Image Classification

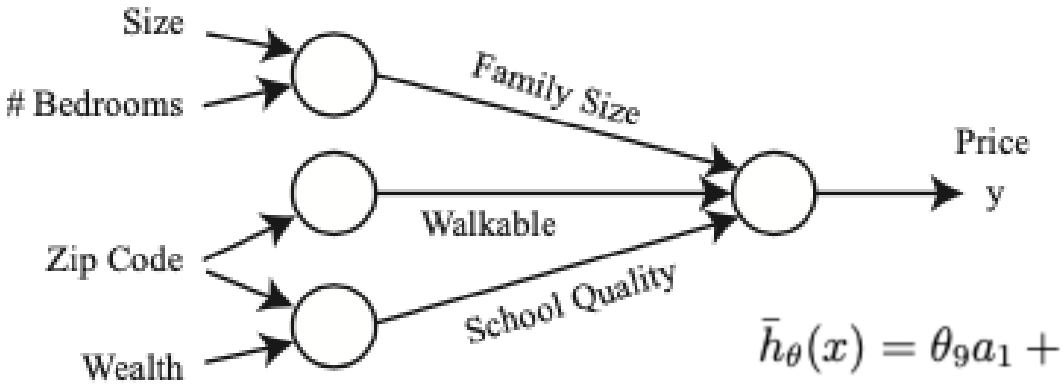
AI/ML Teaching

Goals

- Basic concept of Neural Network
- Dataset – training set, test set
- Training & inference

Small neural network

- $\sigma(\theta^T x) = \frac{1}{1+\exp(-\theta^T x)}$
- $\mathrm{ReLU}(\theta^T x) = \max(\theta^T x, 0)$



$$\bar{h}_\theta(x) = \theta_9 a_1 + \theta_{10} a_2 + \theta_{11} a_3 + \theta_{12}$$

$$a_1 = \mathrm{ReLU}(\theta_1 x_1 + \theta_2 x_2 + \theta_3)$$
$$a_2 = \mathrm{ReLU}(\theta_4 x_3 + \theta_5)$$
$$a_3 = \mathrm{ReLU}(\theta_6 x_3 + \theta_7 x_4 + \theta_8)$$

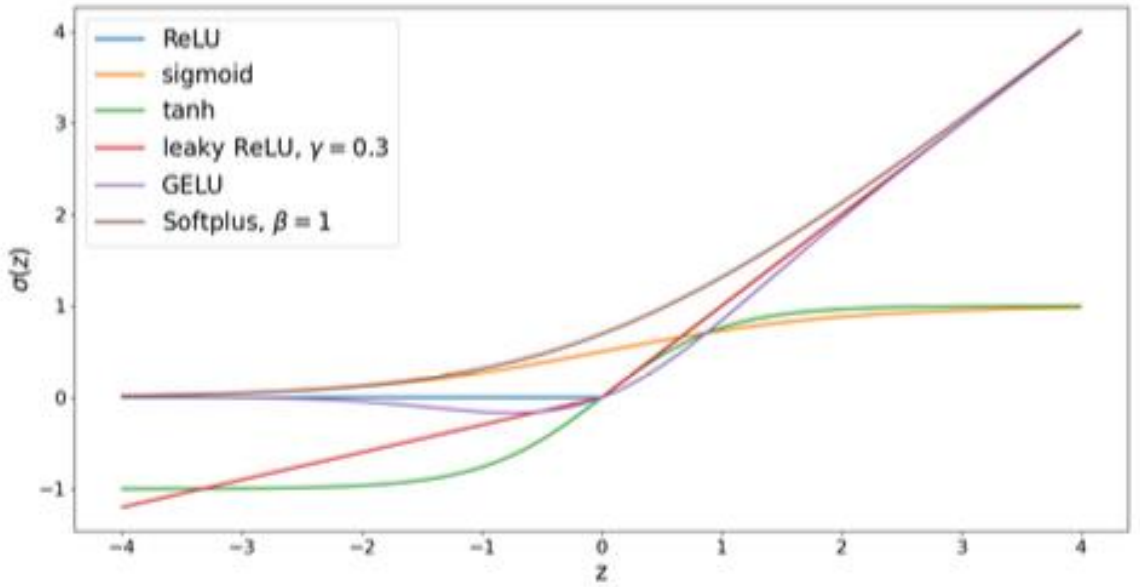
Activation functions

$$\sigma(z) = \frac{1}{1+e^{-z}} \qquad \text{(sigmoid)}$$

$$\sigma(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}} \qquad \text{(tanh)}$$

