Machine Learning Basics

Abstract of Deep Learning

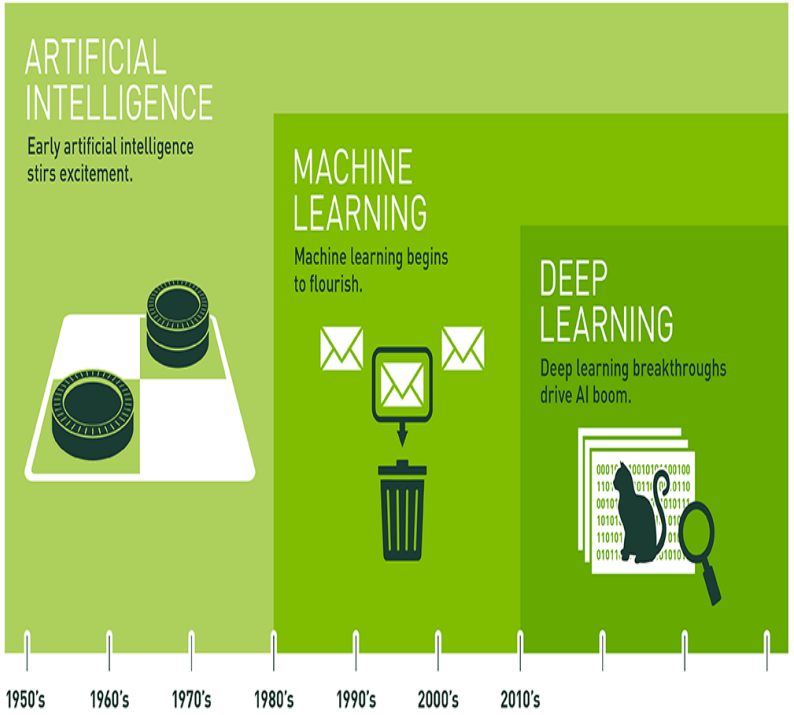
- Problem of data analysis :

  Too many data that include several parameters

  → Deep learning can be appropriate

- Trend : The technique of deep learning has been improved.
  Ex) The victory of artificial intelligence over professionals.
  Improvement of image recognition

