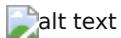


机器学习

监督学习 Supervised learning

输入向量 x 和输出值 y 的训练数据集

模型



- 线性模型
- 梯度下降法(梯度)的损失函数
- 逻辑回归模型

梯度下降

- 梯度下降法的损失函数
- $\frac{1}{2m} \sum_{i=1}^m (\hat{y}_i - y_i)^2$
- $f(x) = wx + b$

梯度

- $\nabla J(w, b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot \frac{\partial}{\partial w} (wx + b)$
- $w \leftarrow w - \alpha \nabla_w J(w, b)$



梯度

- 梯度下降法的损失函数



梯度下降法的损失函数 ##### 梯度 $\nabla J(w, b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot \frac{\partial}{\partial w} (wx + b)$

$\nabla J(w, b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot \frac{\partial}{\partial w} (wx + b)$

$\nabla J(w, b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot \frac{\partial}{\partial w} (wx + b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x$

$\nabla J(w, b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x$



梯度下降法的损失函数 $\nabla J(w, b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot \frac{\partial}{\partial w} (wx + b) = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i) \cdot x$

梯度下降

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

numpy

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

#####

1. \$x_1 = \frac{x_1 - \overline{x}}{x_{max} - x_{min}}\$
2. \$x_2 = \frac{x_2 - \overline{x}}{x_{max} - x_{min}}\$

sigmoid

#####

- sigmoid

•

- sigmoid

• sigmoid

- sigmoid

• sigmoid

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

- sigmoid

softmax

softmax

softmax

softmax

\$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)})))\$

10

sigmoid $\sigma(x) = \frac{1}{1 + e^{-x}}$

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

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

1

- ဗိသုကရာဇ်အတွက်မြန်မာစာ(၁၆၀၈)
 - ဗိသုကရာဇ်အတွက်မြန်မာစာ(၁၆၀၉-၁၆၁၀)
 - ဗိသုကရာဇ်
 1. ဗိသုကရာဇ်
 2. ဗိသုကရာဇ်
 3. ဗိသုကရာဇ်

□□□(□□□)

A horizontal row of 24 small, empty rectangular boxes arranged in a single row.

- $\boxed{\lambda}$
 - $\frac{\partial L}{\partial w_j} = \frac{1}{2m} \sum_{i=1}^m (f_{\{w,b\}}(x^{(i)}) - y^{(i)}) x_j$
 - $\boxed{\lambda} = \frac{1}{2m} \sum_{j=1}^n (w_j)^2$
 - $\boxed{\lambda} = \frac{1}{2m} \sum_{i=1}^m (f_{\{w,b\}}(x^{(i)}) - y^{(i)})^2$

□□□(□□□)

```

\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_{\vec{w}, b}(\vec{x}^{(i)}) - y^{(i)}) x_j^{(i)} + \frac{\lambda}{m} w_j
b = b - \frac{1}{m} \alpha \sum_{i=1}^m (f_{\vec{w}, b}(\vec{x}^{(i)}) - y^{(i)})

```

□□□(□□□)



