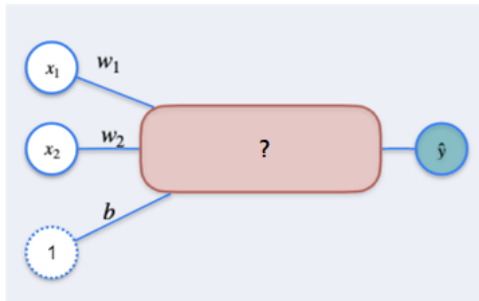


1. Given the Single Layer Perceptron described in the lectures:

1 / 1 point



What should be replaced in the question mark?

- ☐ $w_1 w_2 + x_1 x_2 + b$
- ☐ $w_1 x_1 + w_2 x_2 + b_1 + b_2$
- ☒ $w_1 x_1 + w_2 x_2 + b$
- ☐ $w_1 x_2 + w_2 x_1 + b$

✓ Correct

Correct! In a single layer perceptron, we evaluate a (weighted) linear combination of the inputs plus a constant term, which represents the *bias*!

2. For a Regression using a Single Layer Perceptron, select all that apply:

1 / 1 point

- ☐ The Loss Function used is $L(y, \hat{y}) = -y \ln(\hat{y}) - (1 - y) \ln(1 - \hat{y})$.
- ☒ The Loss Function used is $L(y, \hat{y}) = \frac{1}{2}(y - \hat{y})^2$.

✓ Correct

Correct! This is the mean squared error, usually used as a loss function for regression.

- ☒ To minimize the Loss Function, we consider $L(y, \hat{y})$ as a function of w_1, w_2 and b .

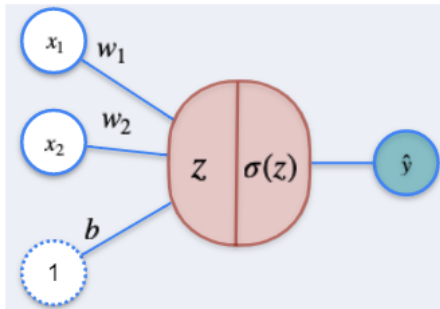
✓ Correct

Correct! We see the Loss Function as a function of w_1, w_2 and b so we can perform Gradient Descent to find the optimal parameters that minimize it!

- ☐ To minimize the Loss Function, we consider $L(y, \hat{y})$ as a function of x_1 and x_2 .

3. Consider the problem of Classification using a Single Layer Perceptron as discussed in the lectures.

1 / 1 point



In the figure above, z and $\sigma(z)$ are, respectively:

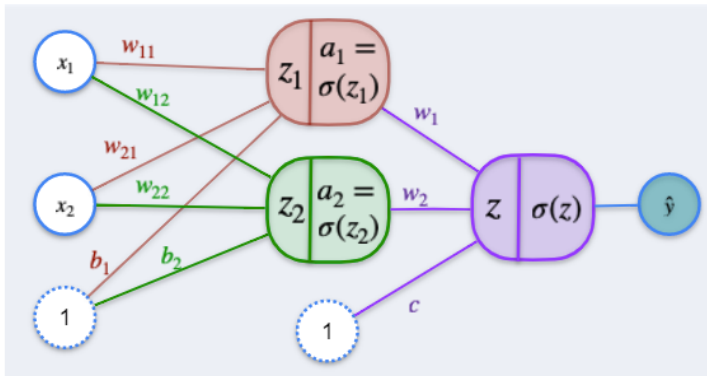
- ☐ $z = w_1x_1 + w_2x_2 + b$ and $\sigma(z) = \frac{1}{2}(z - \hat{z})^2$
- ☐ $z = \frac{1}{1+e^{-z}}$ and $\sigma(z) = w_1x_1 + w_2x_2 + b$
- ☐ $z = x_1 + x_2 + b$ and $\sigma(z) = \frac{1}{2}(z - \hat{z})^2$
- ☒ $z = w_1x_1 + w_2x_2 + b$ and $\sigma(z) = \frac{1}{1+e^{-z}}$

✓ Correct

Correct! In this case, z is a linear combination of the inputs and $\sigma(z)$ is the sigmoid function, so it maps the result to a value between 0 and 1, thus the output can be interpreted as a probability.

4. In the 2,2,1 Neural Network described below

1 / 1 point



How many parameters must be tuned to minimize the Loss Function?

- ☐ 2
- ☐ 3
- ☐ 6
- ☒ 9

✓ Correct

Correct! We have 2 inputs, which will generate 2 constant terms (b_1 and b_2), since the next layer has 2 neurons, each input must have 2 parameters, therefore the first layer has $2 + 2 \times 2 = 6$ parameters. The hidden layer, therefore, has three more parameters since there are 2 neurons. We also must add another constant term c . In total there are 9 parameters.

5. About Backpropagation, check all that apply:

1 / 1 point

- ☐ It is a way to obtain the input values for a given output of a neural network.
- ☒ It is a method to update the parameters of a neural network.

✓ Correct

Correct! This is the method which a neural network updates its parameters.

- ☐ It is the same as gradient descent.
- ☒ It is a method that starts in the output layer and finishes in the input layer.

✓ Correct

Correct! As the name suggests, the backpropagation method iteratively updates the neural network parameters from backwards.