

 README.md

MNIST perceptron

We have perceptron designed to recognize hand-written digits. Changing hyperparameters - number of neurons in hidden layer, number of epochs, learning rate - we may impact accuracy of recognition which we test on a test sample. As an activation function Sigmoid is used. Accuracy score is calculated after testing on a test sample as number of successful recognitions divided by the test sample size.

The results of changing hyperparameters are summarized in the table below:

Activation func	Hidden layer size	Number of epochs	Learning rate	Score	Figure
sigmoid	200	1	0.1	0.9548	Fig.1
sigmoid	100	1	0.1	0.9493	Fig.2
sigmoid	300	1	0.1	0.9549	Fig.3
sigmoid	200	1	1	0.8558	Fig.4
sigmoid	200	1	0.01	0.9067	Fig.5
sigmoid	200	3	0.05	0.9685	Fig.6
sigmoid	300	1	0.1	0.9549	Fig.7
sigmoid	300	3	0.05	0.9696	Fig.8

Conclusion

With increasing the number of hidden layers the score is improving. Increase in the number of epochs leads to higher score as well. Changing learning rate to higher level improves the score, but more epochs needed to obtain good result. Low learning rate (e.g., 0.01) makes the loss function smooth, however, under a high rate (e.g., 1) the function is rocky (seems, the gradient jumps). Thus, the best score (0.9696) has been achieved at 300 hidden layers size, 3 epochs and 0.05 learning rate.

