□□□□

□□□□□Supervised learning□

□□□□□x□y□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□

□□

alt text

- □□□□□□□□□□□□□
- □□□□□□□□□□□□(□□)□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□

□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□y□□□□□□□□□□□□□□□□□□$\hat{y}$□□□□□□□□□
- $f(x) = wx + b$

□□□□

- □□□□□□□□□□□□□□□□□□$J = \frac{\sum_{i=1}^{m}{(\hat{y^(i)} - y^(i))^2}}{2m}$
- □□□□□w□b□□□,□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□□□□ alt text

---

alt text

□□□□

- □□□

alt text * □□ ####### □□□□□□□ $w = w - \alpha\frac{\partial}{\partial w}J(w,b)$ $b = b - \alpha\frac{\partial}{\partial b}J(w,b)$ * □α(□□□)□□□ * □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ * □□□□□□□□□□□□□□□□□□□□□□□□□□ * □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ *

alt text

□□□□□□□□□□□□□□□□□□□□□□ alt text

####### □□□□□□□□□□□□□□ □□□□□□□□□□ $w = w - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)})x^{(i)}$ $b = b - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)})$

□□□□□□

####### □□□□ □□□□□□□□□y□□□□□□□□□□□□□□□□ □□□□□□□$f_{w,b}(x) = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + b$ □□$\vec{w} = [w_1 , w_2 , w_3 , w_4]$ $\vec{x} = [x_1 , x_2 , x_3 , x_4]$ □□□□□□□ $f_{\vec{w},b}(\vec{x}) = \vec{w}\cdot\vec{x}$

####### □□□ □□numpy□□□□□□□□□□□□□ alt text

####### □□□□□□□□□□□□□□□□□□□ □□□□□$J(\vec{w},b)$ □□□□□ $$ w_j = w_j - \alpha\frac{\partial}{\partial w_j}J(w_1,\cdots,w_n,b) \ b = b - \alpha\frac{\partial}{\partial b}J(w_1,\cdots,w_n,b)\ w_j = w_j - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)})x_j^{(i)}\ b = b - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)}) $$

####### □□□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ alt text □□□

1. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□x□□□□
2. □□□□$x_1 = \frac{x_1 - \overline{x}}{x_max - x_min}$

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ alt text

####### □□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- alt text
- □□□□□

□□□□(□□□□□□)

- sigmoid□□□$g(z) = \frac{1}{1 + e^{-z}}$ 0 < g(z) < 1
- □□sigmoid□□□□□□□□□□(0,1),□□□□□□□□

□□□□□□

- $f_{w,b}(\vec{x}) = \frac{1}{1 + e^{-(\vec{w}\vec{x} + b)}}$
- □□□□□□□□□□□□□□□□□□□□□□□□
- □□□□:□□□□□□□□□

□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□ alt text

□□□□□□□□□□□□□□□□□□□□□□□□ alt text

□□□□□□□1□□□□□□□□□□f->1□□□□□□□1□□y□□□□□□□□□□0□□□□□□□□□□ alt text

---

□□□□□□□□□□□□□□□□□□□□**(□□□□)**□□□□ □□□□□ $L(f_{\vec{w},b}(\vec{x}^{(i)}),y^{(i)}) = -y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)})) - (1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$ □□□□□$J(\vec{w},b) = -\frac{1}{m} \sum_{i=1}^{m}y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)})) + (1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$

[illegible]

[illegible]**([illegible])**[illegible]**f**[illegible]**f**[illegible]**f=wx+b**[illegible]**sigmoid**[illegible] [illegible] $$ w_j = w_j - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)})x_j^{(i)}\ b = b - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)}) $$

[illegible]

- [illegible]([illegible])
- [illegible]([illegible])
- [illegible]
  1. [illegible]
  2. [illegible]
  3. [illegible]

[illegible]([illegible])

[illegible]

