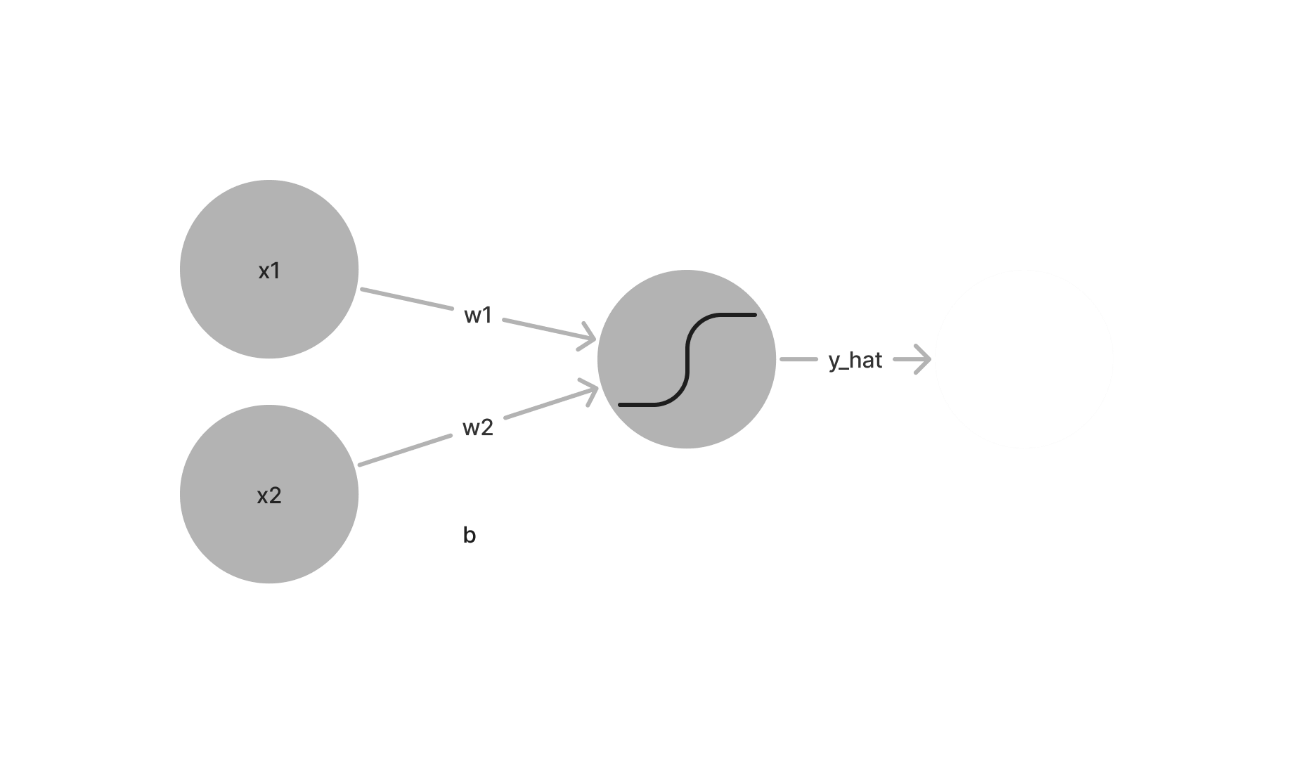
# Tutorial 5

Q1. Refer to slide 61 for the algorithm for computing Stochastic Gradient Descent.

1. Perform a single forward pass.
2. Perform a single backward pass.

Using the weights from the logistic regression setup below:



Given,

*x1 = 1, x2 = 2*

*w1 = 1, w2 = 1*

*b1 = 1*

*ŋ = 0.2*

*y = 2*

*output = sigmoid (w1 x1 + w2 x2+ b )*

Q2. You are using Stochastic Gradient Descent Optimizer to train your Logistic Regression. Consequently, the gradients leading to the parameter updates are computed on a single training example.

For one training example *x = (x1, x2, … , xn)* of dimension *n*, the forward propagation is:

*z = wx + b*

*ŷ = a = σ(z)*

*L = - (ylog(ŷ) + (1-y) log(1- ŷ) )*

1. Compute the derivatives of loss function with respect to *w* and *b.*
2. Prove that *w=w- ŋ (a-y) XT and b=b- ŋ (a-y)*

Hint, use the chain rule:

# Answer Key

a) Forward pass:

*w.x + b =* 1.1 + 1.2 + 1 = 4

b) Backward pass:

The two derivatives,

Then,

*Src:* [section3\_soln.pdf (stanford.edu)](http://cs230.stanford.edu/fall2018/section_files/section3_soln.pdf)