Annex with figures

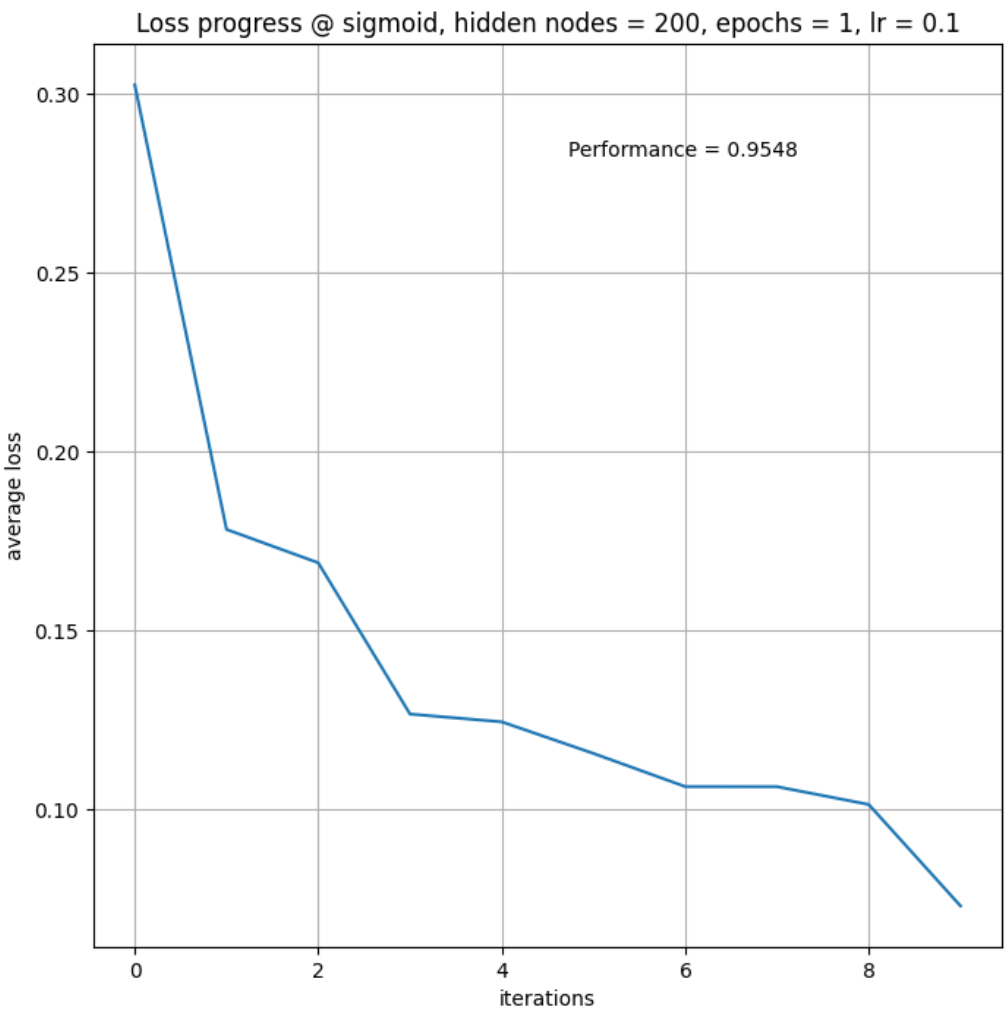


Fig.1