- [illegible]λ
- [illegible]$\frac{\lambda}{2m}\sum_{j=1}^{n}(w_j)^2$
- [illegible]$J(\vec{w},b) = \frac{1}{2m}\sum_{i=1}^{m}(f_{\vec{w},b}(\vec{x}^{(i)}-y^{(i)})^2 + \frac{\lambda}{2m}\sum_{j=1}^{n}(w_j)^2$ [illegible]**w**[illegible]**0**[illegible]**w**[illegible] [illegible]**w**[illegible]**w**[illegible] $$ w_j = w_j - \alpha(\frac{1}{m}\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)})x_j^{(i)} + \frac{\lambda}{m}w_j)\ b = b - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)}) $$

[illegible]([illegible])

[illegible]$J(\vec{w},b) = -\frac{1}{m} \sum_{i=1}^{m}y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)})) + (1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) + \frac{\lambda}{2m}\sum_{j=1}^{n}w_j^2$ $$ w_j = w_j - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)})x_j^{(i)} +\frac{\lambda}{m}w_j\ b = b - \frac{1}{m}\alpha\sum_{i=1}^{m}(f_{w,b}(x^{(i)}) - y^{(i)}) $$

[illegible]([illegible])

alt text

[illegible]

[illegible]

[illegible] [illegible] alt text

softmax[illegible]

$$ z_1 = w_1x + b_1\ z_2 = w_2x + b_2\ z_3 = w_3x + b_3\ \ \ a_1 = \frac{e^{z_1}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y = 1|\vec{x})\ a_2 = \frac{e^{z_2}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y = 2|\vec{x})\ a_3 = \frac{e^{z_3}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y = 3|\vec{x}) $$

softmax[illegible]

alt text

---

softmax[illegible] [illegible]

□□□□□

□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□

□□□□□□Unsupervised Learning□

□□□□□□□x□□□□□□y

□□□□

- □□□□□□□□□□□□□□□□□□ alt text

K-means□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ alt text

□□□□

- $c^{(i)}$□□□□□□□□□□□
- μ_k□□□□□□□k□□□
- $\mu_{c^{(i)}}$□□□□□□□□□□□□□□□□□

□□□□**(□□□□)** $J = (c^{(1)},\dots,c^{(m)},\mu_1,\dots,\mu_k) = \frac{1}{m}\sum_{i=1}^{m}||x^{(i)}-\mu_{c^{(i)}}||^2$

- m□□□□□□□□□x^i□□□□
- □□□□□□□
  - □□□□□□□□□□□□□□c_i□□
  - □□□□□□□□□□□□□□□□□u_i□□

□□□□

□□ □□□□□□□□□□□□□□□□□□□□□

□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□

- □□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□

□□□□

- □□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□

- □□□□□□□□□□□□

- □□□□□□□□□□□□□□□□ alt text

□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□

□□□□□□(□□□□)

- □□□□□□□□□□□□□□

□□□□

□□□□□□w□□□□(□□□)□b□□□□(□□□□)□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□ □□□□□□□□□□□□□□□□□□ □□□□□

- □□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□(□□)□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□

- □□
- □□□ alt text

□□□□□□□□□□□□□□□□□□□□□□□□□□□□ alt text

□□

□□ □□□□□□□□□□□□□□□□□□□

□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□

□□□□□□□□□□

tensorflow

□□**sequential**□□□□□□□□□□□ alt text

□□**tensorflow**□□□□□□□□□□

1. □□□□
2. □□□□□□□□□□□□□
3. □□□□ alt text

□□□□□□□□□

□□□□□□□□□□□□**numpy**□□□□□□□□□□□□□□ alt text

[illegible]

- [illegible]$g(z) = z$
- sigmoid[illegible](0,1),[illegible]
- ReLU[illegible]$g(z) = max(0,z)$

[illegible]

Adam[illegible]([illegible])

[illegible]

1. [illegible](torch.optim)
2. [illegible]
3. [illegible]

[illegible]

- [illegible]
- [illegible]

