Deep Neural Networks

Janghun Jo
Computer Graphics Lab.
jhjo432@postech.ac.kr

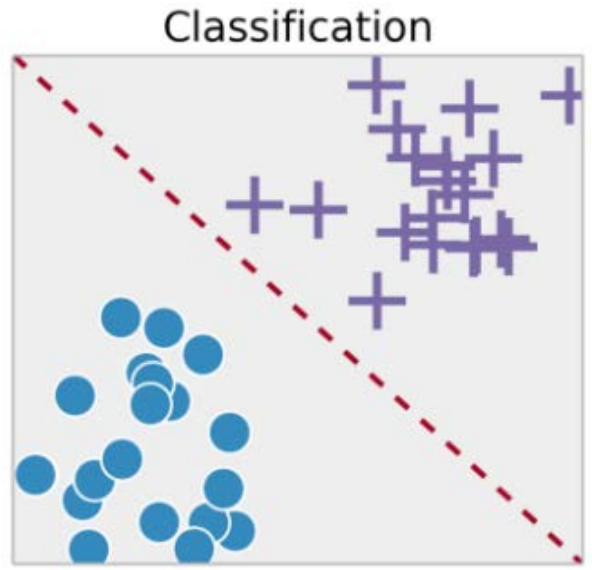
POSTECH

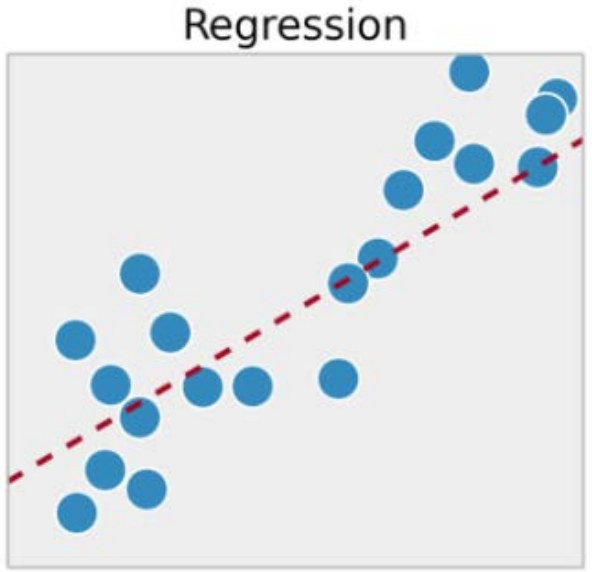
Contents

- Classification
  - XOR problem
- Regression
  - Non-linear regression with DNN
  - Robust regression

Classification and Regression

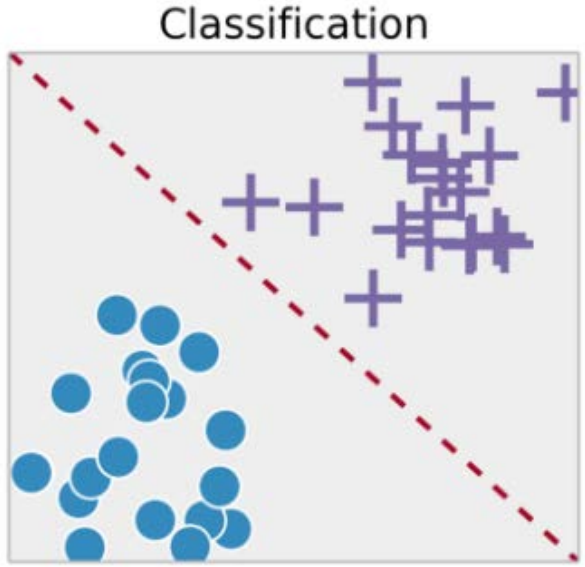
- Machine learning tasks have two types
  - Classification : predict class with categorical data
  - Regression : predict continuous value





Classification

- Classification: predict class with categorical data
  - Logistic regression
    - Single layer
    - Sigmoid activation function
  - Softmax regression (for multi class Classification)
    - Single layer
    - Softmax activation function
  - DNN classifier



◆Note: Logistic regression and Softmax regression are not regression algorithms.