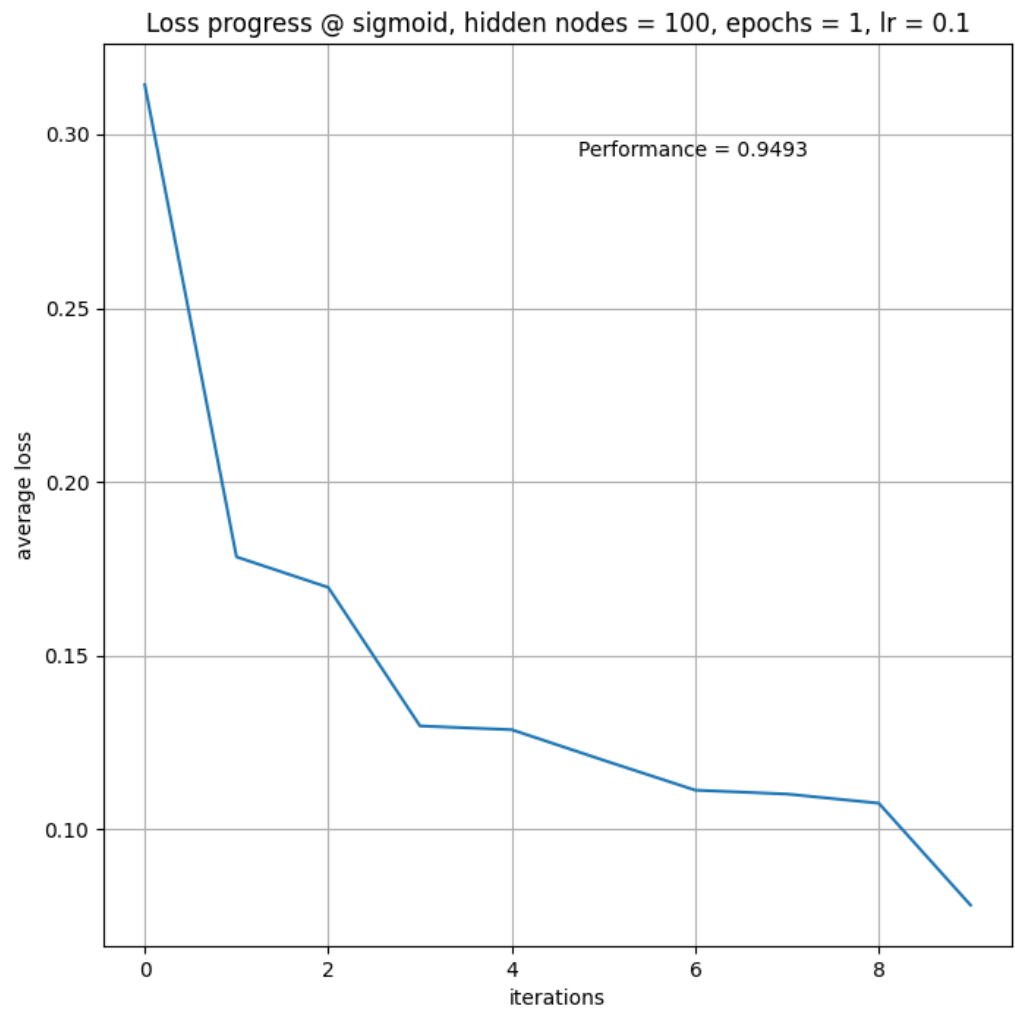


Fig.2

