QUIZ 4 - Neural Network PlayGround Demo

Hoang-Dung Bui,
George Mason University
Fairfax, USA
hbui20@gmu.edu

October 28, 2023

Question 1 To begin, we will try training on the same data - Circle, but with no hidden layers, which means that this problem is a simple (linear) logistic regression objective. Use the '-' button at the top to remove all the hidden layers and try retraining. How is the performance on the "circular" dataset?
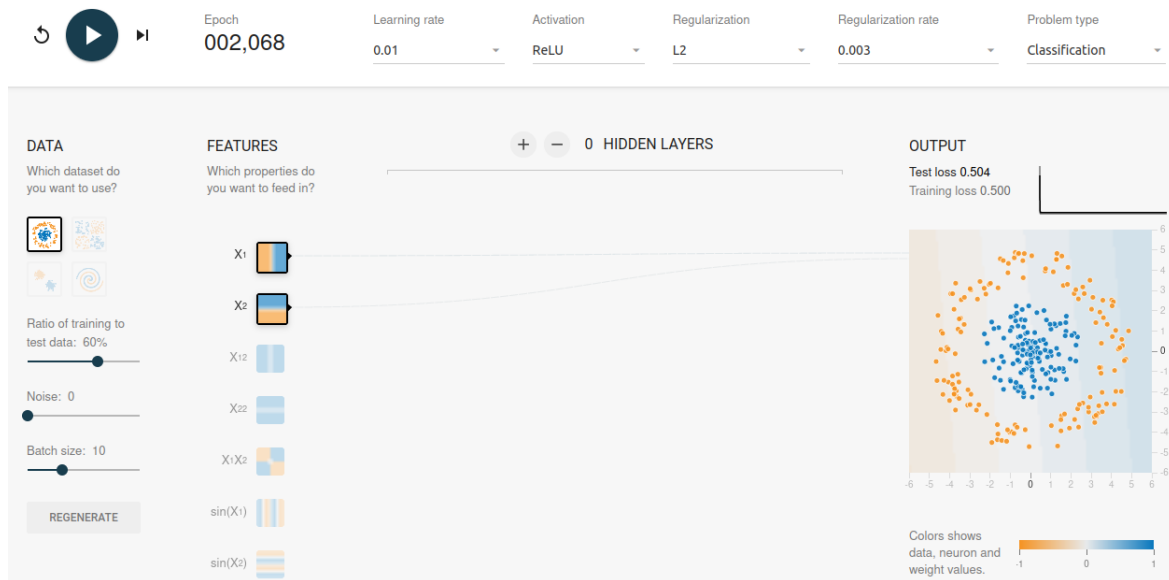

Figure 1: No hidden layers with X1 and X2 inputs

Answer It does not work. The loss is 0.504 (equivalent to random choice) after 2e3 training epoch.

Question 2 : To make this training work, we will need to do some feature engineering, a nonlinear preprocessing step that allows for nonlinear functions of the data to contribute to the learning performance. Try enabling the next two features (by clicking on them): x21 and x22. Retrain and comment on the performance: how does the system do? Include a screenshot of your Playground. (Your screenshot should contain the entire window, like Figure 1, and not just the output).

Answer: The model works much better, it can classify all the data with the loss of 0.001, meaning 0.1% correct. To improve the performance, Regularization $L1$ is used with a rate of 0.003