Loss function for Classification

- Cross entropy loss (*typically for classification*)
  - Entropy of probability distribution P

$$H(P) = E_P[-\log P] = -\sum_X P(X) \log P(X)$$

  - Cross entropy between two probability distributions P and Q

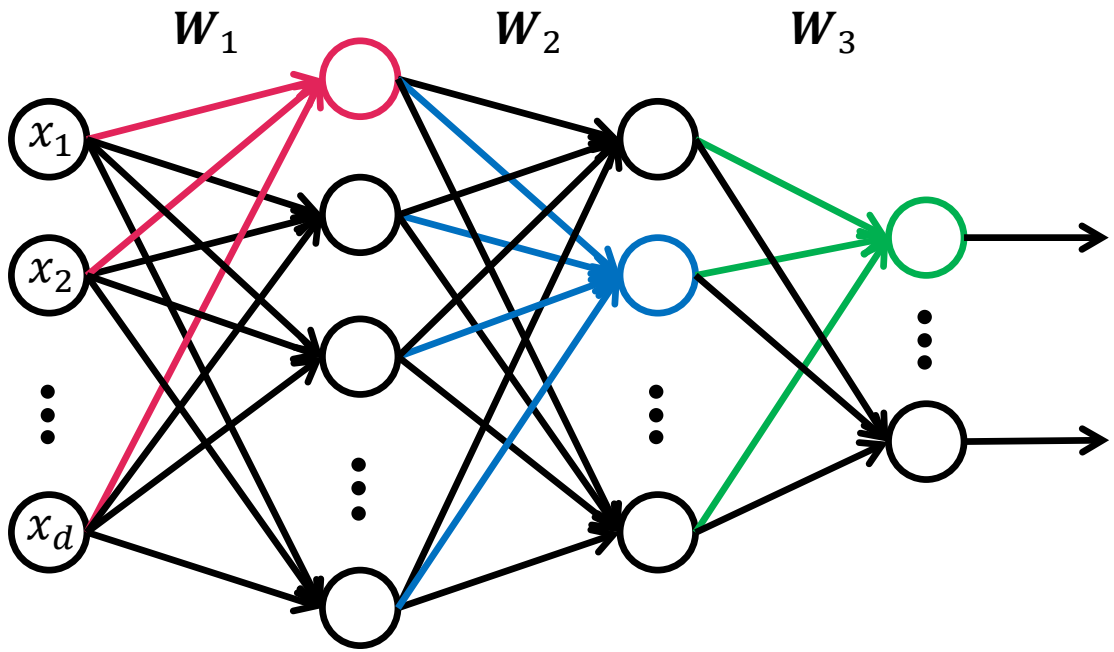
$$H(P, Q) = E_P[-\log Q] = -\sum_X P(X) \log Q(X)$$

  - Cross entropy loss
    - P: Ground truth label distribution
    - Q: Predicted probability distribution

$$L(\boldsymbol{y}, i) = -\log\left(\frac{\exp(y_i)}{\sum_{j=1}^{C} \exp(y_j)}\right) = -y_i + \log\left(\sum_{j=1}^{C} \exp(y_j)\right)$$

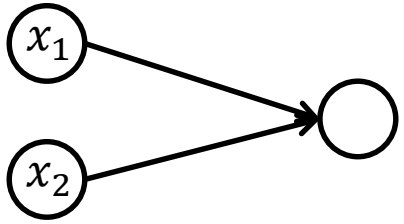
Multilayer Perceptron

- Stacking layers of multiple perceptrons
- Advantages
  - Nonlinear classification:
    More complex decision boundary can be defined using multiple layers.
  - Typically achieves better performance.