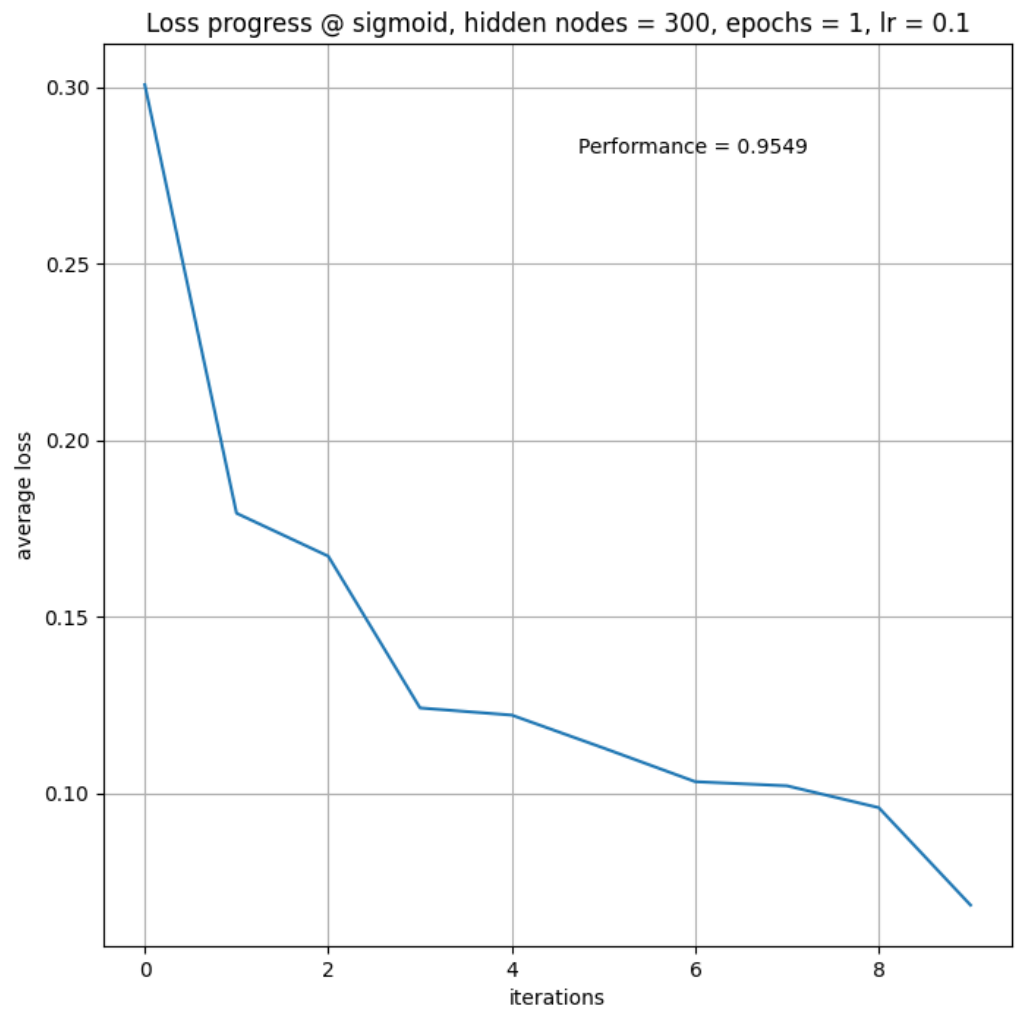


Fig.3

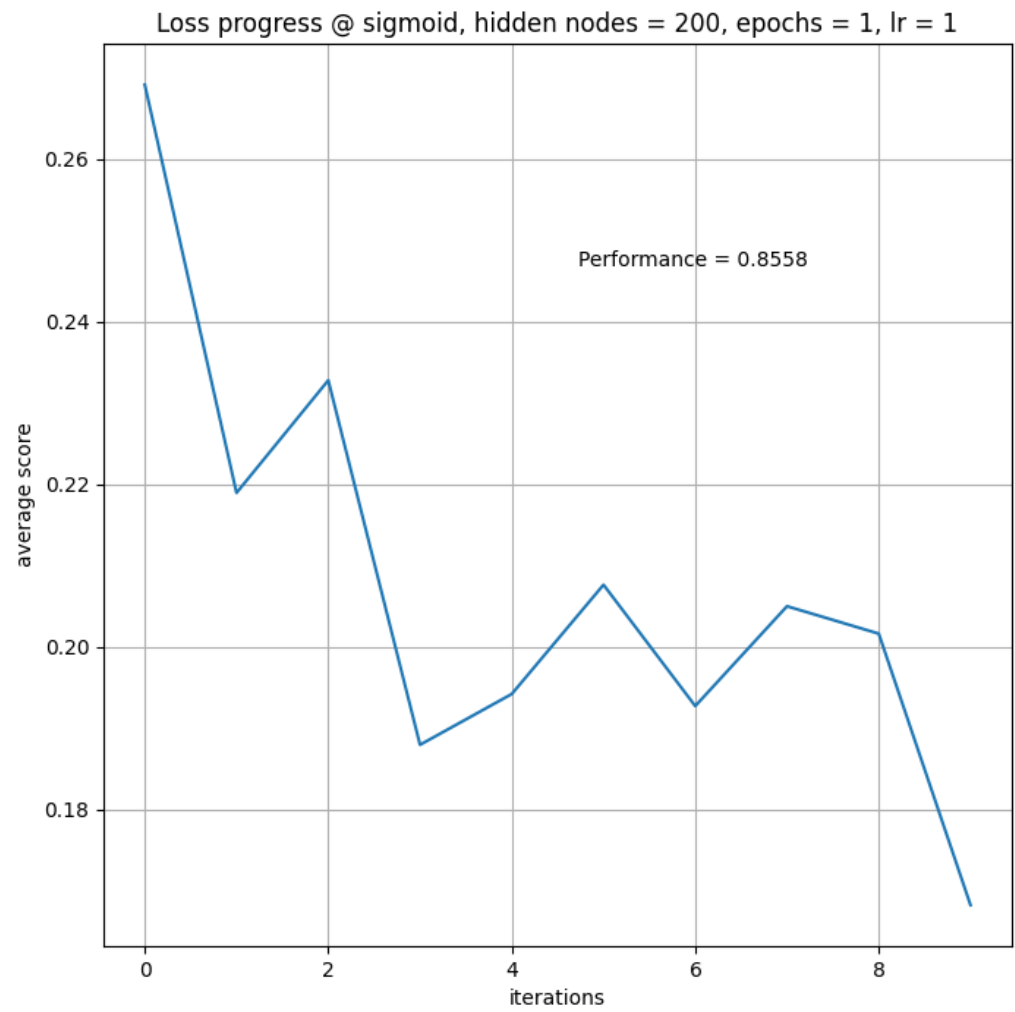


