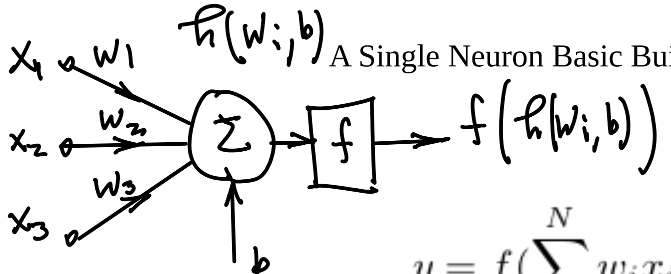
1. Sol:

The Key.

CMPE258
A Single Neuron Basic Building Blocks and Gradient Descent Function
Homework
HL

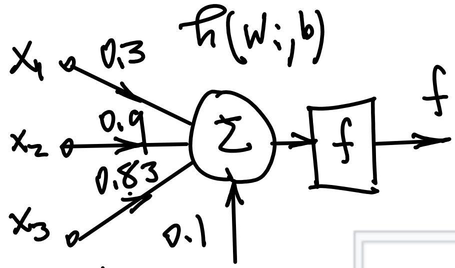


$$y = f(\sum_{i=1}^{N} w_i x_i = W \cdot X + b) = f(h(w_i, b)).$$

From the given condition, we have

Figure 1.

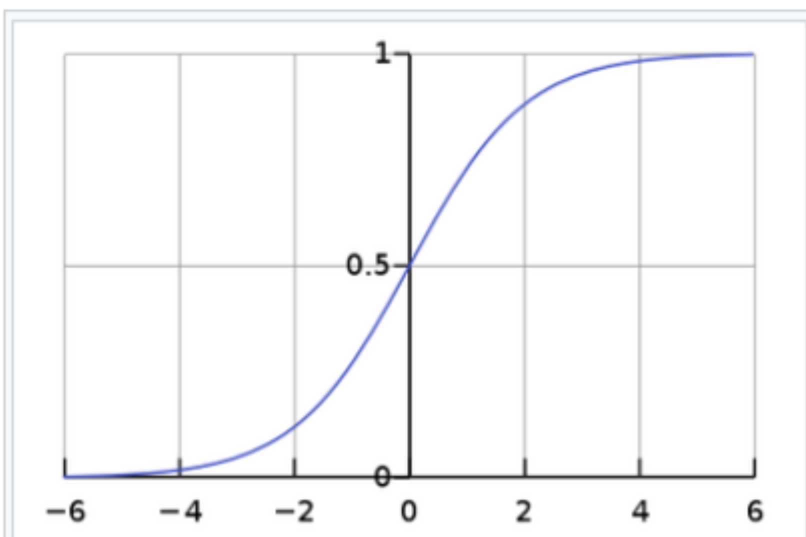
1. Given the equation in Figure 1, design by drawing a single neural, for N=3, and w1=0.3, w2=0.9, w3=0.83, suppose the bias b = 0.1.
2. Based on the equation in Figure 1, explain what is the function h(.), based on the parameters in Question 1 (above), for x1=0.1, x2=14, x3=-7.5, find h=?
3. Suppose we choose the following function for activation function f, find the output of the neuron based on the equation in Figure 1, with the parameters in Question 1 and 2.



$$\boxed{S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}}$$

2. Sol $h(W_i,b)$ is a transfer function;

$$h(W_i,b) = \sum_{i=1}^{3} W_i X_i + b$$
$$= 0.1*0.3 + 14*0.9 + (-7.5)*0.83 + 0.1$$



3. For the Sigmoid Activation Function, from the given condition, we have

https://en.wikipedia.org/wiki/Sigmoid_function

Figure 2. Sigmoid activation function.

$$S(x)\Big|_{x=h(W_i,b)} = \frac{1}{1+e^{-h(W_i,b)}}$$

please be sure to Evaluate Calculate to find its Numerical Value.

$$h(w_i,b) = 0.1*0.3 + 14*0.9 + (-7.5)*0.83 + 0.1$$

4. Define a loss function as follows, answer the following questions:

(4.1) Sol. The Interpretation of $\tilde{y}^j - y^j$: a. $\tilde{y}$ is the Neural Network Output (e.g.) experiment result; b. SuperScript "j" for $\tilde{y}^j$ is the Neural Network Output at Experiment j.

$$L_{total} = \frac{1}{2}\sum_{j=1}^{P}(\tilde{y}^j - y^j)^2. \qquad (23)$$

Since there's no Subscript i, No $\tilde{y}_i^j$, so the Output of the NN has a single output Neuron.

Figure 3.

(4.1) Explain in this equation, what is meaning of the following equation?

(4.2) Sol Summation $\sum_{j=1}^{P}$ over j describes the squared error is Computed for all the experiments (Training) from 1 to P.

$$L = \tilde{y}^j - y^j$$

Figure 4.