- Desire to incorporate this technology into data analysis

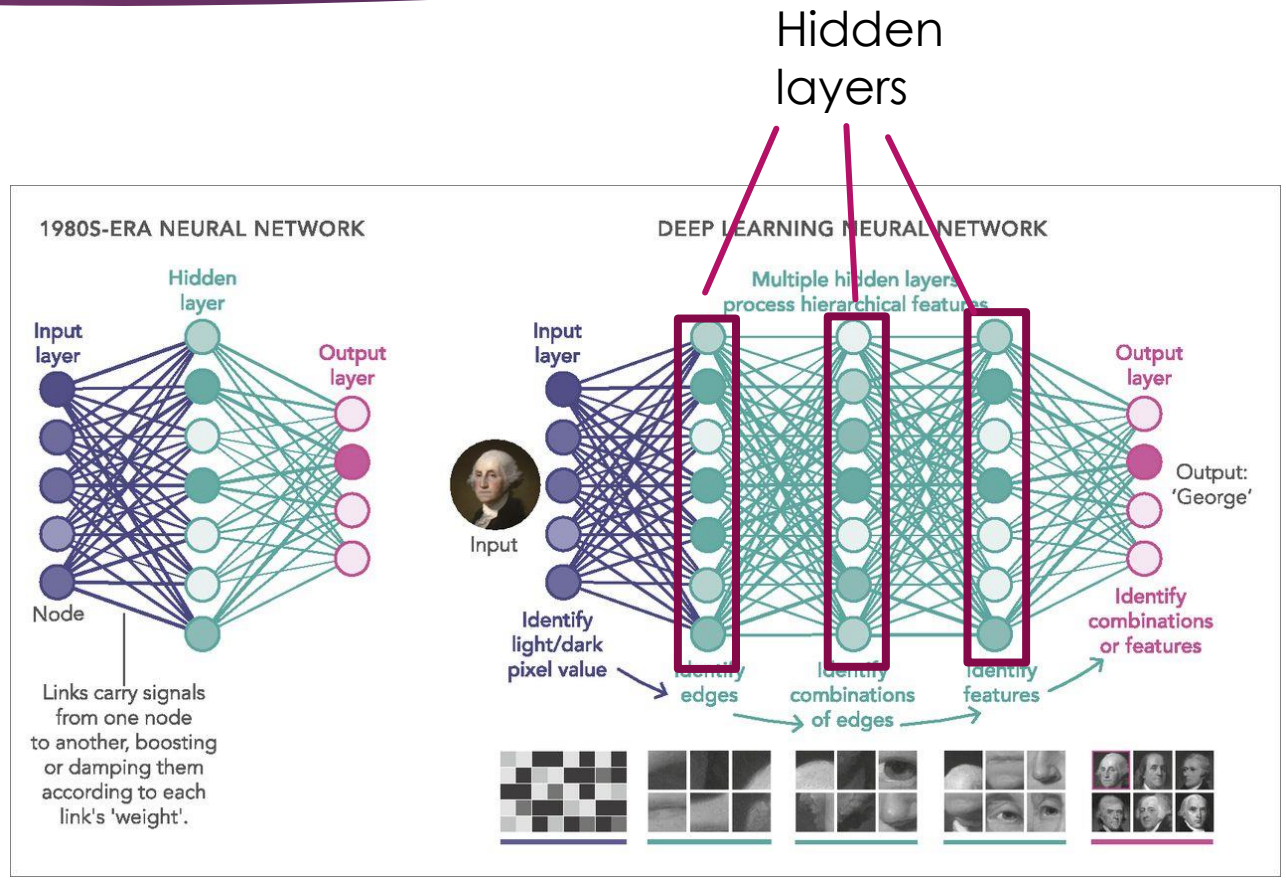


Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

Features of Deep Neural Network

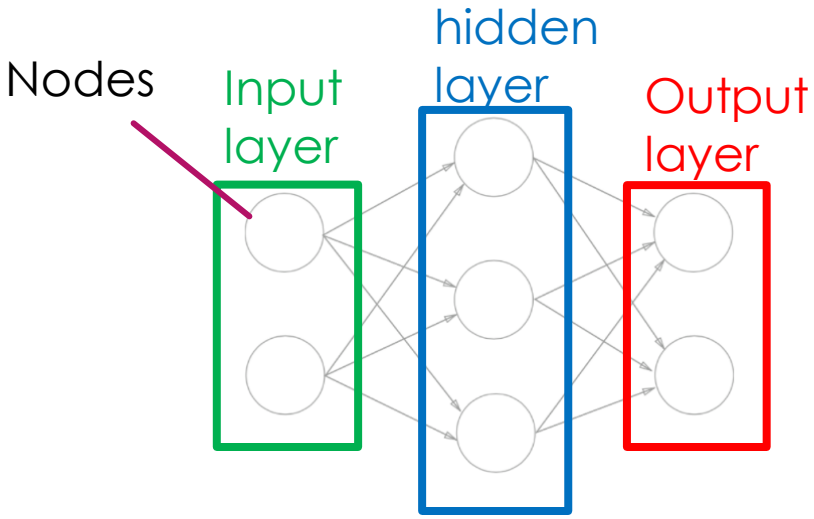
- Features : Multiple hidden layers
  → Expressive power of output
- However, too complex network arises other problems
  Ex) overfitting

