NYCU DL
Lab1 - Backpropagation

TA 陳敬中 Bill

Mar. 4, 2025

Outline

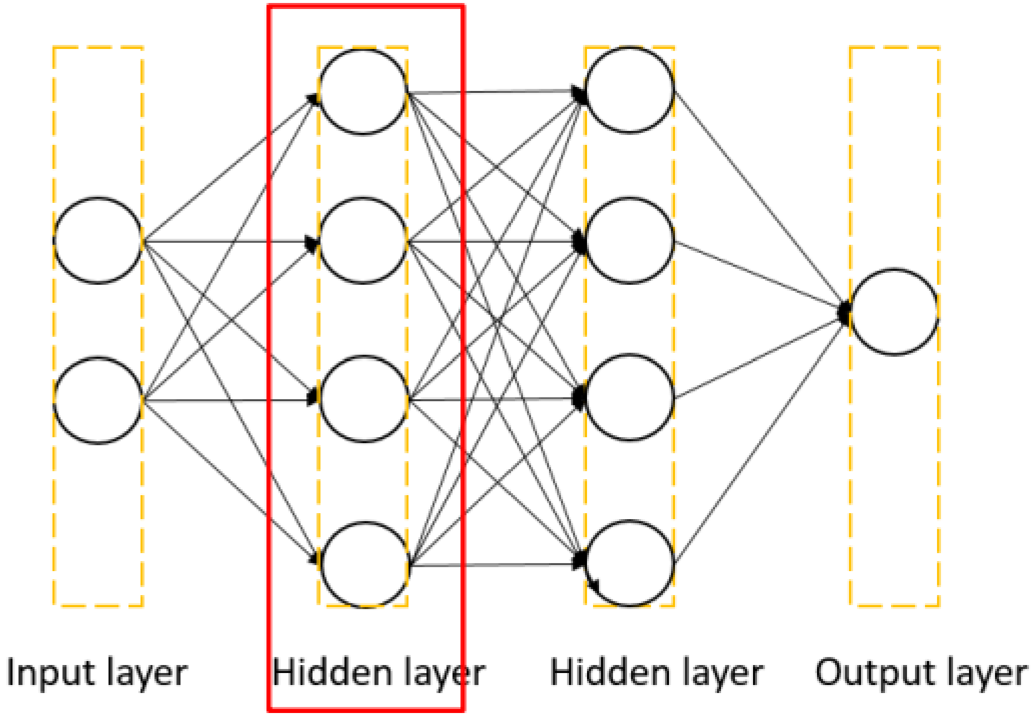
- Lab Description
- Scoring Criteria
- Time Schedule

