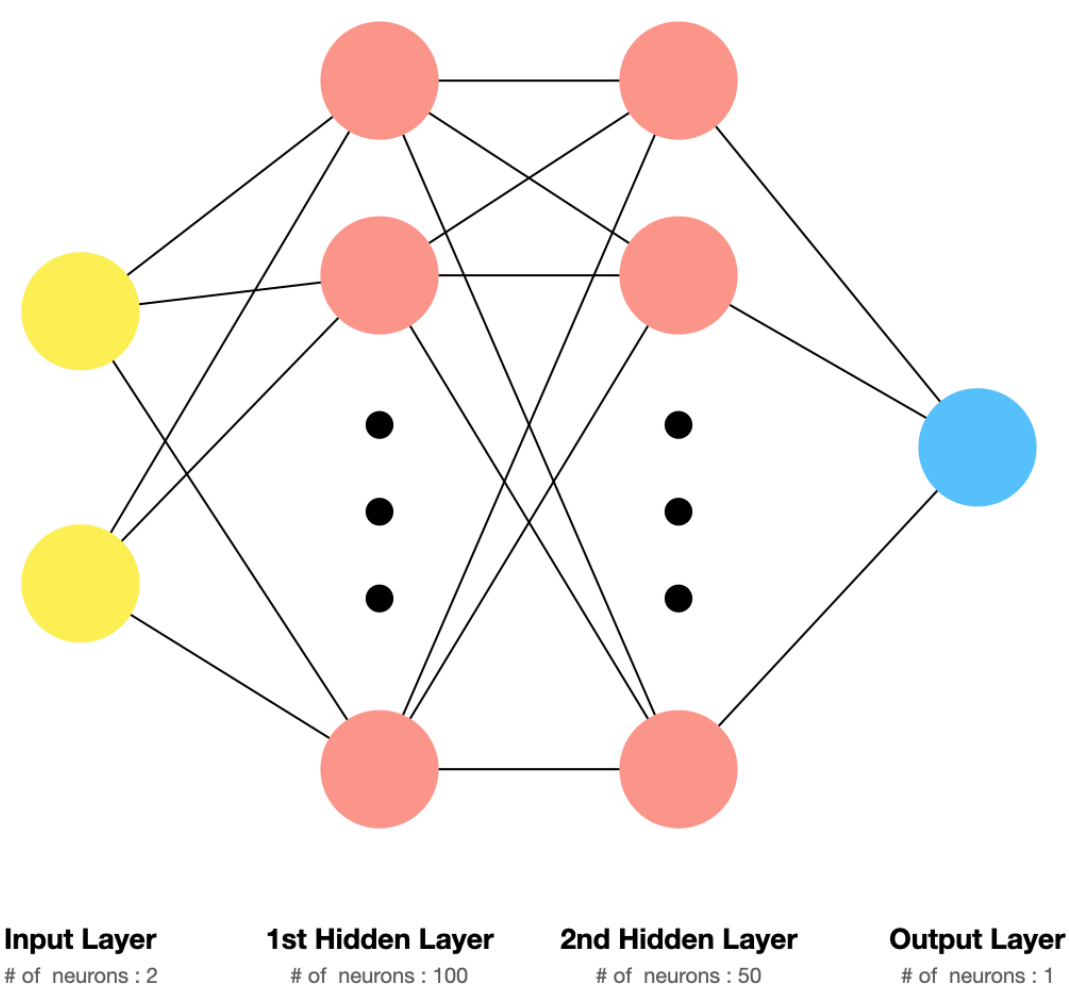


ML CW3 Report

- **Name** : Solang Kim
- **Student ID** : 20181041
- **E-mail** : solangii@unst.ac.kr