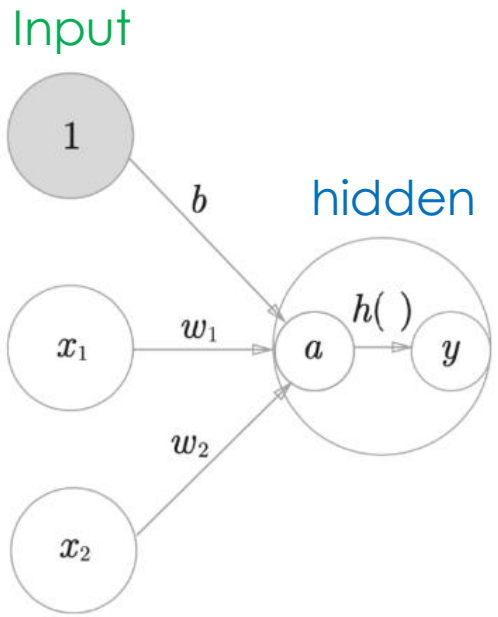
- Contents :
  - Network Construction
  - How to learn the network
  - Confirming how good the neural network is



Neural Network

- Example of network construction

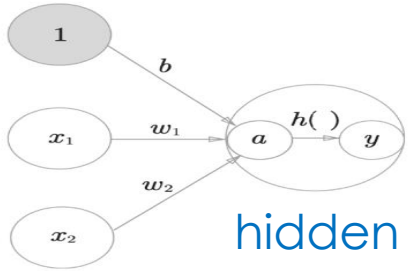


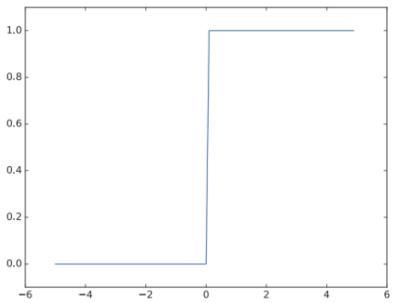


Weight : w
Bias : b
$a = w_1x_1 + w_2x_2$
h() : activation function

Neural Network – Input and hidden

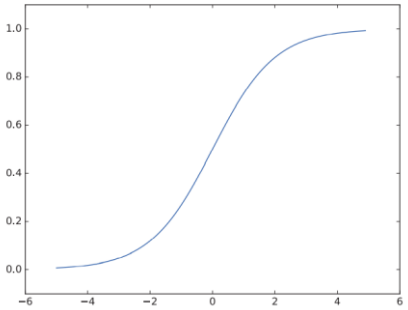
- Weight : Importance on the signals passing through routes.
- Bias : How easy to fire
- Activation function :





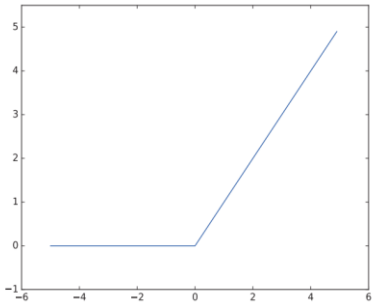
Step function

$$h(x) = \begin{matrix} 0\ (x \leq 0) \\ 1\ (x > 0) \end{matrix}$$



Sigmoid function

$$h(x) = \frac{1}{1+\exp(-x)}$$

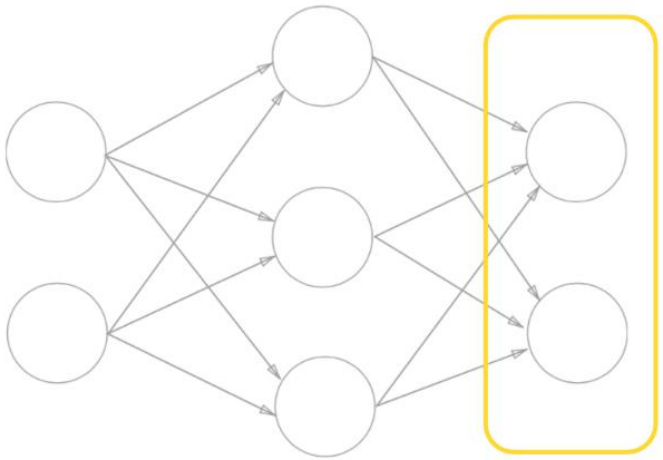


ReLU function

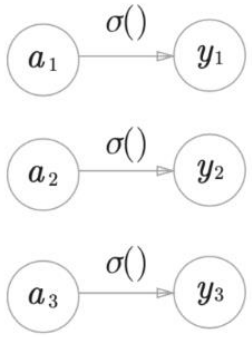
$$h(x) = \begin{matrix} x\ (x > 0) \\ 0\ (x \leq 0) \end{matrix}$$

Neural network - Output

- Output layer uses activation function that is distinct from the one used in the hidden layers.



