature</p>	<p>Here, the weights are brought in a comparable range.</p> <p>This reduces the chance of updates biased to any particular feature</p>