Lab Description

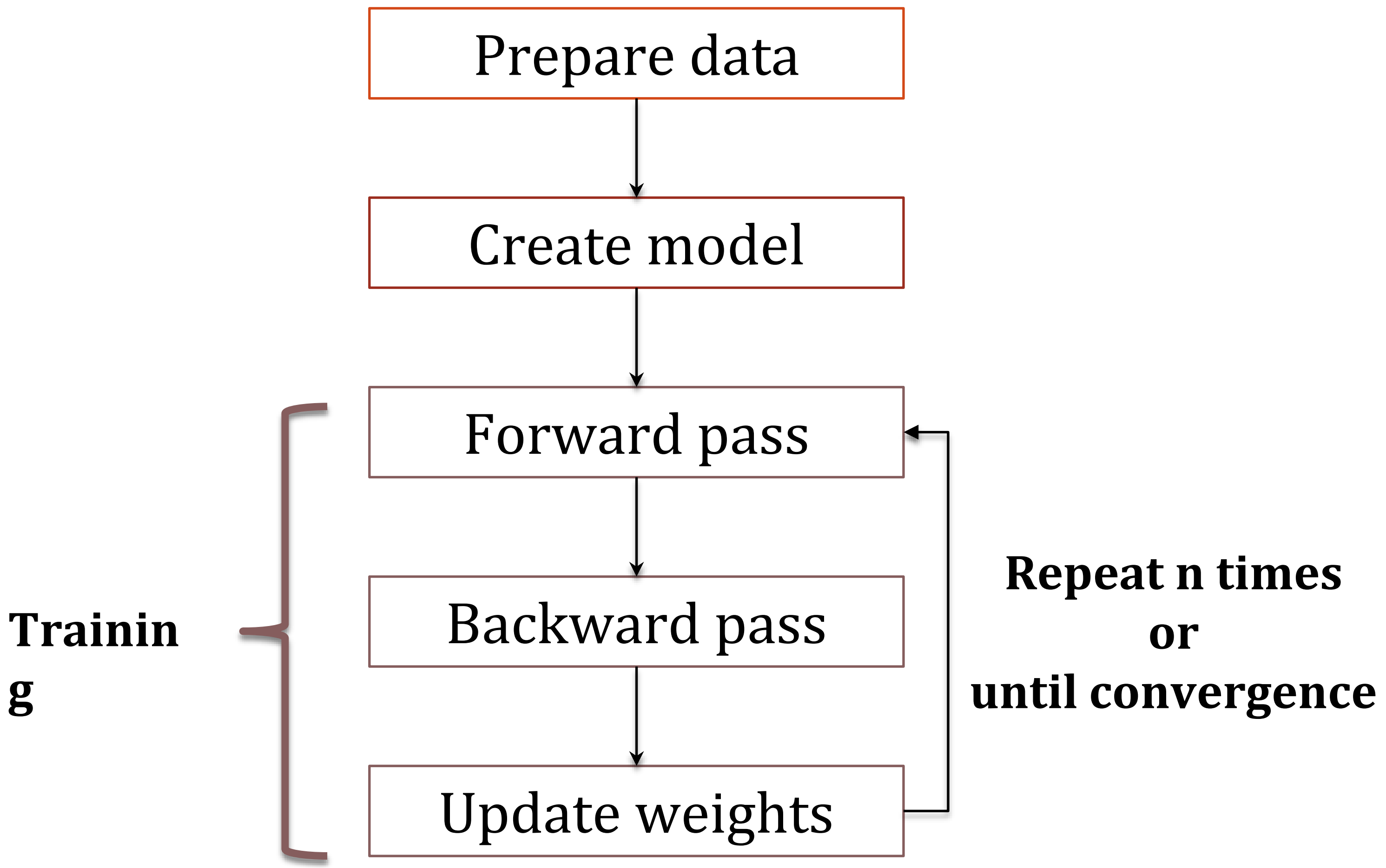
- Implement a simple neural network with two hidden layers
- Perform backpropagation to update model weights
- You can only use **Numpy** and other **python standard libraries**.
    - Pytorch, TensorFlow, … frameworks are **NOT** allowed in this lab
- Visualization
    - Plot comparison figures showing the predictions and ground truth.
    - Plot your learning curve (loss vs. epoch).
    - Print the accuracy of your prediction.
    - You are allowed to use matplotlib