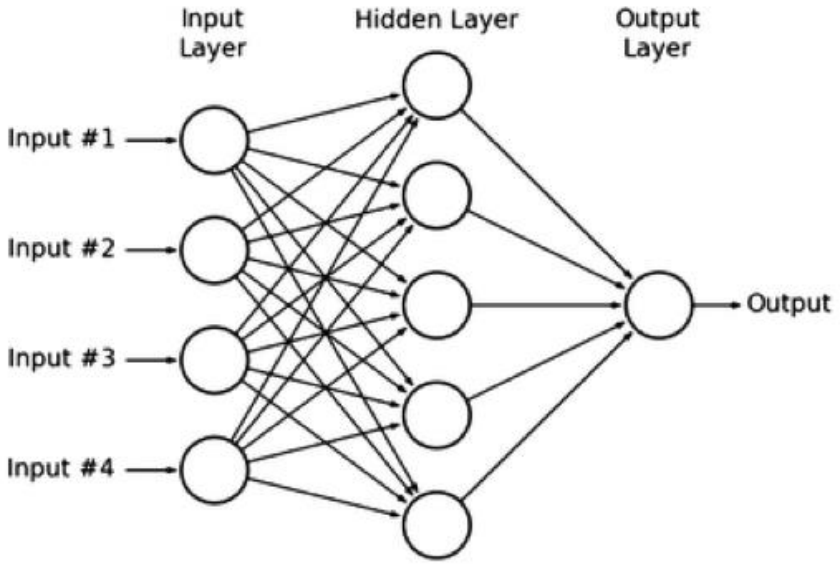
$$\sigma(z) = \max\{z, \gamma z\}, \gamma \in (0,1) \qquad \text{(leaky ReLU)}$$

$$\sigma(z) = \frac{z}{2}\left[1+\text{erf}(\frac{z}{\sqrt{2}})\right] \qquad \text{(GELU)}$$



Deep neural network

- Feed forward network (FFN)
- Multi-layer perceptron (MLP)
- Fully-connected (FC) layer