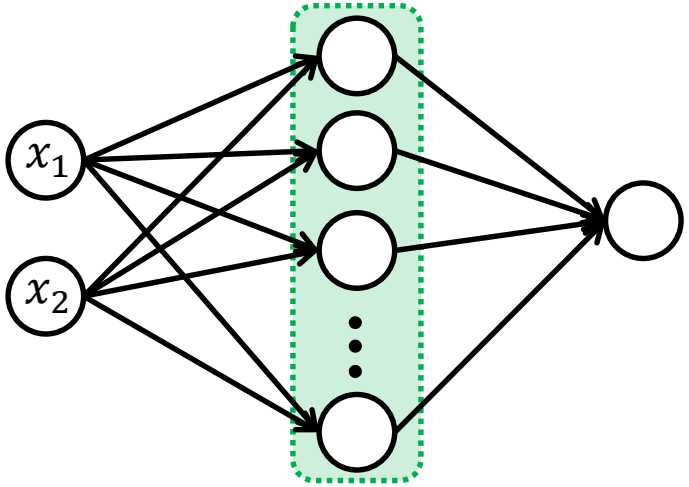


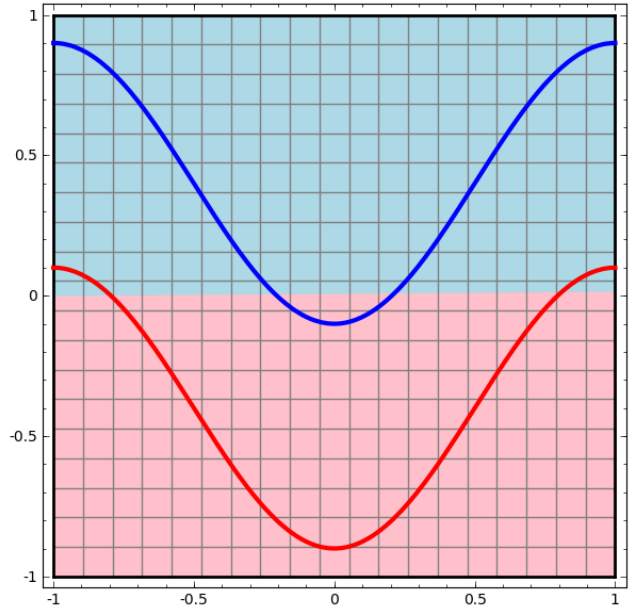
Multilayer Perceptron

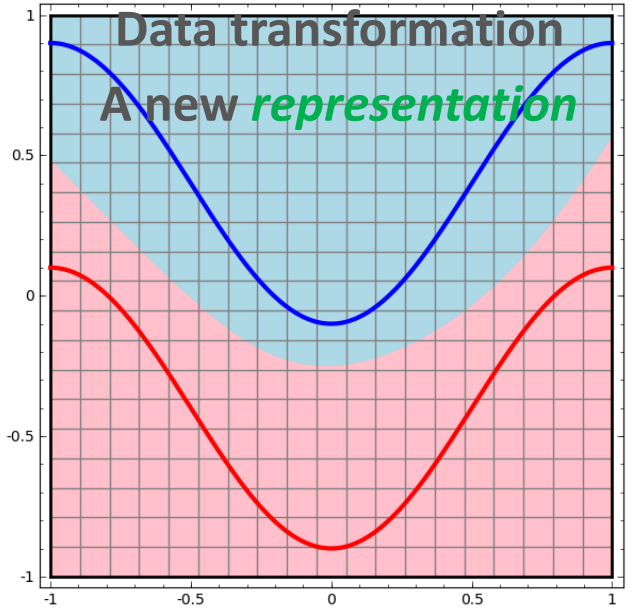
Single Perceptron



Multi-layer perceptron





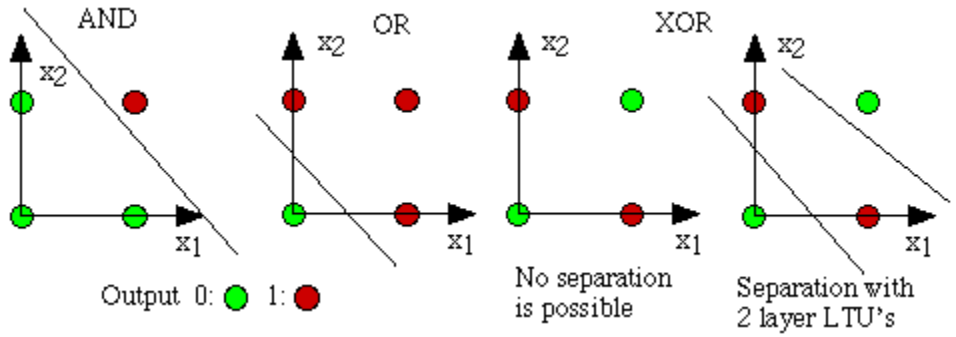


XOR Problem

- What is the XOR operation?
  - Exclusive OR
  - XOR gives a True output when the number of True inputs is odd.

INPUT		OUTPUT
A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0

- XOR problem: predict output of XOR operation using a neural network.
  - Single layer neural networks cannot solve the XOR problem
    - Data points are not linearly separable.
  - Deep Neural Networks can solve the XOR problem.