Lab Description

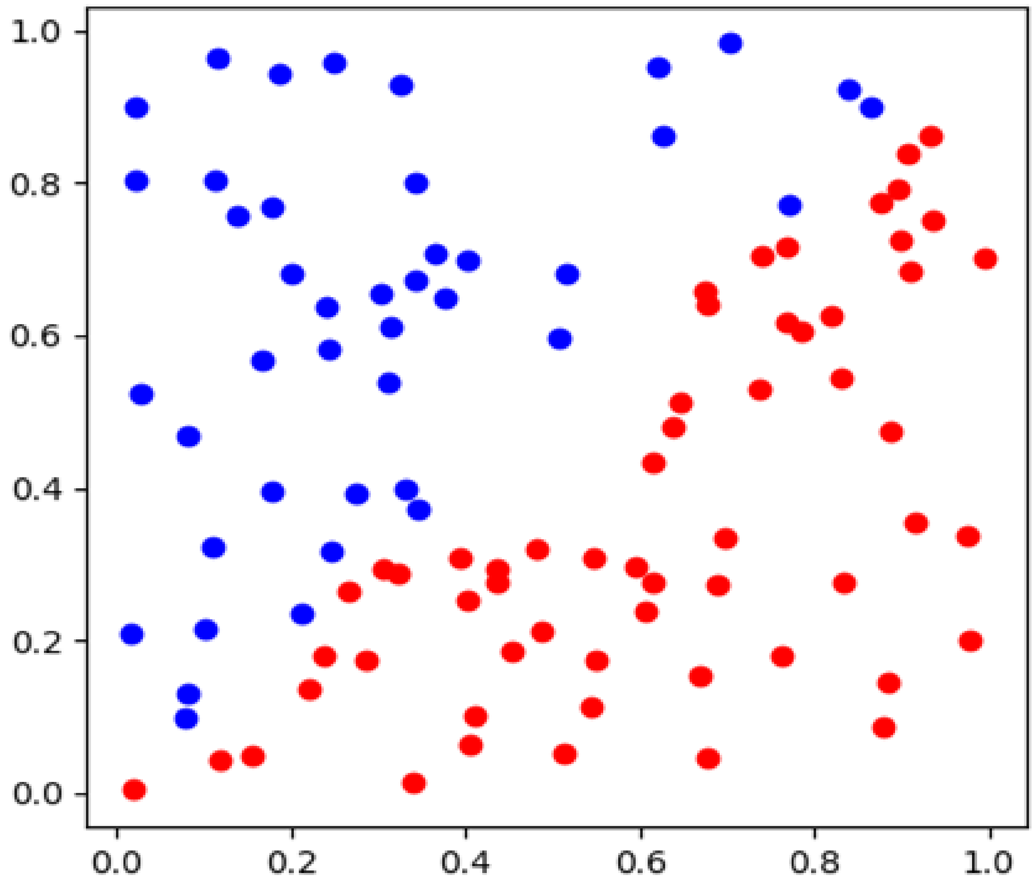
- Each layer should contain
  - At least one transformation (e.g., Linear, CNN, …)
  - One activation function (e.g., sigmoid, tanh, relu, …)