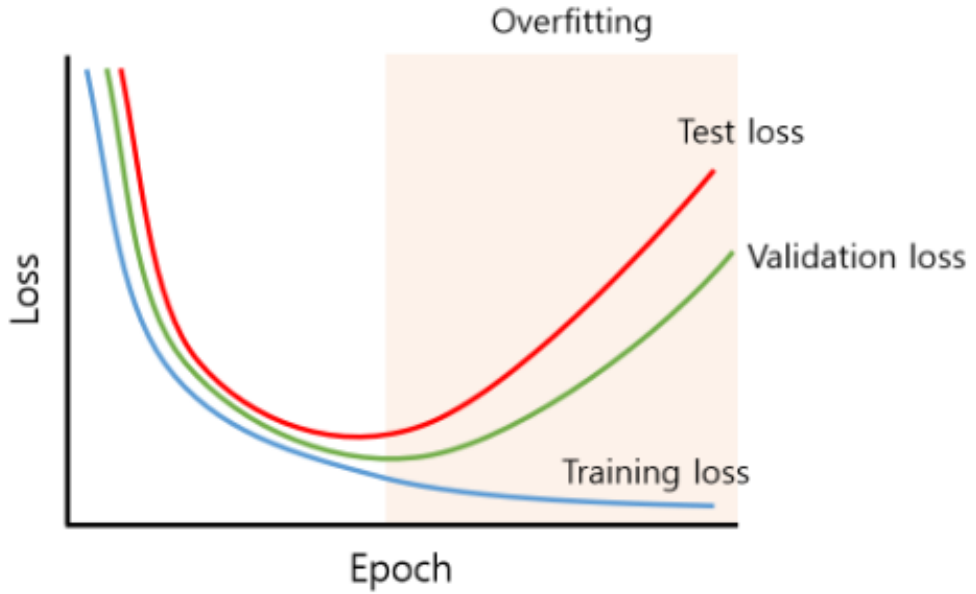


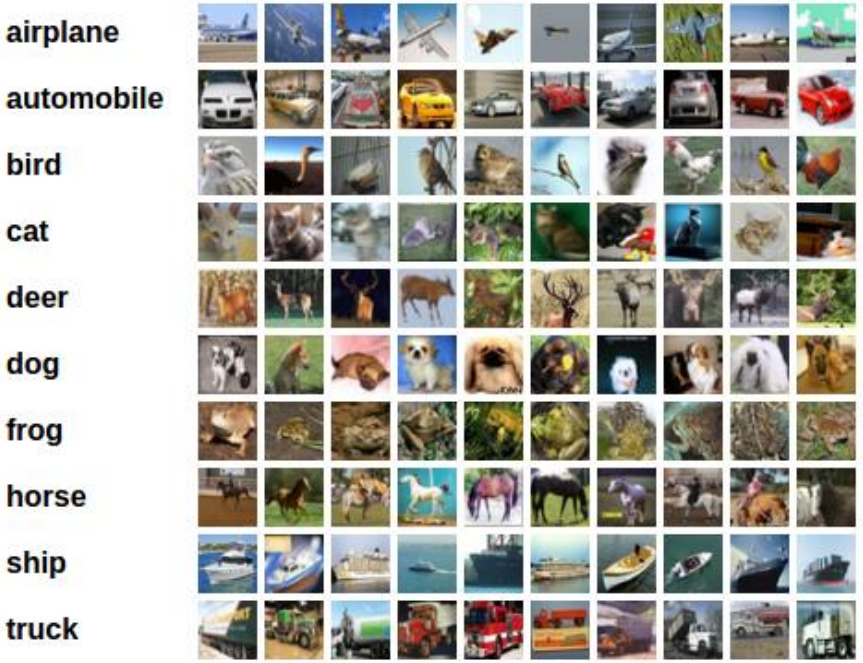
$$W^{[1]} = \begin{bmatrix} — & w_1^{[1]\top} & — \\ — & w_2^{[1]\top} & — \\ & \vdots & \\ — & w_m^{[1]\top} & — \end{bmatrix} \in \mathbb{R}^{m \times d}$$

$$\begin{aligned} a^{[1]} &= \mathrm{ReLU}(W^{[1]}x + b^{[1]}) \\ a^{[2]} &= \mathrm{ReLU}(W^{[2]}a^{[1]} + b^{[2]}) \\ &\cdots \\ a^{[r-1]} &= \mathrm{ReLU}(W^{[r-1]}a^{[r-2]} + b^{[r-1]}) \\ \bar{h}_\theta(x) &= W^{[r]}a^{[r-1]} + b^{[r]} \end{aligned}$$

Dataset

- MNIST: 10 classes of 60000 samples
- Fashion MNIST
- CIFAR10/CIFAR100
- ImageNet-1K: 1.2M samples of 1000 classes