Fig.4

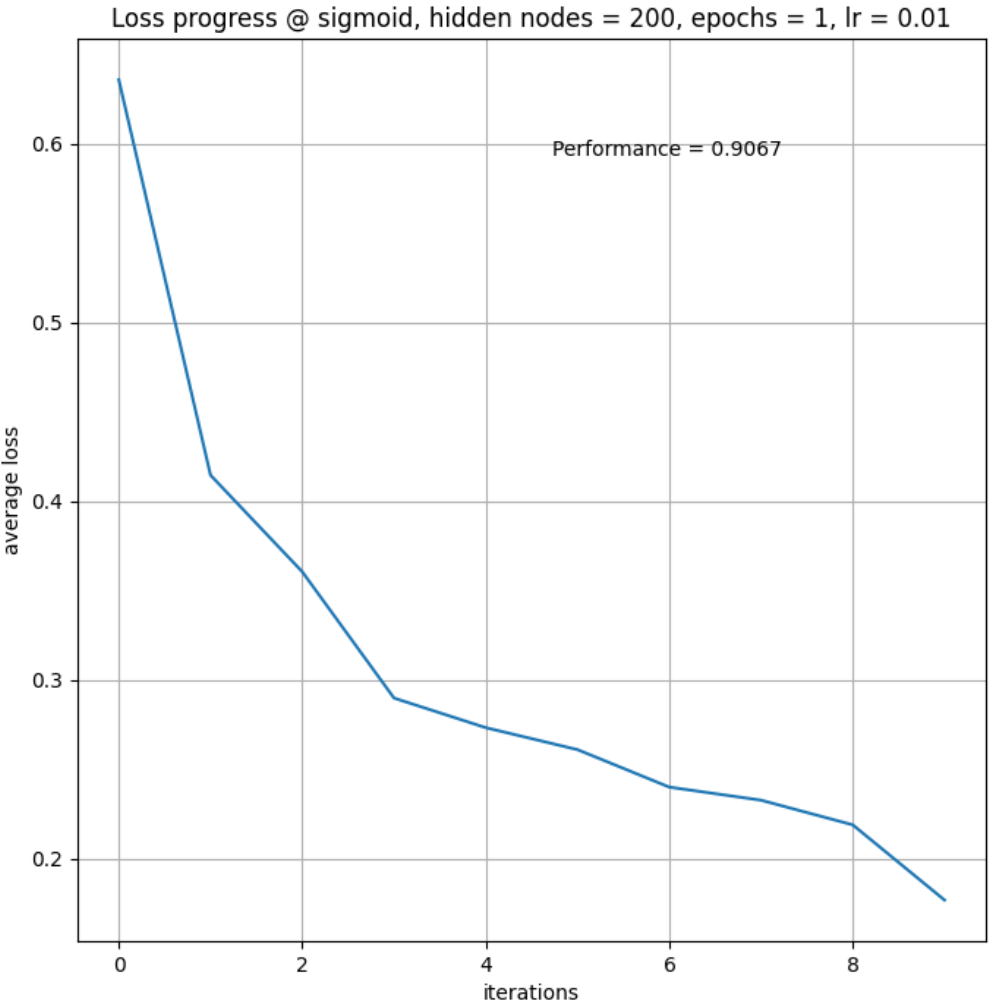


Fig.5