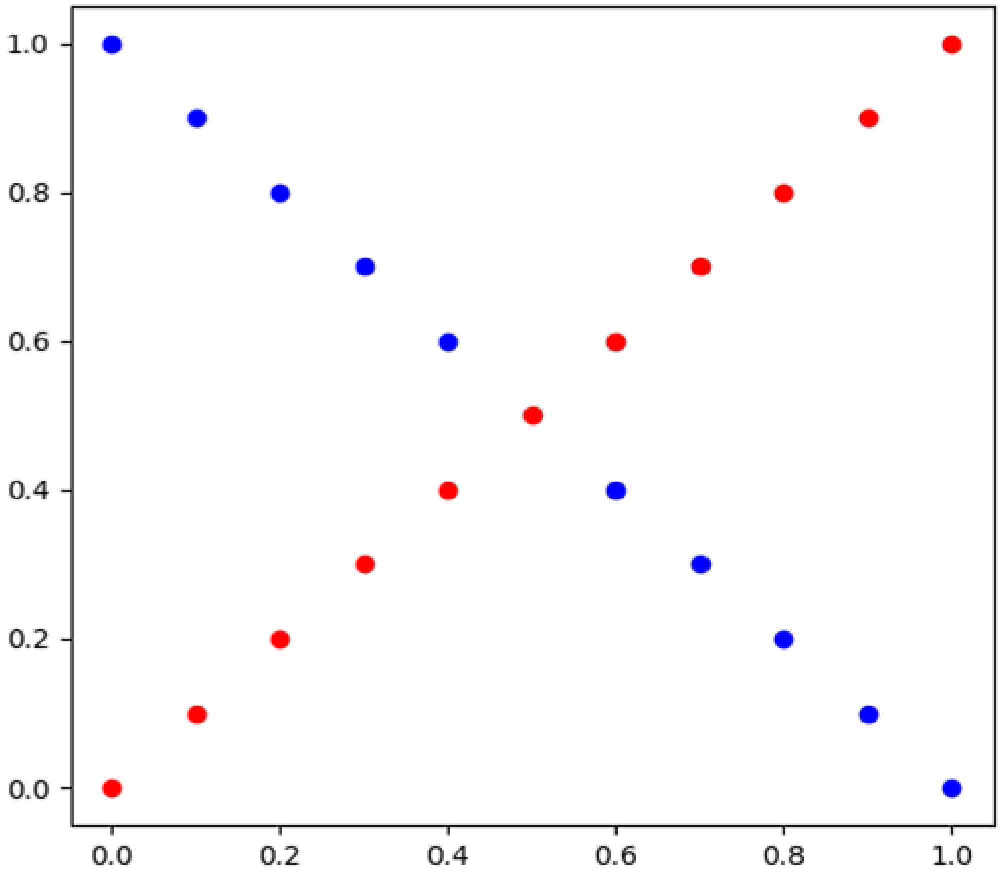
Lab Description – Flowchart



Lab Description - Data



generate_linear()



generate_XOR_easy()

Data Generation

Function usage

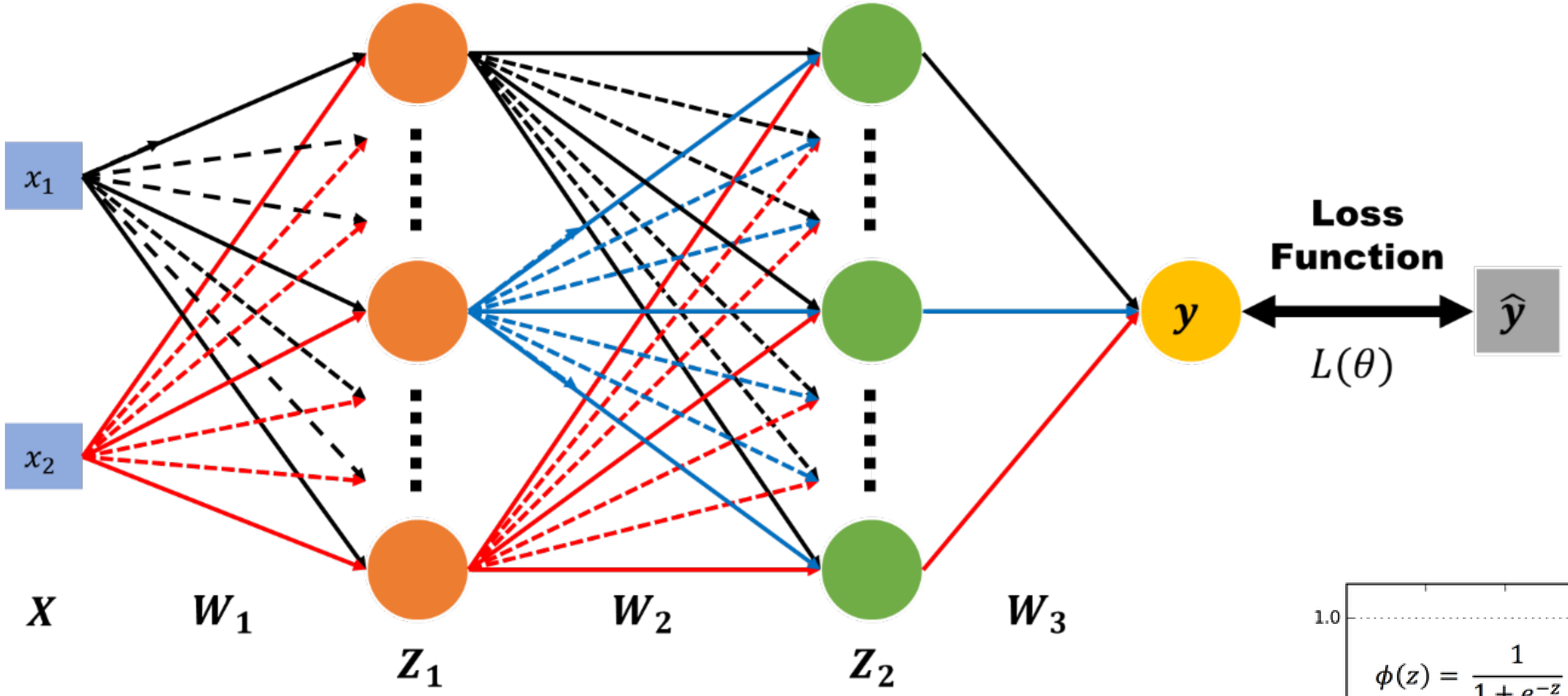
```python
x, y = generate_linear(n=100)
x, y = generate_XOR_easy()
```

