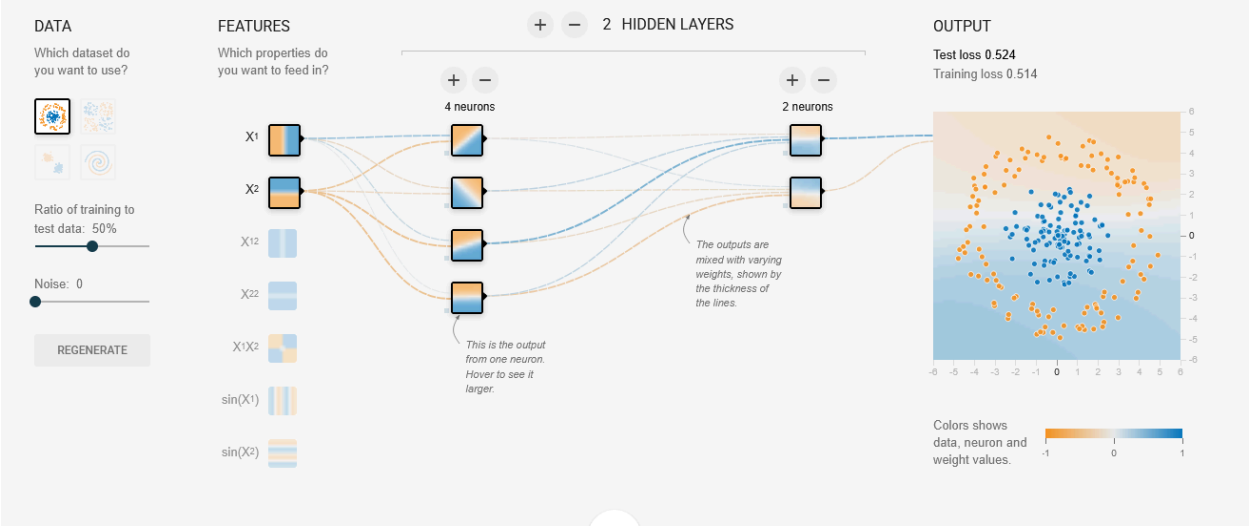
Question 1- Basic Training

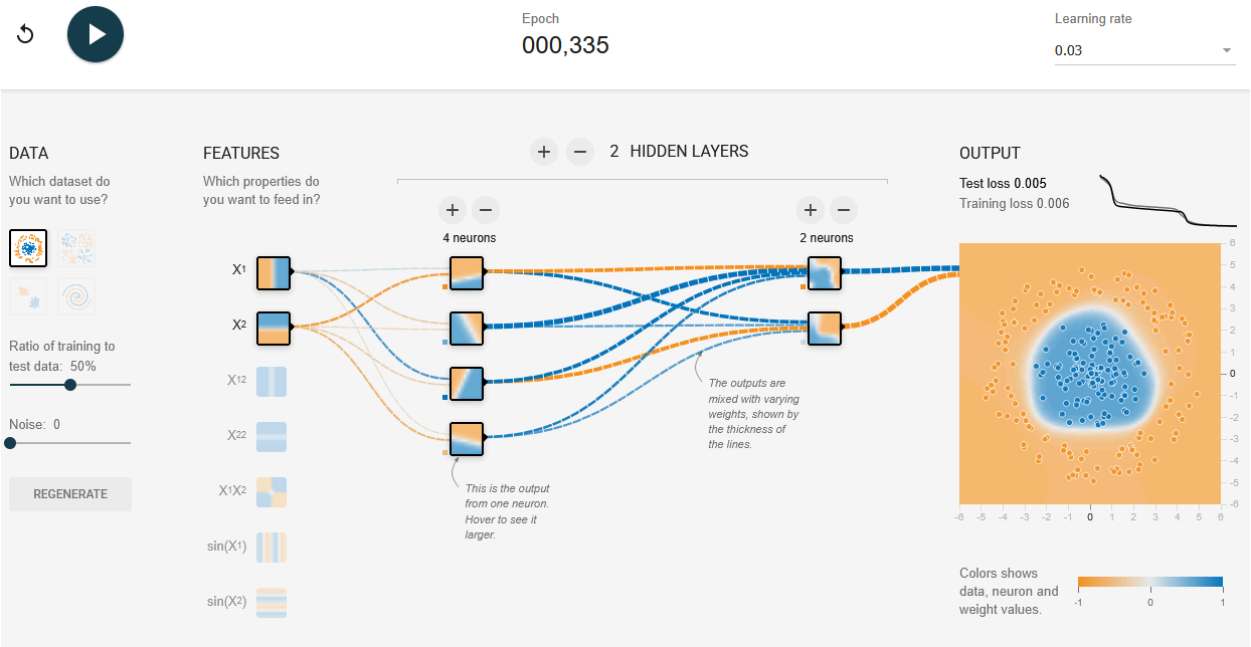


Take note of the background shading in the "Output" pane. What does it look like before training? Does it look like it will make a very accurate prediction?