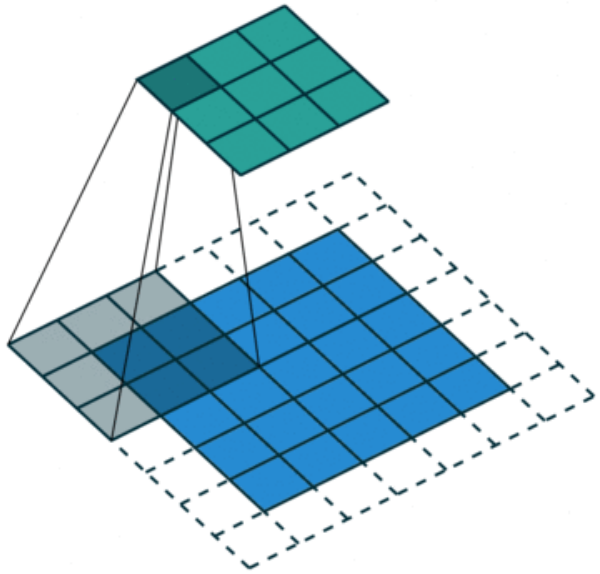


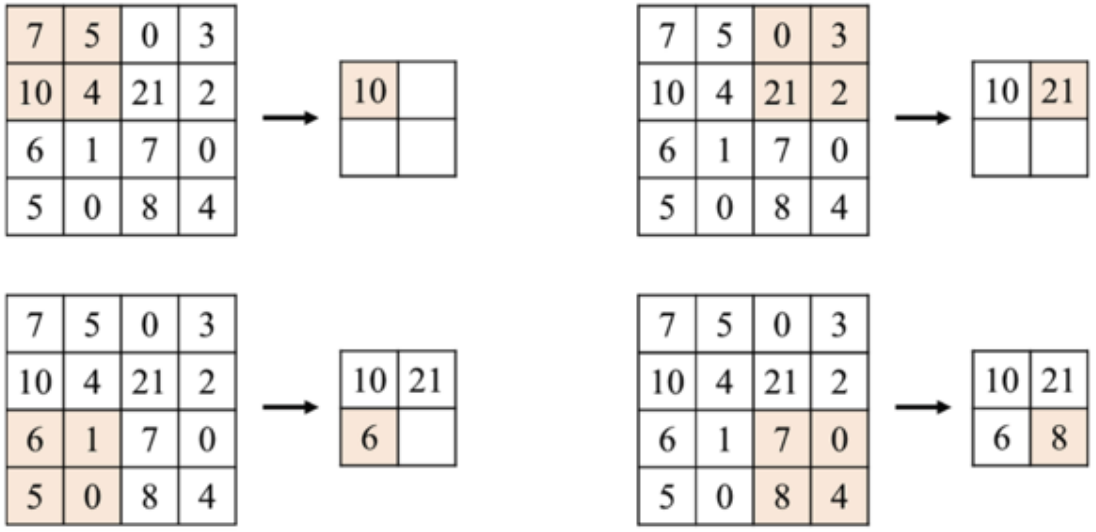

airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck


IMAGENET

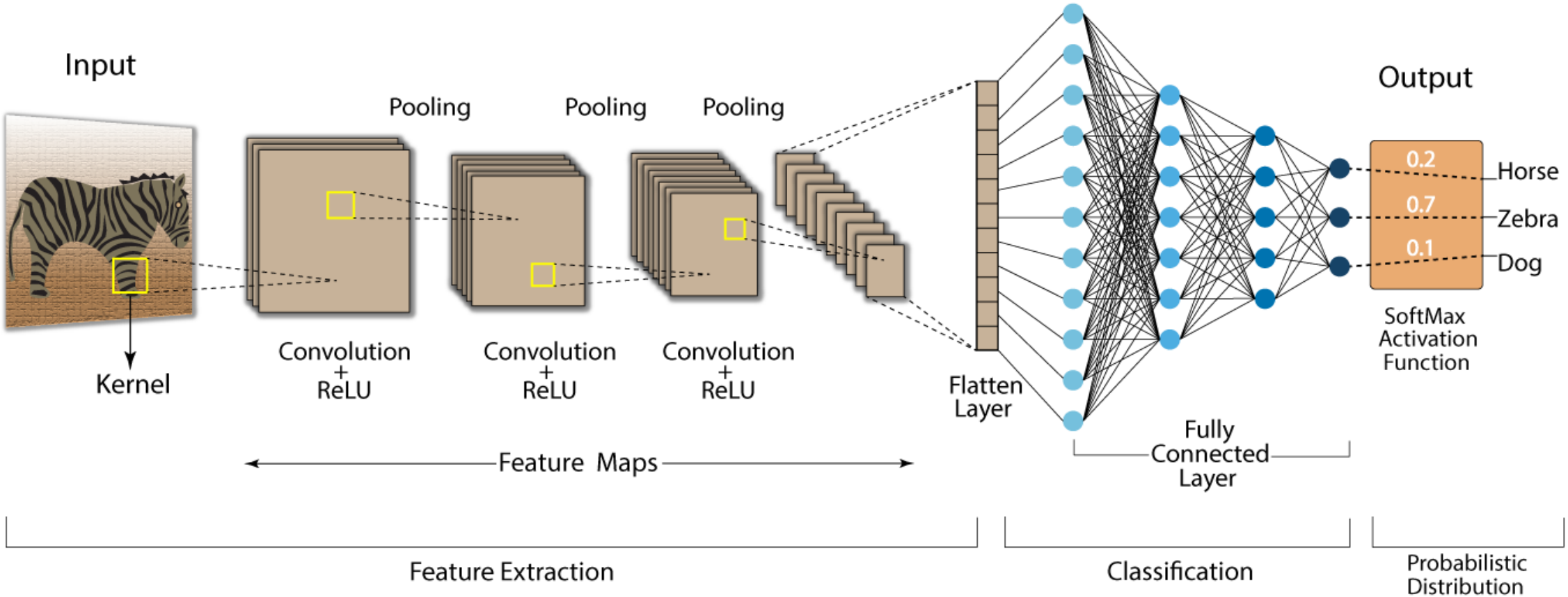
CNN

- Convolutional neural network
  - Kernel/filter
  - Pooling





Convolution Neural Network (CNN)



Reference

- Andrew Ng, CS229 Lecture Notes