- Do **NOT** overwrite these functions
- Training and Testing with the same data

```python
def generate_linear(n=100):
    import numpy as np
    pts = np.random.uniform(0, 1, (n, 2))
    inputs = []
    labels = []
    for pt in pts:
        inputs.append([pt[0], pt[1]])
        distance = (pt[0]-pt[1])/1.414
        if pt[0] > pt[1]:
            labels.append(0)
        else:
            labels.append(1)
    return np.array(inputs), np.array(labels).reshape(n, 1)
```

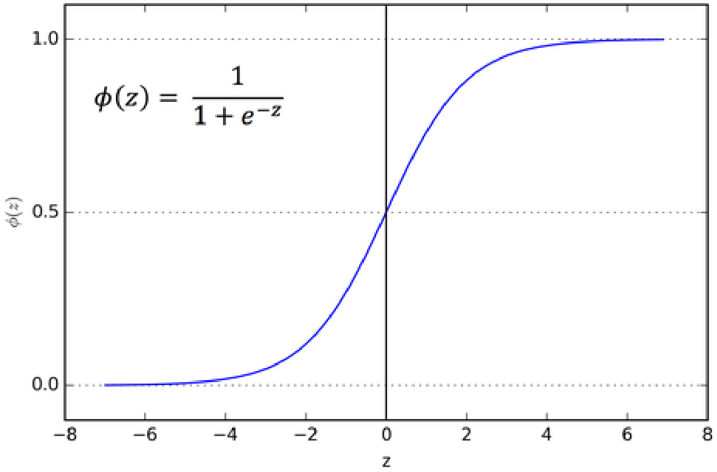
```python
def generate_XOR_easy():
    import numpy as np
    inputs = []
    labels = []

    for i in range(11):
        inputs.append([0.1*i, 0.1*i])
        labels.append(0)

        if 0.1*i == 0.5:
            continue

        inputs.append([0.1*i, 1-0.1*i])
        labels.append(1)

    return np.array(inputs), np.array(labels).reshape(21, 1)
```

Lab Description – Forward

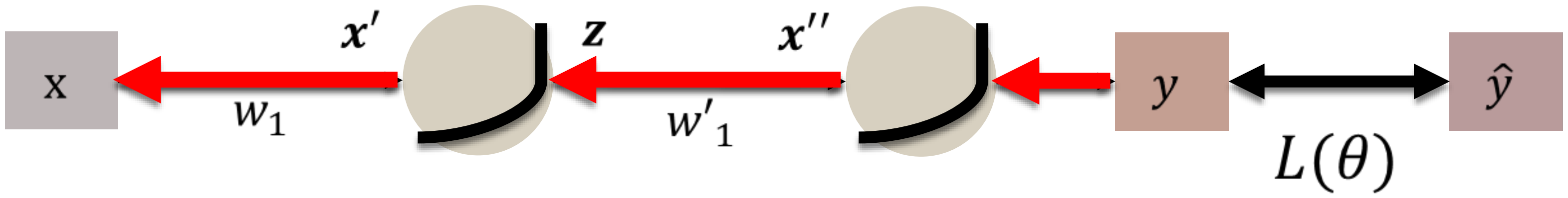


$$Z_1 = \sigma(XW_1) \qquad Z_2 = \sigma(Z_1W_2) \qquad y = \sigma(Z_2W_3)$$

$$\sigma(\mathrm{x}) = \frac{1}{1+e^{-x}}$$



Lab Description – Backward



$\boldsymbol{x' = xw_1}$ $\boldsymbol{z = \sigma(x')}$ $\boldsymbol{x'' = zw'_1}$ $\boldsymbol{y = \sigma(x'')}$

