

Assignment 10

Ziqi Zhao

November 23, 2019

- 10.1. (a) Most **machine learning** methods work well because of human-designed representations and input features. Machine learning becomes just optimizing weights to best make a final prediction. However, as a machine learning subfield of learning representations of data, **deep learning** algorithms attempt to learn multiple levels of representation by using a hierarchy of multiple layers. It will begin to understand the data and respond in useful ways when it is provided with tons of information.
- (b) DL is useful because (of) ...
- i. Manually designed features are often over-specified, incomplete and take a long time to design and validate.
 - ii. Learned Features are easy to adapt, fast to learn.
 - iii. Deep learning provides a very flexible, universal, learnable framework for representing world, visual and linguistic information.
 - iv. It can learn both unsupervised and supervised.
 - v. Effective end-to-end joint system learning.
 - vi. Utilize large amounts of training data.

- 10.2. (a)

$$\text{output} = \sigma(x_1 \times w_1 + x_2 \times w_2 + x_3 \times w_3 + b)$$

- (b)

$$t_1 = \sigma(1 \times 1 + (-1) \times 2 + 2) = \frac{1}{1 + e} = 0.73$$

$$t_2 = \sigma(1 \times 2 + (-1) \times (-1) - 4) = \frac{1}{1 + e^{-1}} = 0.27$$

$$y = \sigma(t_1 \times 1 + t_2 \times 1 + 1) = \frac{1}{1 + e^{2.00}} = 0.88$$

- 10.3. (a) Training:

$$t_1 = R(-1) = 0$$

$$t_2 = R(4) = 4$$

$$t_4 = R(1) = 1$$

$$y_1 = R(t_1 \times (-1) + t_2 \times 2 + t_4 \times (-4)) = 4$$

$$y_2 = R((t_1 \times 1 + t_2 \times 0 + t_4 \times (-2))) = 0$$

(b) Testing:

$$\begin{aligned}t_1 &= R(0.75 \times 4) = 3 \\t_2 &= R(0.75 \times 4) = 3 \\t_3 &= R(0.75 \times (-1)) = 0 \\t_3 &= R(0.75 \times 1) = 0.75\end{aligned}$$

$$\begin{aligned}y_1 &= R(0.75 \times (t_1 \times (-1) + t_2 \times 2 + t_3 \times 0 + t_4 \times (-4))) = 0 \\y_2 &= R(0.75 \times (t_1 \times 1 + t_2 \times 0 + t_3 \times (-1) + t_4 \times (-2))) = 1.125\end{aligned}$$

10.4. (a)

$$\begin{aligned}y_A &= \sigma(20) = 1 \\y_B &= \sigma(-20) = 0 \\y_C &= \sigma(-20) = 0\end{aligned}$$

(b) Since

$$\begin{aligned}y &= \sigma(z) = \frac{1}{1 + e^{-z}} \\z &= t_1 \times 40 + t_2 \times 40 + t_3 \times 40 - 100\end{aligned}$$

let $y > 0.999$, we have:

$$z > \ln\left(\frac{0.999}{0.001}\right) = 6.91$$

thus,

$$\begin{cases} t_1 = 1 \Rightarrow x_1 \geq 1 \\ t_2 = 1 \Rightarrow x_2 \geq 1 \\ t_3 = 1 \Rightarrow x_1 \times (-10) + x_2 \times (-10) + 300 \geq 7 \Rightarrow x_1 + x_2 \leq 29 \end{cases}$$

Therefore, we obtain the decision boundary:

$$y = \begin{cases} 1, & X \in \mathbb{S} \\ 0, & \text{others} \end{cases}$$

$$\text{where } \mathbb{S} \text{ is an area which satisfies } \begin{cases} x_1 \geq 1 \\ x_2 \geq 1 \\ x_1 + x_2 \leq 29 \end{cases}.$$