(4.3) Sol: Total Loss $= \frac{1}{2}\sum_{j=1}^{3}(\tilde{y}^j - y^j)^2 = \frac{1}{2}[(\tilde{y}^1 - 11.3)^2 + (\tilde{y}^2 - 0.2)^2 + (\tilde{y}^3 - 1)^2]$ //

(4.2) What is the meaning of taking the summation defined in equation (23)?

(4.3) Suppose there are 3 experiments (training) are performed, with the known ground truth as (11.3, 0.2, 1), use the equation in Figure 3, find the total loss. You can use abstract output symbol such as y^2, and y^3 in your result provided you have evaluated y^1 output based on the given parameters in this assignment.

5. Given the equation below

(5.1) Sol: Output function $y = f(h(w_{i,k}, b))$, hence

$$\frac{\partial}{\partial w_{i,k}}\left[\frac{1}{2}\sum_{j=1}^{P}\sum_{i=1}^{M}(f(h(w_{i,k},b)) - y_i^j)^2\right] = \sum_{j=1}^{P}\sum_{i=1}^{M}(f(h(w_{i,k},b)) - y_i^j)\,\frac{\partial f(h(w_{i,k},b))}{\partial h(w_{i,k},b)}\cdot\frac{\partial h(w_{i,k},b)}{\partial w_{i,k}}$$ //

$$\frac{\partial L}{\partial w_{i,k}} = \frac{\partial}{\partial w_{i,k}}\frac{1}{2}\sum_{j=1}^{P}\sum_{i=1}^{M}(\tilde{y}_i^j - y_i{}^j)^2$$

(5.2) Sol: "M" is the No. of Output Neurons. we Can derive this from the double Summation in Fig. 5.; Again "P" is the No. of Experiments //

Figure 5.

(5.1) use output function y defined in this homework, substitute it into the above equation, (you do not have to evaluate this partial derivative here), to check and verify the basic concept of understanding this equation.

(5.2) based on the given condition in this assignment, what is M=? and why? And what is P=? and why?

6. A gradient is defined as

(6.1) Sol for 3 inputs x_i, i=1,2,3, we have 3 weights w_i, i=1,2,3. hence,

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial w_1} \\ \frac{\partial f}{\partial w_2} \\ \frac{\partial f}{\partial w_3} \end{pmatrix};$$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \cdots \\ \frac{\partial f}{\partial x_i} \\ \cdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

(6.2) Sol From Taylor Expansion $f(x_1, x_2) = f(a,b) + \frac{1}{1!}\frac{\partial f}{\partial x_1}(x_1 - a) + \frac{1}{1!}\frac{\partial f}{\partial x_2}(x_2 - b) + R_n(x_1, x_2)$

or

$f(x_1,x_2) \cong f(a,b) + \frac{\partial f}{\partial x_1}(x_1-a) + \frac{\partial f}{\partial x_2}(x_2-b)$, Since $f(x_1,x_2) - f(a,b) = \frac{\partial f}{\partial x_1}(x_1-a) + \frac{\partial f}{\partial x_2}(x_2-b)$

$= f_{x_1}(x_1-a) + f_{x_2}(x_2-b) \quad \dots (*)$,

Figure 6.

(6.1) based on the given conditions in this homework assignment, rewrite the gradient for a single neuron (Hint: with 3 inputs, therefore 3 weights).
(6.2) given the following equation for 2 inputs single neuron, (suppose we denote weights here as x1 and x2, in facts, in our class we use w1 and w2 for weights and x1, x2 for inputs). Explain why the selection of the gradient f will lead the reduction of loss function by using the equation in Figure 8.

if we update $x_1 - a = -f_{x_1}$, $x_2 - b = -f_{x_2}$, as in Eqn (5), then, Eqn (*) becomes

$$(x_1^{k+1}, x_2^{k+1}) = (x_1^k, x_2^k) + [-\eta(\nabla f)^t] \qquad (5)$$

$f_{x_1}(x_1-a) + f_{x_2}(x_2-b) = -(f_{x_1} + f_{x_2})^2$

Figure 7.

Therefore, we have the Error (Loss function) Reduction, e.g. "$f(x_1,x_2) - f(a,b) \le 0$"

$$f(x_1, x_2) - f(a, b) = -(f_{x_1}^2 + f_{x_1}^2) < 0 \qquad (13)$$

(6.3) Sol: "$(w_1^{k+1}, w_2^{k+1}, w_3^{k+1}) = (w_1^k, w_2^k, w_3^k) + (-\eta(\nabla f)^t)$, where $\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_1} \end{pmatrix}$"

Figure 8.

(6.3) Suppose we call the equation in Figure 7 as a training update equation, rewrite it for the given parameters in this assignment (change notation x_i to w_i if needed to reflect on the actual weights in this formulation).

7. Submit your work in one PDF file, then zip it. Use the following file naming convention:
firstName_lastName_SID(last-4-digits)_cmpe258_CondaOpenCV.pdf.
Submit it to the class canvas.

(END)