[illegible]

[illegible]([illegible]),[illegible] alt text

[illegible] [illegible] alt text

[illegible]

- [illegible]([illegible])[illegible]
- [illegible]([illegible])[illegible]

[illegible]**Jcv**[illegible] alt text

[illegible]

[illegible] [illegible] alt text

[illegible]

[illegible] [illegible] alt text

[illegible] [illegible] alt text

[illegible] [illegible] alt text

[illegible] alt text

[illegible]

[illegible]

- [illegible]
- [illegible]
- [illegible]
- [illegible]([illegible])[illegible]

[illegible]

[illegible] [illegible] alt text

- [illegible]
- [illegible]
- [illegible]$\frac{[illegible]}{[illegible]+[illegible]}$
- [illegible]$\frac{[illegible]}{[illegible]+[illegible]}$ [illegible]
- [illegible]
- [illegible]
- [illegible]F1[illegible]([illegible])[illegible] $F1_SCORE = \frac{1}{\frac{1}{2}(\frac{1}{P} + \frac{1}{R})} = 2\frac{PR}{P+R}$

[illegible]

- [illegible]
- [illegible]([illegible])[illegible]

[illegible] [illegible] alt text [illegible] $$ p_0 = 1 - p_1\ H(p_1) = -

p_1log2(p_1)-p_0log2(p_0)\ = -p_1log2(p_1)-(1-p_1)log2(1 - p_1) $$

[illegible] [illegible]([illegible])[illegible] alt text

p[illegible]**w**[illegible] [illegible] $$ H(p_1^{root}) - (w^{left}H(p_1^{left} + w^{right}H(p_1^right))) $$

[illegible] [illegible]**k**[illegible]**k**[illegible]**0**[illegible]**1**[illegible] [illegible]**0**[illegible]**1**[illegible]

- [illegible]

[illegible]

[illegible]**Y**[illegible] [illegible]([illegible])[illegible] alt text

[illegible]**(**[illegible]**)**

[illegible]

[illegible]([illegible])

- [illegible]
- [illegible]
- [illegible]

[illegible]

- [illegible]([illegible]100[illegible])
- [illegible]k[illegible]k[illegible]

XGBoost

- [illegible]
- [illegible]
- [illegible]

[illegible]

- [illegible]
- [illegible]

[illegible]

- [illegible]
- [illegible]

[illegible]

- [illegible]([illegible])
- [illegible]

[illegible]

- $r(i,j)$[illegible]1[illegible]
- $y^{(i,j)}$[illegible]
- $w^{(j)},b^{(j)}$[illegible]
- $X^{(i)}$[illegible]i[illegible]
- $m^{(j)}$[illegible]

[illegible]$w^{(j)}\cdot{X^{(i)}}+b^{(j)}$

[illegible]([illegible])[illegible] alt text

[illegible]

- [illegible]
- [illegible]
- [illegible]w[illegible]wx+b[illegible]x[illegible]

[illegible] alt text

[illegible]**w**[illegible]**b**[illegible]**x**[illegible] alt text

[illegible] alt text

□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□
- □□□□□□□□sigmoid□□
- □□□□□□□□□□□□□□□ alt text

□□□□□

□□□□□□□□□□□□□□□□□ alt text

- □□□□□□□□□□□□□□□□□□0□□□w□x□□0□□□□□□□□wx+b□0

□□□□□ alt text

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□
- □□□□□□□$wx+b+\mu$□μ□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□$v_u^{(j)}\cdot v_m^{(j)}$,□□□□□□□□□□□□□□□
- □□□□□□□□□□ alt text

Momentum□□□□□□□□

- □□□□SGD□□

alt text

- □□□□□□□□□□□□□□□□□□β□□□□□□□0.9
- □□□
- $$v_t = \beta v_{t-1} + (1 - \beta)\frac{\partial J}{\partial w} \ w = w - \alpha v_t\(\alpha □□□ □)$$