Chain rule

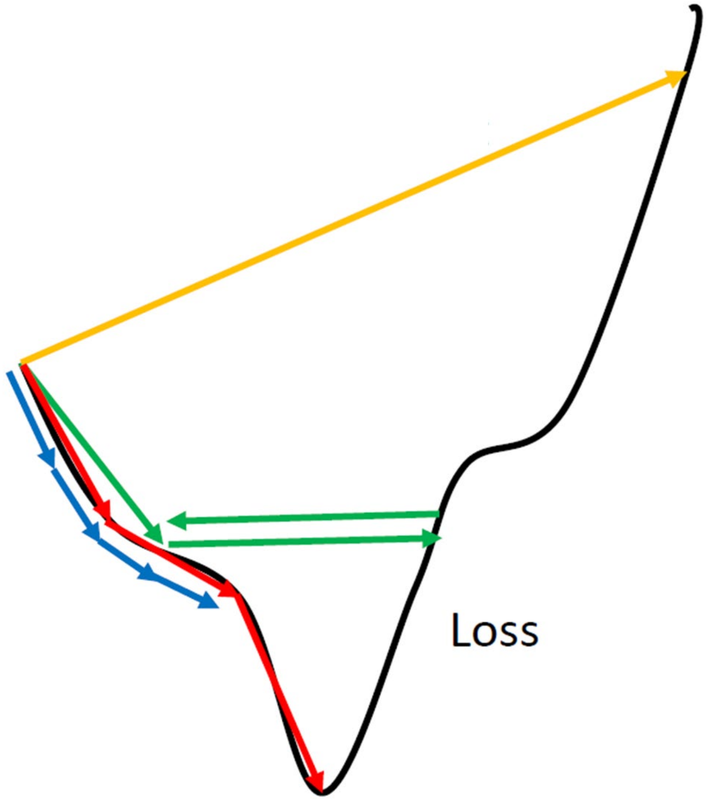
$$y = g(x) \quad z = h(y)$$

$$\mathbf{x} \xrightarrow{\mathbf{g}()} \boldsymbol{y} \xrightarrow{\boldsymbol{h}()} \boldsymbol{z} \qquad \frac{\mathbf{dz}}{\mathbf{dx}} = \frac{\mathbf{dz}}{\cancel{\mathbf{dy}}}\frac{\cancel{\mathbf{dy}}}{\mathbf{dx}}$$

$$\begin{aligned}
\frac{\partial L(\theta)}{\partial w_1} &= \frac{\partial y}{\partial w_1}\frac{\partial L(\theta)}{\partial y} \\
&= \frac{\partial x''}{\partial w_1}\frac{\partial y}{\partial x''}\frac{\partial L(\theta)}{\partial y} \\
&= \frac{\partial z}{\partial w_1}\frac{\partial x''}{\partial z}\frac{\partial y}{\partial x''}\frac{\partial L(\theta)}{\partial y} \\
&= \frac{\partial x'}{\partial w_1}\frac{\partial z}{\partial x'}\frac{\partial x''}{\partial z}\frac{\partial y}{\partial x''}\frac{\partial L(\theta)}{\partial y}
\end{aligned}$$