Before the training has begun, the orange-blue shading is evenly split across the output pane down the middle. The learning that happens is called backpropagation . Backpropagation is a process that a network goes through. The network calculates the error within its prediction, and corrects the weighting of the calculations to reduce future errors. It works from the output layer and corrects towards the input layer, and keeps adjusting the weight according to the current loss. The orange and blue shading does not cover half of the orange and blue dots. The network looks like it won't be making very accurate predictions as the shading of the pane does not correspond with the placement of the orange and blue dots.

Take note also of the connections between nodes. These are the initial, random connections. Which neurons are most strongly connected to which other neurons?

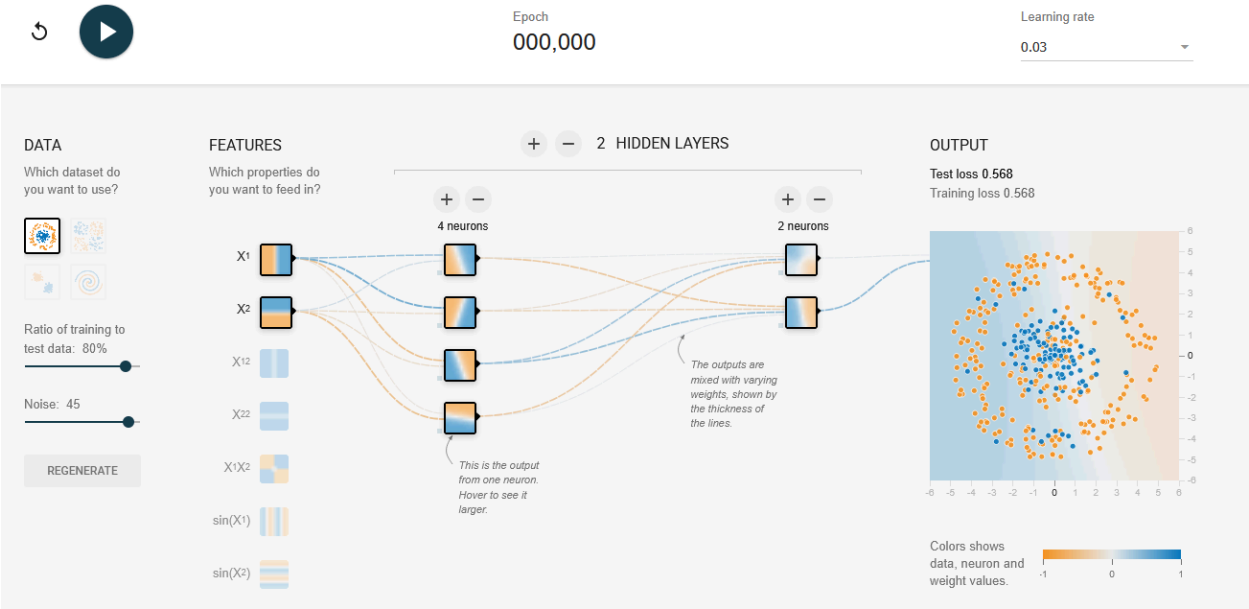
The connections of the initial nodes seem to be the strongest amongst the rest. The X^2 seems to have strong connections to the first, second, and third neurons in the first hidden layer. The X^1 seems to only have one mildly strong connection to the first neuron in the first hidden layer, and then subsequently weaker connections with the next nodes. The random connection between the third neuron in the first hidden layer, and the first neuron in the second hidden layer seems to be the strangest with it being the thickest line. The rest of the connections are weaker than this one.



How did the network change during this process? How did the predictions, the test and training loss, and the connections between neurons change? How do the patterns encoded in the hidden layer neurons change? Write your answers in a short paragraph, and attach an image of how the network looks now.

After training, the network improves significantly. The network shifted the shading, at first getting some of the blue and yellow shading correct, and after a few seconds, slowly adjusting to the final product seen on the screenshot. After 000,335 epochs, the shading in the output pane is more accurate than it was before training. The predictions now coincide with the blue and orange dots, the shading the colours correspond exactly with the area the dots occupy. The test and training loss was extremely high at first, close to 0.500 for both, and after a bit of training, it became as low as 0.005 and 0.006, and still getting lower with time. It seemed like it exponentially decreased, decreasing extremely fast with the initial training, then decreasing slower until it reached 0. The connections between the neurons all became stronger, becoming stronger than any connection before training. Usually the strongest pre-training connection between hidden layers often becomes even more dominant after training, and seems to develop the fastest amongst the rest. These favored connections likely reflect the most critical pathways for accurate predictions.