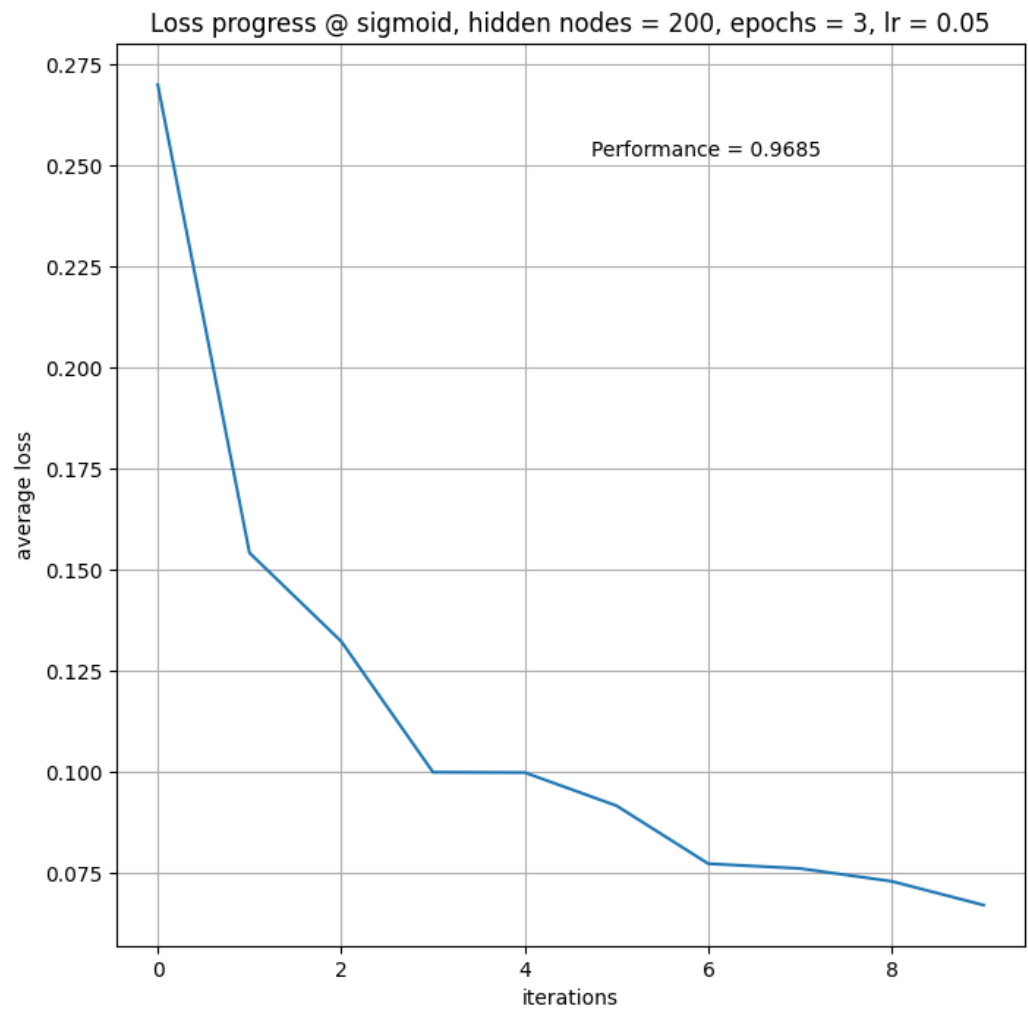


Fig.6

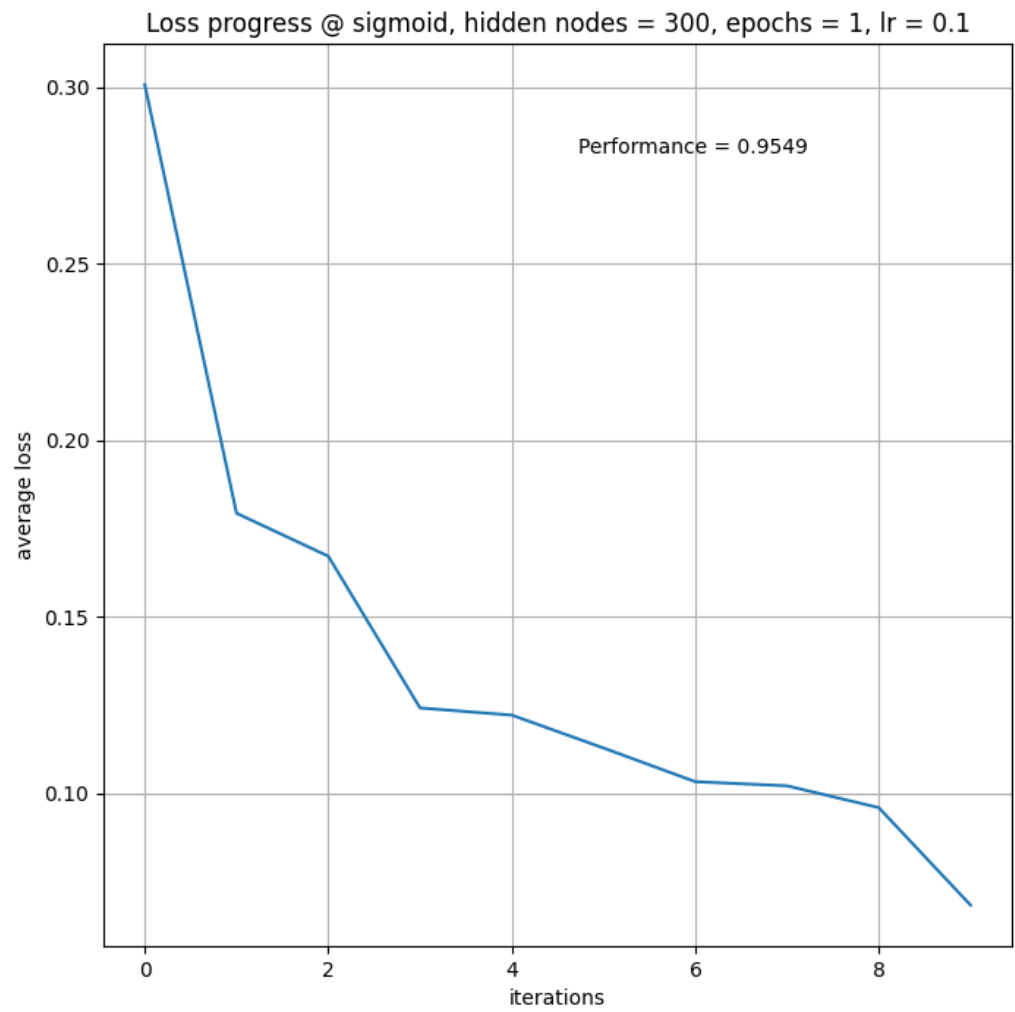