Lab Description – Gradient descent

Network Parameters $\theta = \{w_1, w_2, w_3, w_4, \cdots\}$

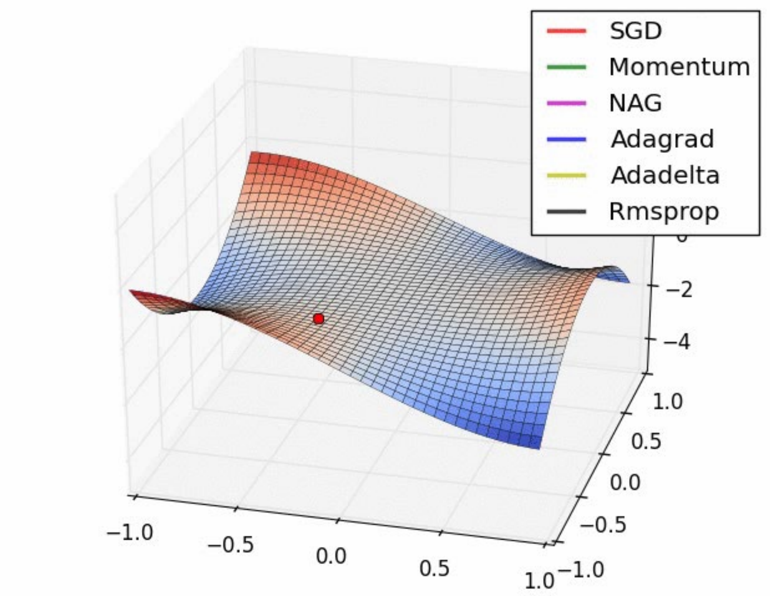


$$\theta^1 = \theta^0 - \rho \, \nabla L(\theta^0)$$

$$\theta^2 = \theta^1 - \rho \, \nabla L(\theta^1)$$

$$\theta^3 = \theta^2 - \rho \, \nabla L(\theta^2)$$

ρ : Learning rate



http://www.denizyuret.com/2015/03/alec-radfords-animations-for.html

Lab Description - Prediction

- During training, print the loss

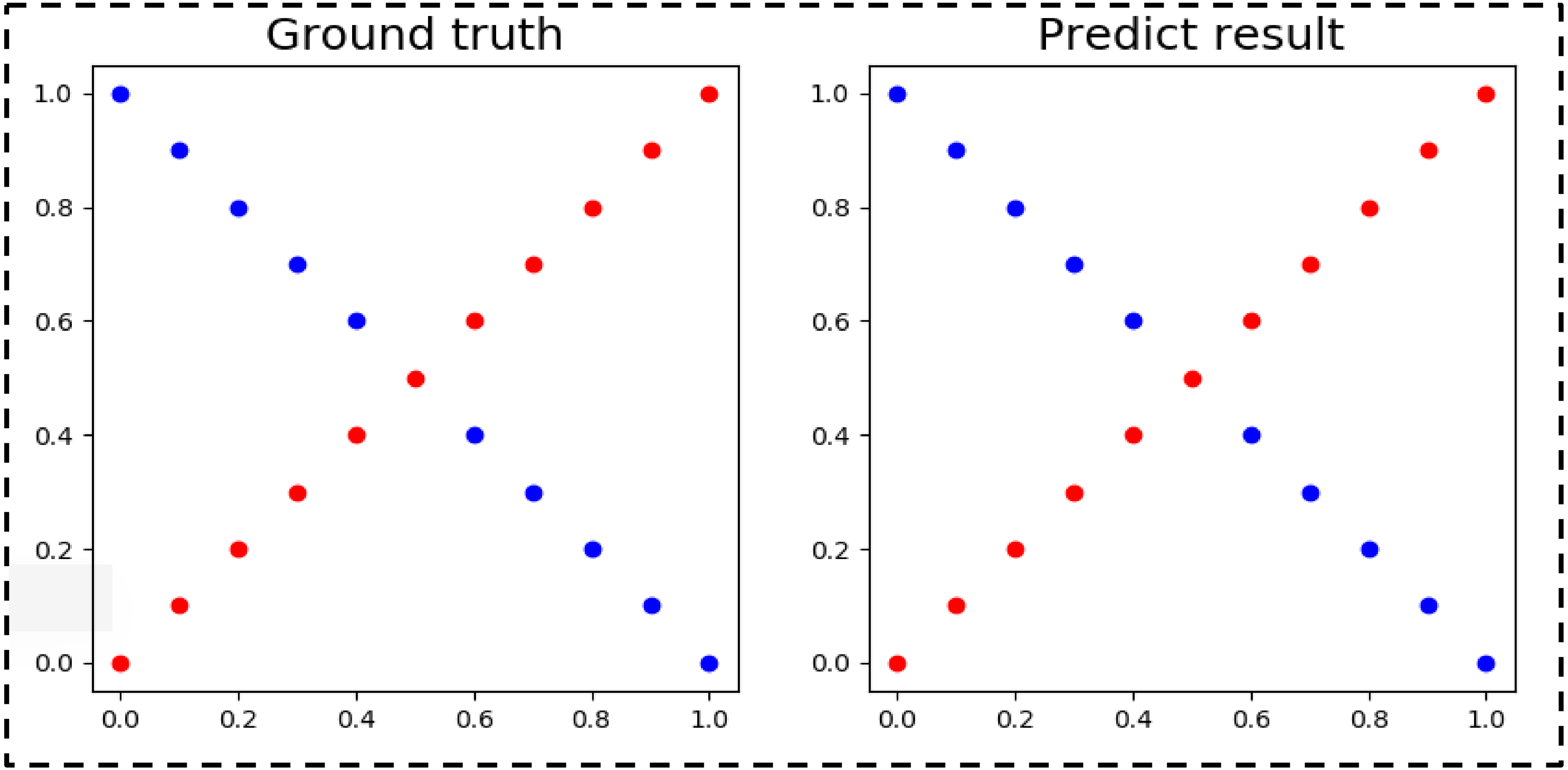
```
epoch 10000 loss : 0.16234523253277644
epoch 15000 loss : 0.2524336634177614
epoch 20000 loss : 0.1590783047540092
epoch 25000 loss : 0.2209944703023485З
epoch 30000 loss : 0.3292173477217561
epoch 35000 loss : 0.40406233282426085
epoch 40000 loss : 0.43052897480298924
epoch 45000 loss : 0.4207525735586605
epoch 50000 loss : 0.3934759509342479
epoch 55000 loss : 0.3615008372106921
epoch 60000 loss : 0.33077879872648525
epoch 65000 loss : 0.30333537090819584
epoch 70000 loss : 0.2794858089741792
epoch 75000 loss : 0.25892812312991587
epoch 80000 loss : 0.24119780823897027
epoch 85000 loss : 0.22583656353511342
epoch 90000 loss : 0.21244497028971704
epoch 95000 loss : 0.2006912468389013
```