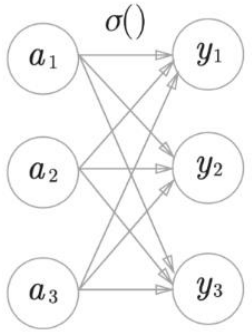
Identify function



Regression problem

Softmax function

$$y_k = \frac{\exp(a_k)}{\sum_{i=1}^{n} \exp(a_i)}$$



Classification problem

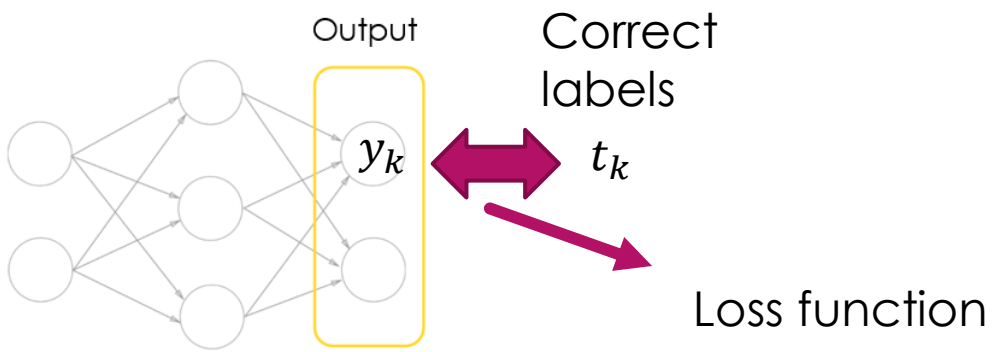
Output is between 0 and 1 →Probability

Learning of Network

- Case: Image analysis on MNIST data and classifying the number
  MNIST (Modified National Institute of Standards and Technology database)

  600000 pieces of training data・100000 pieces of test data・Label data for the numbers written on the image

- Loss function:
  Difference between the output from the network and the correct labels of the training data



Learning of Network

=

Decreasing Loss function

Loss Function and Method of Learning

- Loss function :

Mean squared error : $\mathrm{L} = \frac{1}{2}\sum_k (y_k - t_k)^2$ Regression Problem

Cross entropy error : $\mathrm{L} = -\sum_k t_k \log y_k$ Classification Problem

MNIST classification

- Method of decreasing loss function

Gradient method

Backpropagation method

Optimization -Gradient Method

- The optimal parameter is the value of the parameter when the loss function takes the minimum value.

- Differentiate a loss function and find parameters with 0 gradients

  Replacement of parameter $\mathbf{w}$

$$\mathbf{w} = \mathbf{w} - \eta \frac{\partial \mathrm{L}}{\partial \mathbf{W}}$$