Fig.7

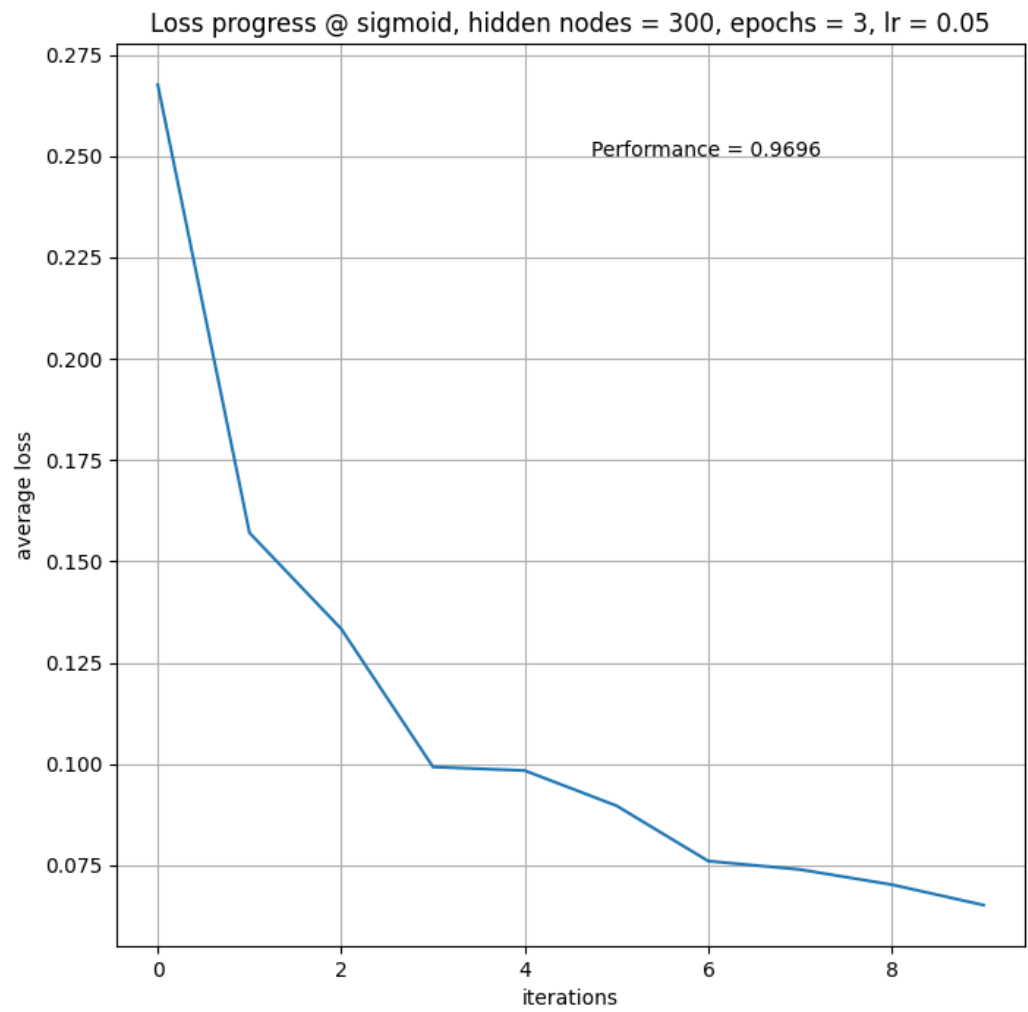


Fig.8