- During testing, show the predictions and the accuracy

```
Iter91 |    Ground truth: 1.0 |    prediction: 0.99943 |
Iter92 |    Ground truth: 1.0 |    prediction: 0.99987 |
Iter93 |    Ground truth: 1.0 |    prediction: 0.99719 |
Iter94 |    Ground truth: 1.0 |    prediction: 0.99991 |
Iter95 |    Ground truth: 0.0 |    prediction: 0.00013 |
Iter96 |    Ground truth: 1.0 |    prediction: 0.77035 |
Iter97 |    Ground truth: 1.0 |    prediction: 0.98981 |
Iter98 |    Ground truth: 1.0 |    prediction: 0.99337 |
Iter99 |    Ground truth: 0.0 |    prediction: 0.20275 |
loss=0.03844 accuracy=100.00%
```

Lab Description - Prediction

- Visualize the predictions and ground truth at the end of the training process



Scoring Criteria

- Achieve 90% accuracy in testing to get 40% experimental results score

Report Spec:

1. Introduction (5%)
2. Implementation Details (15%):
   A. Sigmoid function
   B. Neural network architecture
   C. Back-propagation
3. Experimental Results (45%)
   A. Screenshot and comparison figure
   B. Show the accuracy of your prediction (40%) (achieve 90% accuracy)
   C. Learning curve (loss-epoch curve)
   D. Anything you want to present
4. Discussion (15%)
   A. Try different learning rates
   B. Try different numbers of hidden units
   C. Try without activation functions
   D. Anything you want to share
5. Questions (20%)
   A. What is the purpose of activation functions? (6%)
   B. What might happen if the learning rate is too large or too small? (7%)
   C. What is the purpose of weights and biases in a neural network? (7%)
6. Extra (10%)
   A. Implement different optimizers. (2%)
   B. Implement different activation functions. (3%)
   C. Implement convolutional layers. (5%)

Important Date

- Report Submission Deadline: 3/11 (Tue.) 23:59
- Zip all files in one file
  - Report (report.pdf)
  - Source code
- Name it like「DL_LAB1_yourstudentID_name.zip」
  - E.g.,「DL_LAB1_313551157_陳敬中.zip」
- If there are any format errors in your files, you will receive a **5-point** penalty

Reference

1. *http://www.denizyuret.com/2015/03/alec-radfords-animations-for.html*
2. *http://speech.ee.ntu.edu.tw/~tlkagk/courses_ML17_2.html*