  η: the learning rate

- The right picture : Example of Loss Function is $f(x_0, x_1) = x_0^2 + x_1^2$

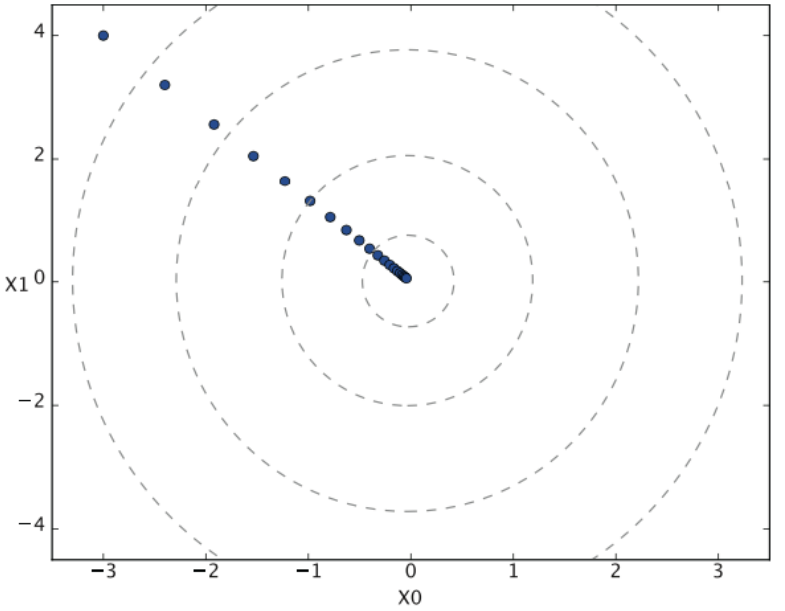
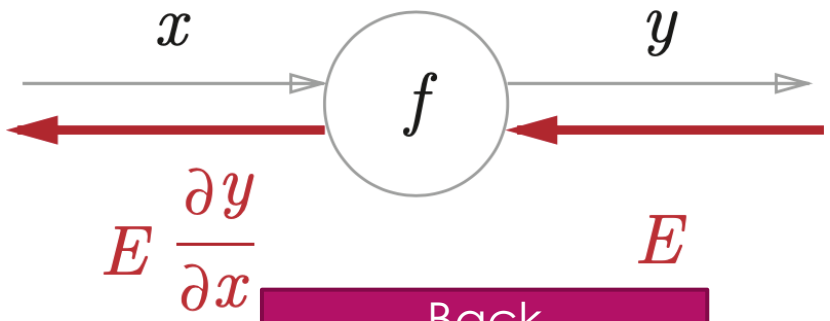


Fig. The process of updating $f\ (x_0\ ,\ x_1)\ =\ x_0^2\ +\ x_1^2$by the gradient method: the dashed line shows the contours of the function.

Optimization -back propagation-

Forward propagation

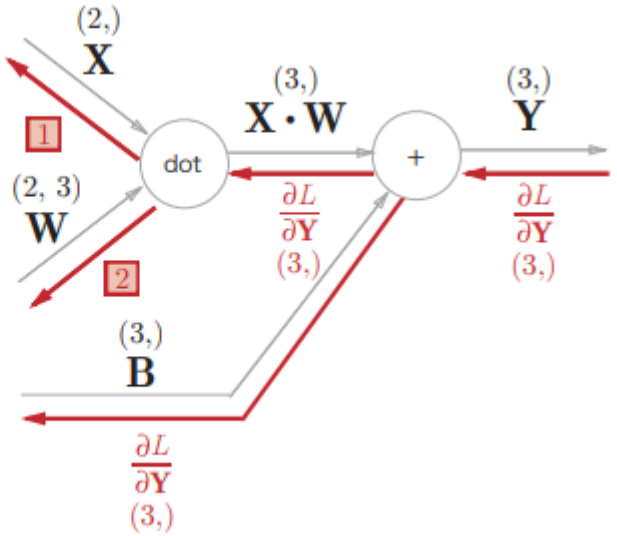


Back propagation

- Computes the gradient of the loss function with respect to the weights of the network for a single input–output
- More efficiently than gradient Method

$$\boxed{1}\ \underset{(2,)}{\frac{\partial L}{\partial \mathbf{X}}} = \underset{(3,)}{\frac{\partial L}{\partial \mathbf{Y}}} \cdot \underset{(3,\ 2)}{\mathbf{W}^{\mathrm{T}}}$$

$$\boxed{2}\ \underset{(2,\ 3)}{\frac{\partial L}{\partial \mathbf{W}}} = \underset{(2,\ 1)}{\mathbf{X}^{\mathrm{T}}} \cdot \underset{(1,\ 3)}{\frac{\partial L}{\partial \mathbf{Y}}}$$



Deep Learning Procedure

1. Splitting the data into training data and test data.
2. Selecting randomly from the training data and forward propagation to the network
3. Updating the parameters (Learning)
4. Inputting test data into trained network and check if network is working properly