Question 3 : Now, let's change the data and see what we can do with a few hidden layers. Switch to using the Exclusive Or data (the one with + and - examples at alternating corners) and limit yourself to using only the two linear features: x1 and x2. Add a few hidden layers and play around with the settings, including the number of neurons. Can you get nearly perfect performance on the test data? Include a screenshot of your final configuration, showing the performance of your trained model.
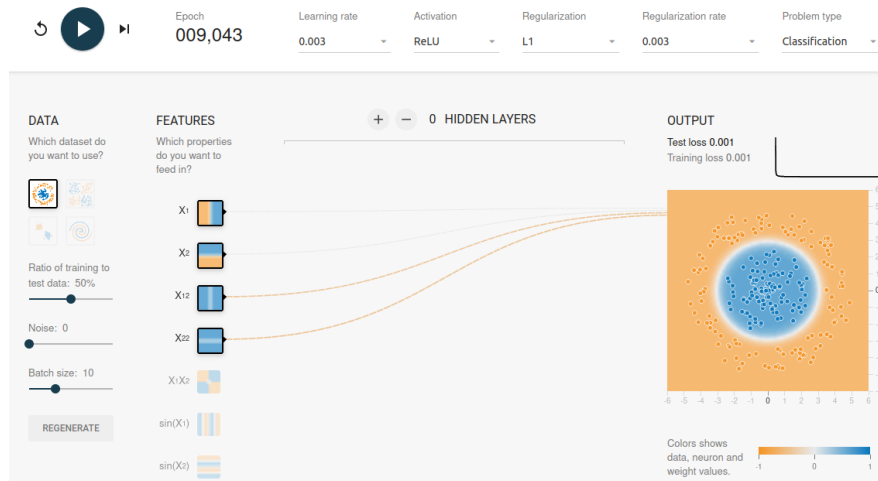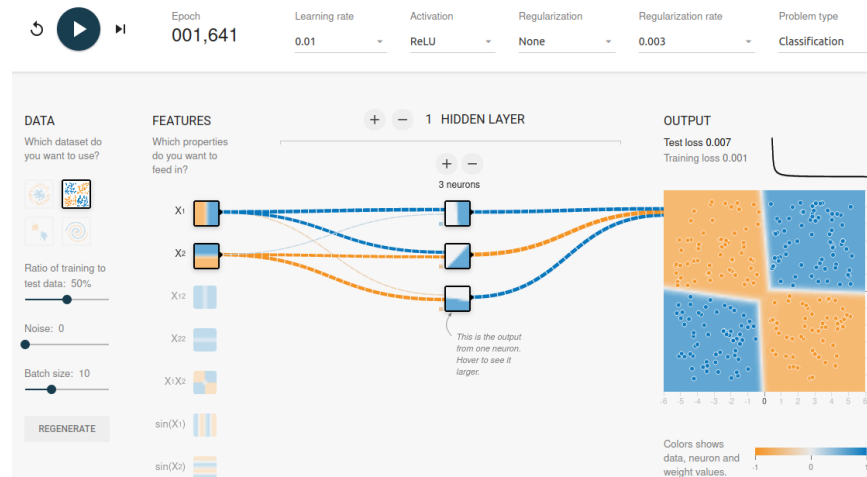
Figure 2: No hidden layers with X12 and X22 inputs

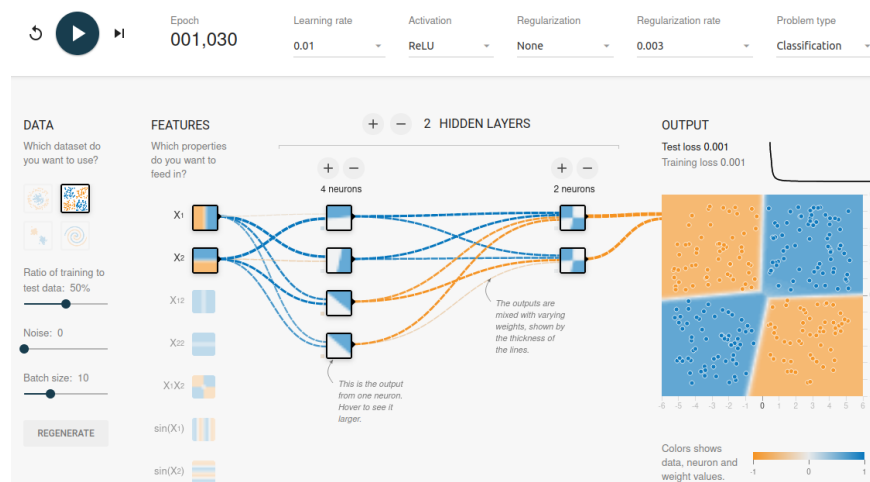Figure 3: 1 hidden layer with Exclusive OR data