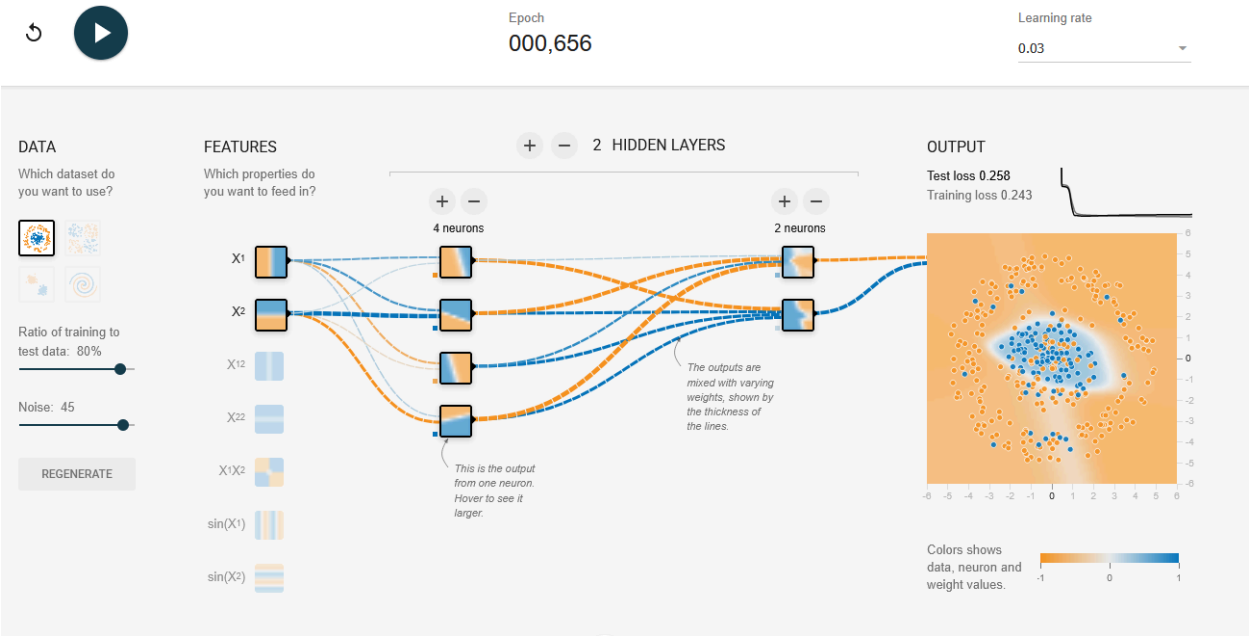
Question 2 - Background & Changing Parameters



Choose at least two of these six parameters and change them in some way. You can also change to a different data set if you would like, although this does not count as one of your two required changes. Write down which parameters you decided to change and why.

I decided to change the **test data** and the **noise** levels to see how different parameters act underneath the same network. I changed the test data from 50% to 80%, and the noise level from 0 to 45.

A prediction of what will occur is that the network will struggle to make a proper final solution as the blue points are within the orange cluster and vice versa. The end product will be unsatisfactory and will not have a low test and training loss.



What was different this time? How did your choices change how the network learned, in terms of accuracy (test and training loss), speed, or what happened to the connections between the nodes?

The things that changed this time were the time it took to get to a stable state, the amount of uncertainty and the irregularity occurred during training, the increase of time to get to a stable state, and the test and training loss result. It seemed to take a longer amount of time to get to a stable state (a state that has barely any movement). The movement that occurred to get to this state was very irregular and had a lot of erratic changes. The final result is more inaccurate than the lower level noise in question 1. The end result of the test and training loss was 0.258 and 0.243 respectively. The reason for this is because of the noise levels being high. Due to the dispersion of the dots, and some blue dots appearing in the orange dot cluster and vice versa, it would be impossible for this program to have complete accuracy. The connections between the nodes act the same (relative to the initial and final product) , with the connections strengthening over time.

What do you think of these changes? What have you learned about the importance of these parameters for a neural network, and the different things that are possible when they are set up in different ways? If nothing important really seemed to change, why do you think that is?

The change of noise and test data really changed the output of the network. The end result was inaccurate, and the development was erratic. Changing the parameters made the network struggle

a lot more to evolve and the end result was less clean. Due to the change in the test data and the noise (orange points being within the blue cluster, and vise versa), it is impossible to get a 100% accurate and 0.000 test and training loss, and due to this, the network will continuously struggle to decrease the losses and will end with an unsatisfactory result. Overall, changing any of the data will drastically change the output and the process. Changing the data will affect how much the networks will need to work to get to a stable solution, and how satisfactory the end product would be.