Implement Detail

My MLP Architecture



Weight Initialization

- Using **Xavier Normal Initialization** for weight initialization

Xavier Normal Initialization[**Glorot & Bengio, AISTATS 2010**]

$$W \sim N(0, Var(W))$$
$$Var(W) = \sqrt{\frac{2}{n_{in} + n_{out}}}$$

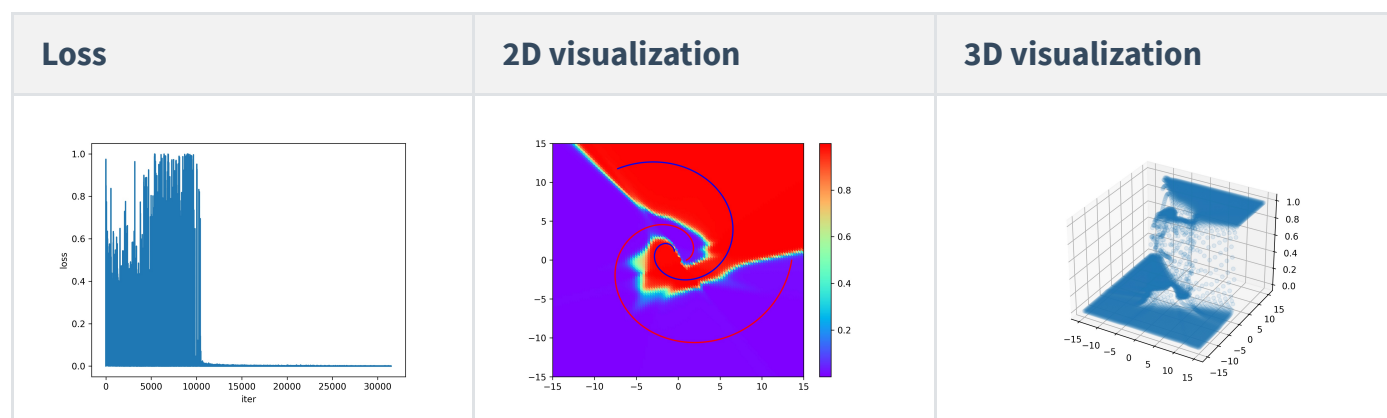
(n_{in} : # of previous layer's neurons, n_{out} : # of next layer's neurons)

- Bias initialization as **zero**

Hyper-Parameter Setting

- Learning Rate** : 1
- epochs** : 50
- batch size** : 1 (fully-SGD)
- using sigmoid for activation function