Figure 4: 2 hidden layers with Exclusive OR data

Answer: For Exclusive OR data, we have tested with 1, 2, and 3 hidden layers. The more hidden layers the network has, the better results it performs. With 3 hidden layers, the network can generalize exactly the data structure, and perform 100 % correctness.
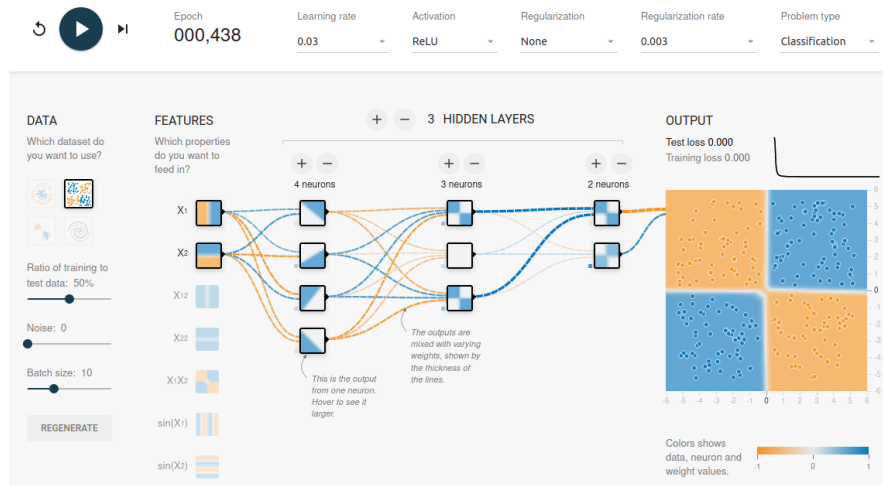
Figure 5: 3 hidden layers with Exclusive OR data

Question 4 Finally, switch to using the spiral data and try to recover good (perfect is not required) performance. Feel free to tune any of the parameters you'd like (including enabling the non-linear features). Include a screenshot of your trained system.
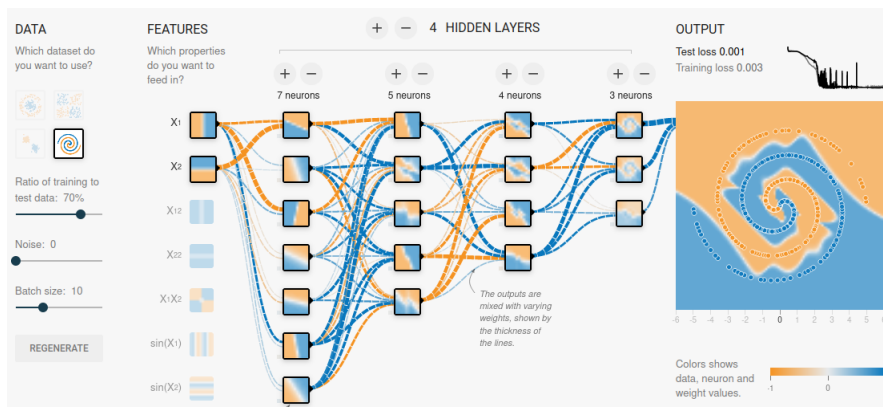

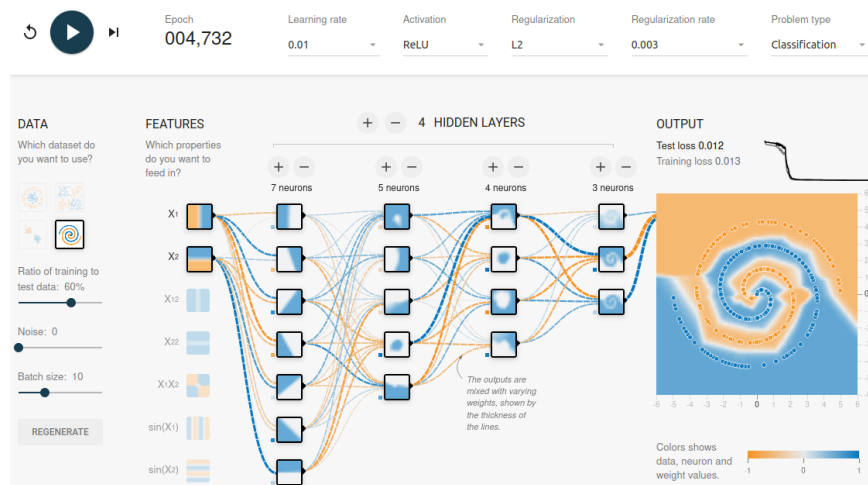Figure 6: 4 hidden layers Spiral data

Figure 7: 4 hidden layers Spiral data and Regularization

Answer: Spiral data is so difficult to be classified and the neural network needs 4 hidden layers to work

on the data. The learning rate is set to 0.1, and the loss on training data is 0.003 (Fig. 6). We tried with Regularization L2, with a rate of 0.003, the network performed even better with 100 % accurate (Fig. 7). It also generalizes better the data structure.