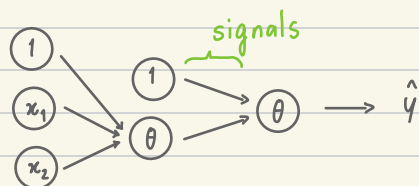


Neural Network

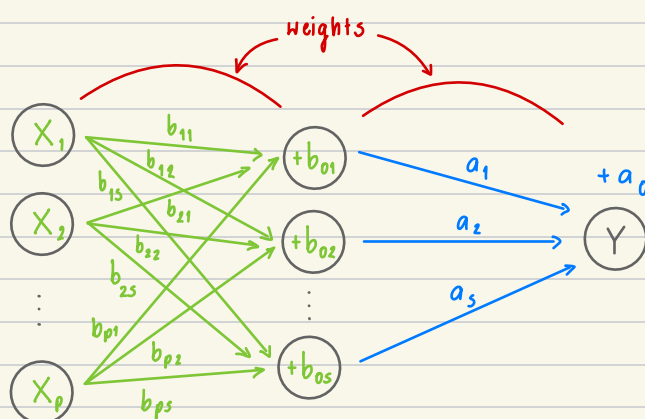
The neural network structure :



θ is a transformation or activation function

θ is at every node with an input

For classification problem $\hat{y} = \theta(s)$, where s is a signal, $\theta(s) = \tanh(s)$, logistic(s), etc.



$$\hat{y} = a_0 + \sum_{k=1}^s \left[a_k \left(b_{0k} + \sum_{j=1}^p b_{jk} x_j \right) \right]$$

$+a_0, +b_{01}, +b_{02}, \dots, +b_{0s}$ are constant terms (biases)

Training process

- **epochs** : number of times our network have seen the model
 higher # of epochs lead to better accuracy and also, lead to overfitting
 lower # of epochs, the model may not learn the underlying patterns (underfitting)
- **batch-size** : number of samples used in one forward and backward pass through the network
 larger batch size, faster, but lower accuracy and overfitting
 smaller batch size, longer (computationally expensive), better accuracy

Iterations : number of batches required to complete one epoch

$$\frac{\text{total number of samples}}{\text{batch size}} = \# \text{ of iterations}$$

number of iterations is equal to number of updates made to the weights of the model per epoch

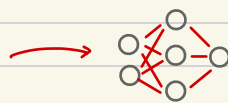
Example

200 observations in training dataset

5 epochs

50 batch_size

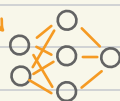
50	50
50	50



first iteration

- 1) → feed forward the input
- 2) ← back propagation

1 update on the weights



second iteration

- 1) → feed forward the input with the updated weights
- 2) ← back propagation

⋮

⋮

A total of 4 iterations

→ = 1 epoch

Repeated above 5 times