

Neural Network

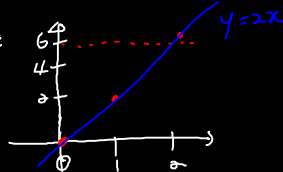
Week1: Basics

Supervised Learning: Need a $x - y$ data set \rightarrow most of Machine Learning (ML)

Example: Decision tree, neural net, SVM, Bayes, etc.

$$X = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \quad Y = \begin{bmatrix} 0.6 \\ 1.49 \\ 6.2 \end{bmatrix}$$

$$Y = 2x + e$$

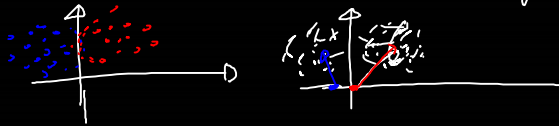


$$X = \begin{bmatrix} \text{img0} \\ \text{img1} \end{bmatrix} \quad Y = \begin{bmatrix} \text{"cup"} \\ \text{"cat"} \end{bmatrix} \rightarrow \begin{bmatrix} \text{"0"} \\ \text{"1"} \end{bmatrix}$$

8 bit unsigned integers.
0 ~ 255
= $w \cdot h + \text{img0}, x_0$

Unsupervised Learning: Need a x data set, no need for label. However, you need a measure to tell whether the output is proper.

Example: genetic algorithm, KNN \rightarrow k-nearest neighbor



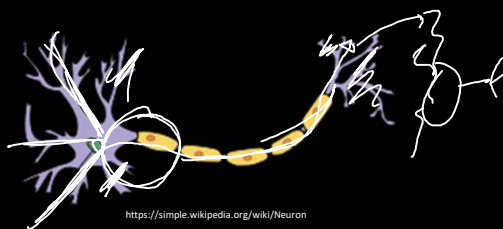
wish i knew

Neural Network

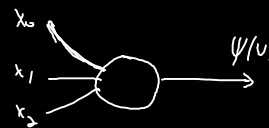
Week1: Basics

Perceptron

Neuron


<https://simple.wikipedia.org/wiki/Neuron>

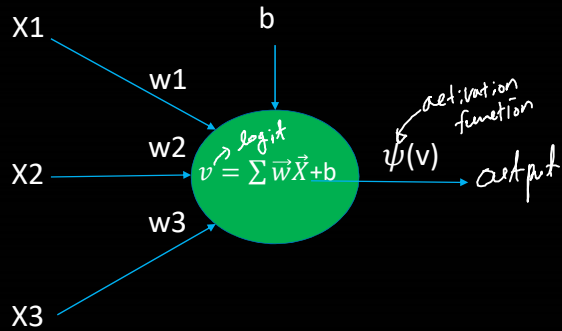
McCulloch-Pits Perceptron



wish i knew

Neural Network

Week1: Basics

Perceptron: a basic unit for neural net

Math notation:

$$\begin{aligned}
 v &= x_1 w_1 + x_2 w_2 + x_3 w_3 + b \\
 &= \sum_{i=0}^n x_i w_i + b \\
 &= x_1 w_1 + x_2 w_2 + x_3 w_3 + 1 \cdot w_0 \\
 &= [1 \ x_1 \ x_2] \begin{matrix} \nearrow x_0 \\ \nwarrow w_0 \end{matrix} [w_0 \ w_1 \ w_2 \ w_3]^T \\
 &= \vec{x} \cdot \vec{w}^T \\
 \psi(v) &= \begin{cases} 1 & v \geq 0 \\ 0 & \text{o.w.} \end{cases}
 \end{aligned}$$

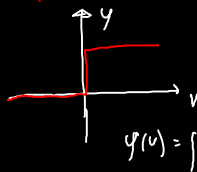
wish i knew

Neural Network

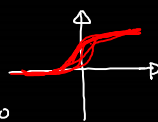
Week1: Basics

Activation Functions

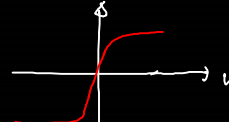
Step Function



Sigmoid Function



Hyperbolic Tangent Function



Relu Function



wish i knew

Neural Network

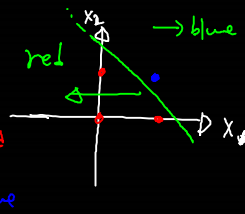
Week1: Basics

Perceptron

Linear Separability

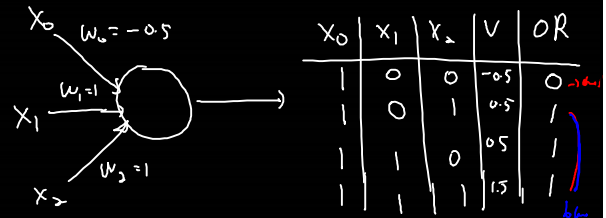
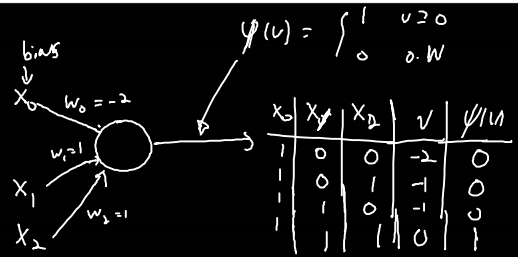
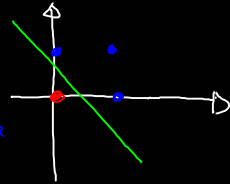
AND Gate

| x_1 | x_2 | AND |
|-------|-------|-----|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |



OR Gate

| x_1 | x_2 | OR |
|-------|-------|----|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |



wish i knew

Neural Network

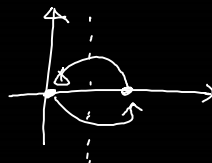
Week1: Basics

Perceptron

Linear Separability

NOT Gate

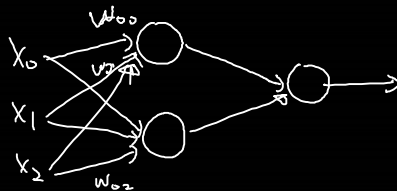
| x_1 | NOT |
|-------|-----|
| 0 | 1 |
| 1 | 0 |



| x | v |
|-----|-----|
| 0 | 1 |
| 1 | 0 |

$\psi(v) = \begin{cases} 1 & v \geq 0 \\ 0 & \text{otherwise} \end{cases}$

XOR Gate



wish i knew

Neural Network

Week1: Basics

Perceptron

Perceptron Learning Rule:

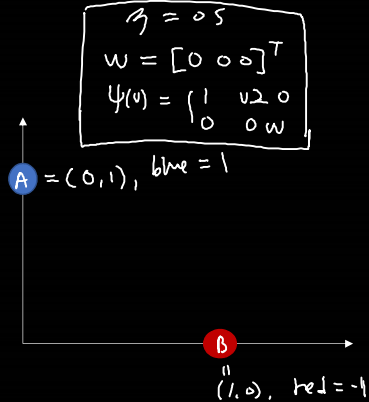
$$w(n+1) = \overline{w(n)} + \Delta w \leftarrow \text{change of } w \text{ val.}$$

$$= \overline{w(n)} + \eta(y - \hat{y})x(n)$$

next weight values
current w. val.

Code: HW

$\eta = \text{learning rate, eta}$
 $y - \hat{y} = \text{desired} - \text{predicted, Error}$
 $x(n) = \text{current input}$ *caused*



| | X | W | V = wx | y | \hat{y} | Error $y - \hat{y}$ | dw | Next w |
|----------------|---|---|--------|----|-----------|------------------------|--|---|
| A=(0,1) b=1 | $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ | 0 | 1 | 1 | 0 | $\Delta w = \eta \cdot E \cdot \text{input}$ $= 0.5 \cdot (0) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ $= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ | $\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$ |
| B=(1,0) b=1 | $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ | 0 | -1 | 1 | -2 | $\Delta w = 0.5 \cdot (-2) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ $= \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$ | |

$$\Sigma \Delta w = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \rightarrow w(n)$$

wish i knew

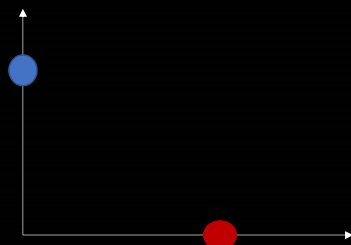
Neural Network

Week1: Basics

Perceptron

Code: HW

Perceptron Learning Rule: $w(n+1) = w(n) + \Delta w$
 $= w(n) + \eta(y - \hat{y})x(n)$



| | X | W | V = wx | y | \hat{y} | Error $y - \hat{y}$ | dw | Next w |
|----------------|---|---|-----------------------|----|-----------|------------------------|---|---|
| A=(0,1) b=1 | $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$ | $V = -1 + 0 + 0 = -1$ | 1 | -1 | 2 | $\Delta w = 0.5 \cdot (2) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ $= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ | $w + \Delta w$ $\begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ |
| B=(1,0) b=1 | $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$ | $V = -1 - 1 = -2$ | -1 | -1 | 0 | $\Delta w = 0$ | $= \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$ |

$$\Sigma \Delta w = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

wish i knew

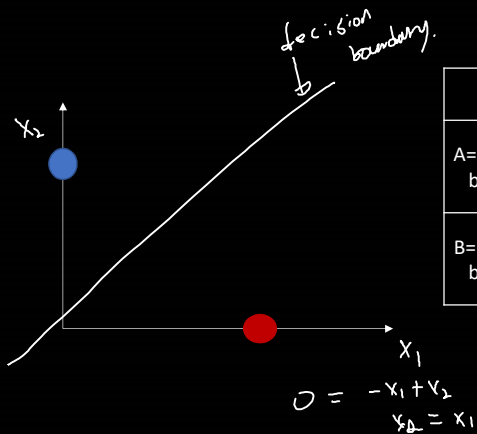
Neural Network

Week1: Basics

Code: HW

Perceptron

Perceptron Learning Rule: $w(n+1) = w(n) + \Delta w$
 $= w(n) + \eta(y - \hat{y})x(n)$



| | X | W | V = wx | y | \hat{y} | Error $y - \hat{y}$ | dw | Next w |
|----------------|---|--|----------|----|-----------|------------------------|----|--------|
| A=(0,1) b=1 | $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ | $\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$ | $v = 1$ | 1 | 1 | 0 | 0 | X |
| B=(1,0) b=1 | $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ | $\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$ | $v = -1$ | -1 | -1 | 0 | 0 | |

$$V = X \cdot w = \begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad w = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad \text{bias} = 0$$

$$= -x_1 + x_2$$

$$\psi = \begin{cases} v \geq 0 \\ v < 0 \end{cases}$$

wish i knew