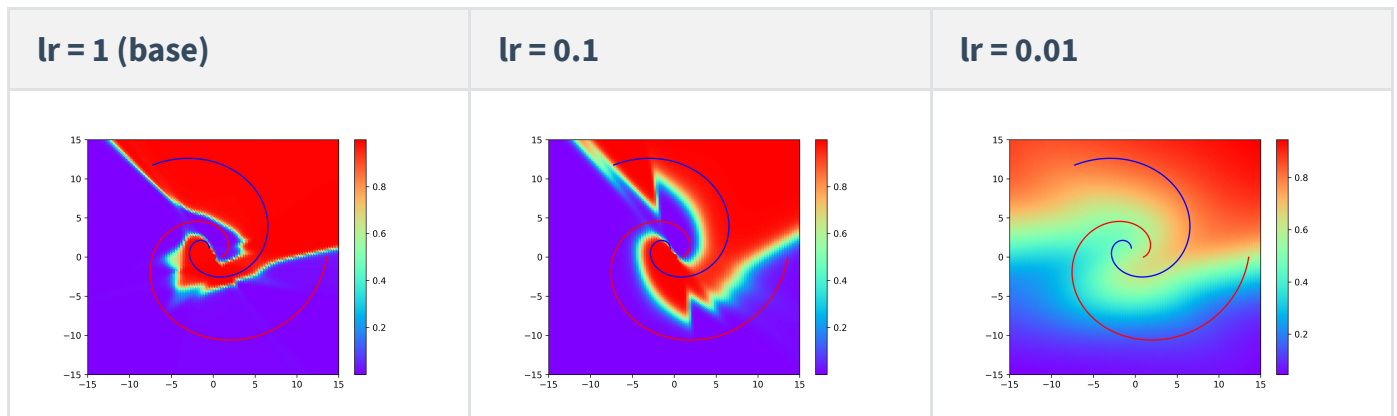
Performance



Additional Analysis

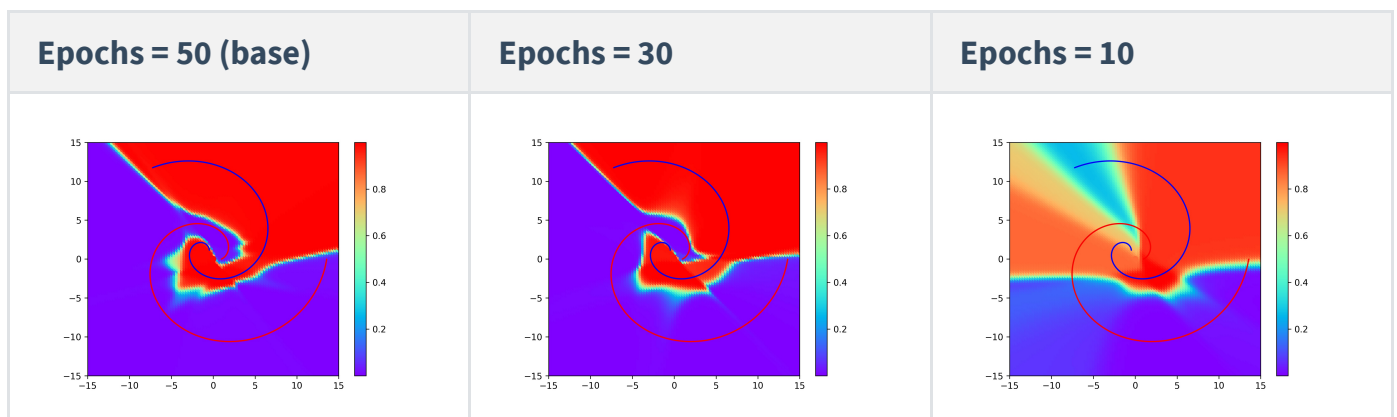
Except for the variable values regulated in the experiment, all values(hyper-params) are set to be the same.

Learning Rate



It appears to have found the appropriate learning rate value.

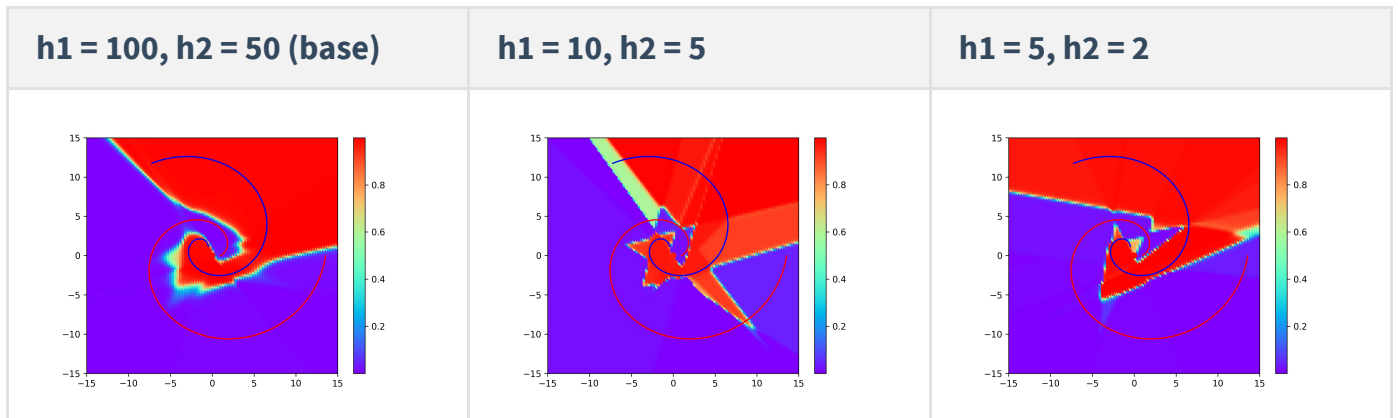
Epochs



Given the loss plot(loss graph), there has been no significant difference since it dropped sharply from about 10000-20000(itsers), indicating that learning has been stable since epochs was about 15-30. Therefore, there is no significant difference between 30 and 50. I think It's okay to stop learning at a location where loss is moderately reduced(about epochs 15-30).

Number of Neurons (for Hidden Layer)

- **h1** : # of 1st hidden layer's neuron
- **h2** : # of 2nd hidden layer's neuron



It appears to have found the appropriate # of neuron value.