Quiz 1. Solving XOR problem

- XOR 데이터 셋 구현

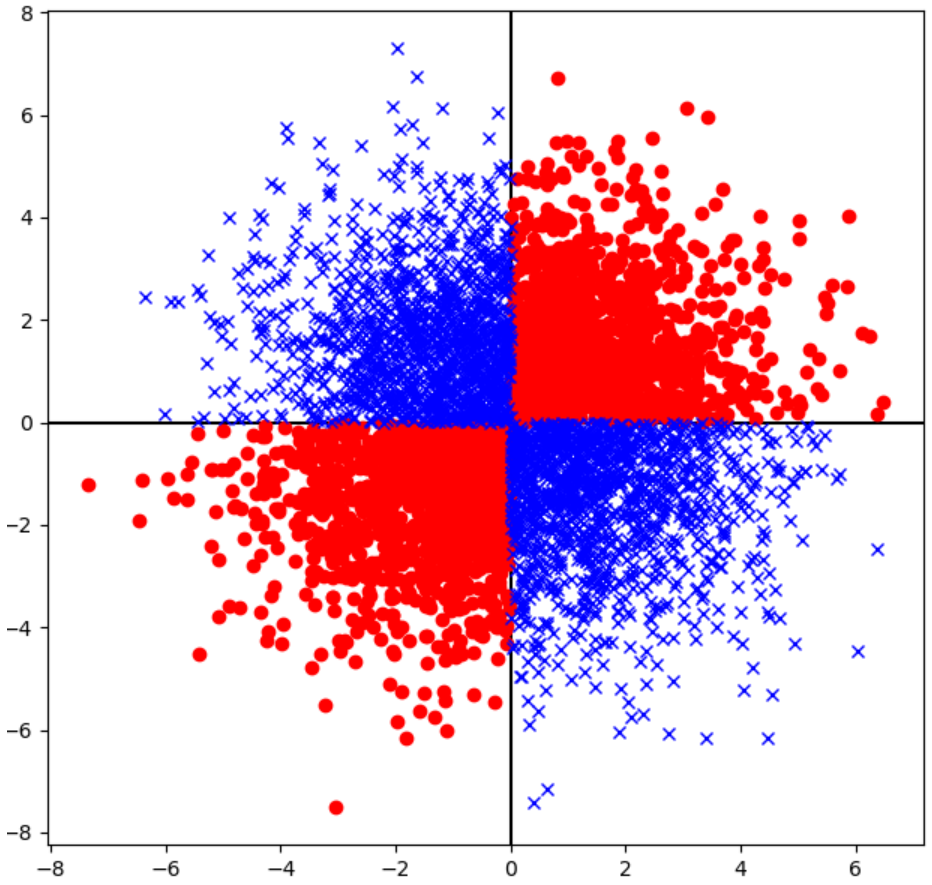
- Network 구현
  - single layer network구현
    - Fully connected layer – in_features: 2, out_features: 1
  - DNN 구현
    - Fully connected layer – in_features: 2, out_features: 20
    - Fully connected layer – in_features: 20, out_features: 20
    - Fully connected layer – in_features: 84, out_features: 1
  - Note
    - Apply ReLU activation function for hidden layers.
    - Apply Sigmoid activation function for output layers.

- Define a Loss function and optimizer
  - Cross-Entropy loss
  - SGD with learning rate 0.01 and momentum 0.5

- Train the network on the training data
- Test the network on the test data

Creating XOR Dataset

- We can make XOR dataset using torch.randn function
  - Generate random (x, y) using torch.randn function
  - if x * y < 0 then label should be 0
  - else (x * y >= 0) then label should be 1



Recap: How to make datasets?

- All datasets are subclass of **torch.utils.data.Dataset**
  - Make new class and inherit **torch.utils.Dataset**
  - It should have **__len__(self)** method and **__getitem__(self, idx)** method
    - **__len__(self)**: returns size of dataset
    - **__getitem__(self, idx)**: returns (**idx**)th data sample in Dataset

- **torch.utils.data.DataLoader** reads datasets and make the batch

```
CLASS  torch.utils.data.DataLoader(dataset, batch_size=1, shuffle=False,
```

  - **dataset** – dataset from which to load the data. (**torch.utils.data.Dataset**)
  - **batch_size** – how many samples per batch to load (default: 1)
  - **shuffle** – set to True to have the data reshuffled at every epoch (default: False).

  - Usage: for batch_idx, data, label in enumerate(trainloader)