CZ4041: Tutorial Week 7

Due on February 25, 2021 at 8:30am

Assoc Prof Pan, Sinno Jialin - CS4

Pang Yu Shao
U1721680D

25/02/2021

Problem 1

On the 21th page of the lecture notes "Lecture 7b", we have shown how to use backpropagation to update parameter of the ANN with one initialization setting for w. Suppose now we initialize w with another set of values: w13 = -1, w14 = -1, w23 = -1, w24 = -1, w35 = -1, and w45 = -1. Run one epoch (i.e., run through the whole training dataset once), to show how the parameters are updated at each iteration.

Solution
Training Record: (0,0,-1) Forward pass:

$$
\begin{aligned}
h_1 &= 0 \\
h_2 &= 0 \\
h_3 &= sign(0 * -1 + 0 * -1) = 1 \\
h_4 &= sign(0 * -1 + 0 * -1) = 1 \\
h_5 &= sign(1 * -1 + 1 * -1) = -1
\end{aligned}
$$

No weights updated.
Training Record: (1,0,1) Forward pass:

$$
\begin{aligned}
h_1 &= 1 \\
h_2 &= 0 \\
h_3 &= sign(1 * -1 + 0 * -1) = -1 \\
h_4 &= sign(1 * -1 + 0 * -1) = -1 \\
h_5 &= sign(-1 * -1 + -1 * -1) = 1
\end{aligned}
$$

No weights updated.
Training Record: (0,1,1) Forward pass:

$$
\begin{aligned}
h_1 &= 0 \\
h_2 &= 1 \\
h_3 &= sign(0 * -1 + 1 * -1) = -1 \\
h_4 &= sign(0 * -1 + 1 * -1) = -1 \\
h_5 &= sign(-1 * -1 + -1 * -1) = 1
\end{aligned}
$$

No weights updated.
Training Record: (1,1,1) Forward pass:

$$
\begin{aligned}
h_1 &= 1 \\
h_2 &= 1 \\
h_3 &= sign(1 * -1 + 1 * -1) = -1 \\
h_4 &= sign(1 * -1 + 1 * -1) = -1 \\
h_5 &= sign(-1 * -1 + -1 * -1) = 1
\end{aligned}
$$

No weights updated.
End of epoch 1

Problem 2

Consider a 2-dimensional dataset for three-class classification by ANN, as shown in Figure 1. Which ANN model as shown in Figure 2 is proper to solve the classification problem? and why?

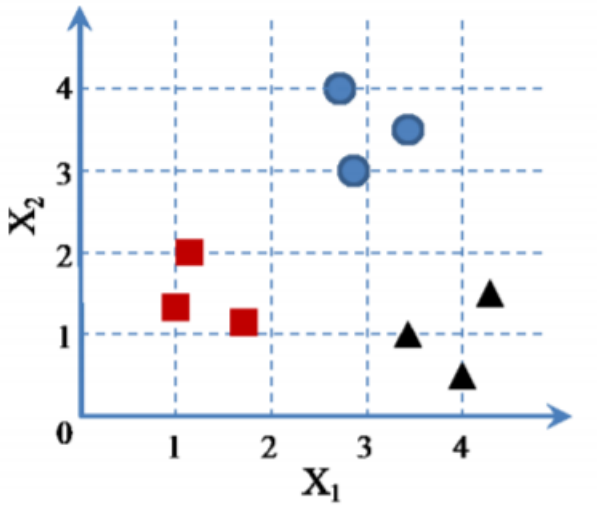
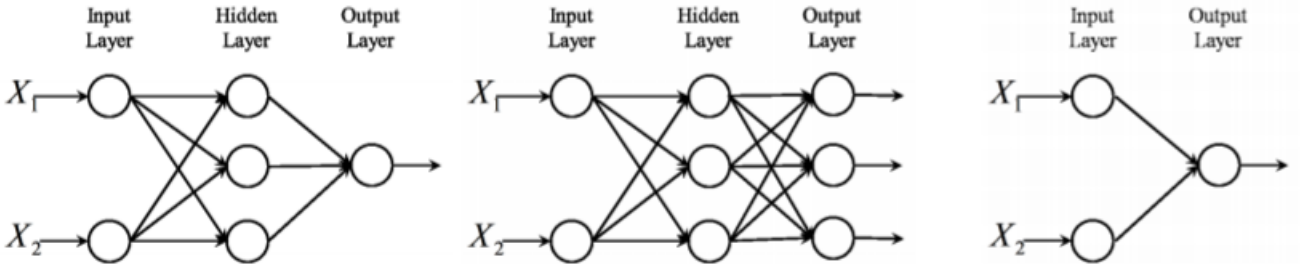


Figure 1: Dataset for Question 2.



(a) ANN model A. (b) ANN model B. (c) ANN model C.

Solution

ANN model B is the proper model to be used as it has 3 output neurons, which is the proper implementation for a three-class classification problem. The output would be probabilities (or prediction) of the input features and the output having the highest value (i.e, highest Probability) would be chosen as the predicted label.

Problem 3

Compute the derivative of the sigmoid function $f(z) = \frac{1}{1+e^{-z}}$ w.r.t z.

Solution
Let $u = 1 + e^{-z}$

$$f(u) = 1/u$$
$$f'(u) = -1/u^2$$

$$\frac{du}{dz} = -e^{-z}$$

$$\begin{aligned}
\therefore f'(z) &= f'(u) * \frac{du}{dz} \\
&= \frac{e^{-z}}{u^2} \\
&= \frac{e^{-z}}{(1+e^{-z})^2} \\
&= \frac{1}{1+e^{-z}}(\frac{e^{-z}}{1+e^{-z}}) \\
&= f(z)(1 - \frac{1}{1+e^{-z}}) \\
&= f(z)(1 - f(z))
\end{aligned}$$