1

□□□□□□□□□□

 alt text

softmax

```
## 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



softmax

10

Unsupervised Learning

□□□□□x□□□□y

4

- alt text

K-means

A horizontal row of 20 empty square boxes, intended for students to write their answers in a grid format.

4

- $\$c^{\{i\}}\$$ 000000000000
 - $\$\mu_k\$$ 00000000k000
 - $\$\mu_{c^{\{i\}}}\$$ 0000000000000000

- $m \in \{0, 1\}^n$, $x \in \{0, 1\}^m$
 - $\text{Pr}_{c \leftarrow \mathcal{C}}[c_i = 1]$
 - $\text{Pr}_{c \leftarrow \mathcal{C}}[c_i = 1] = \frac{1}{2}$
 - $\text{Pr}_{c \leftarrow \mathcal{C}}[c_i = 1] = \frac{1}{3}$

1

1

- 

1

- 二〇一〇年九月三十日
 - 二〇一〇年十月一日

1

1

- 二〇二〇年九月二十一日
 - 二〇二〇年九月二十二日

4

-  alt text

ဗိုလ်ချုပ်အတွက်

- ဗိုလ်ချုပ်အတွက်
• ဗိုလ်ချုပ်အတွက်

ဗိုလ်ချုပ်(ဗိုလ်)

- ဗိုလ်ချုပ်

ဗိုလ်

ဗိုလ်**w**ဗိုလ်(ဗိုလ်) **b**ဗိုလ်(ဗိုလ်) ဗိုလ်

ဗိုလ်ချုပ်အတွက် ဗိုလ်ချုပ်အတွက် ဗိုလ်

- ဗိုလ်ချုပ်
- ဗိုလ်ချုပ်(ဗိုလ်) ဗိုလ်
- ဗိုလ်ချုပ်
- ဗိုလ်

-
- ဗိုလ်ချုပ်
 - ဗိုလ်ချုပ်  alt text
-

ဗိုလ်ချုပ်အတွက် ဗိုလ်ချုပ်  alt text

၁၁

ဗိုလ်ချုပ်အတွက် ဗိုလ်ချုပ် ဗိုလ်

၁၂

ဗိုလ်ချုပ်အတွက် ဗိုလ်

၁၃၁၁

ဗိုလ်

tensorflow

၁၄**sequential**ဗိုလ်ချုပ်  alt text

၁၅**tensorflow**ဗိုလ်

1. ဗိုလ်
2. ဗိုလ်
3. ဗိုလ်  alt text

ဗိုလ်

ဗိုလ်**numpy**ဗိုလ်  alt text

•

- $g(z) = z$
- sigmoid(0,1),
• ReLU $g(z) = \max(0,z)$

•

Adam(α)

•

1. $\text{Adam}(\text{torch.optim})$
2. Adam
3. Adam

•

-
-

•

•(Adam),


•(Adam)


•

- (Adam)
- (Adam)

•(Adam)


•

•(Adam)


•

•(Adam)


•(Adam)


•(Adam)


•(Adam)


Algorithm

1. Initialization

- Initialize $\text{P} = \text{P}_0$
- Initialize $\text{H} = \text{H}_0$
- Initialize $\text{W} = \text{W}_0$
- Initialize $\text{F} = \text{F}_0$ (Initial F1 score)

2. Iteration

Algorithm   alt text

- $\text{P} = \text{P}_0$
- $\text{H} = \text{H}_0$
- $\text{W} = \text{W}_0$
- $\text{F} = \text{F}_0$
- $\text{P} = \text{P}_0 + \text{P}_1$
- $\text{H} = \text{H}_0 + \text{H}_1$
- $\text{W} = \text{W}_0 + \text{W}_1$
- $\text{F} = \text{F}_0 + \text{F}_1$ (Initial F1 score)
- $\text{F} = \frac{\text{P}}{\text{P} + \text{R}}$ (Final F1 score)
- $\text{P}_1 = \frac{\text{P}}{\text{P} + \text{R}}$
- $\text{H}_1 = \frac{\text{H}}{\text{P} + \text{R}}$
- $\text{W}_1 = \frac{\text{W}}{\text{P} + \text{R}}$
- $\text{F}_1 = \frac{2\text{P}\text{H}}{\text{P} + \text{R}}$

3. Termination

- $\text{P} = \text{P}_0$
- $\text{H} = \text{H}_0$

Algorithm   alt text $\text{P} = \text{P}_0$ $\text{H} = \text{H}_0$ $\text{W} = \text{W}_0$ $\text{F} = \text{F}_0$

$\text{p_0} = 1 - \text{p_1}$ $\text{H}(\text{p_1}) = -\text{p_1} \log_2(\text{p_1}) - (1 - \text{p_1}) \log_2(1 - \text{p_1})$

Algorithm   alt text $\text{P} = \text{P}_0$ $\text{H} = \text{H}_0$ $\text{W} = \text{W}_0$ $\text{F} = \text{F}_0$

$\text{p} = \text{p}^{\text{root}}$ $\text{w} = \text{w}^{\text{root}}$ $\text{H}(\text{p}) = -\text{p} \log_2(\text{p}) - (1 - \text{p}) \log_2(1 - \text{p})$

$\text{p} = \text{p}^{\text{left}}$ $\text{w} = \text{w}^{\text{left}}$ $\text{H}(\text{p}) = -\text{p} \log_2(\text{p}) - (1 - \text{p}) \log_2(1 - \text{p})$

- $\text{p} = \text{p}^{\text{right}}$ $\text{w} = \text{w}^{\text{right}}$ $\text{H}(\text{p}) = -\text{p} \log_2(\text{p}) - (1 - \text{p}) \log_2(1 - \text{p})$

4. Output

Algorithm   alt text $\text{Y} = \text{Y}_0$ $\text{P} = \text{P}_0$ $\text{H} = \text{H}_0$ $\text{W} = \text{W}_0$ $\text{F} = \text{F}_0$

Algorithm   alt text

Algorithm   alt text

ကြောင်း(ပုဂ္ဂနိုင်)

- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ

ပုဂ္ဂနိုင်

- အမြတ်ဆင့်သွေးစွဲများ(ပုဂ္ဂနိုင် 100)
- အမြတ်ဆင့်သွေးစွဲများ

XGBoost

- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ

ပုဂ္ဂနိုင်

- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ

ပုဂ္ဂနိုင်

- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ

ပုဂ္ဂနိုင်

- အမြတ်ဆင့်သွေးစွဲ(ပုဂ္ဂနိုင်)
- အမြတ်ဆင့်သွေးစွဲ

ပုဂ္ဂနိုင်

- \$r(i,j)\\$ \cdot 1_{\{x_i = x_j\}}
- \$y^{\{i,j\}}\\$
- \$w^{\{j\}}, b^{\{j\}}\\$
- \$X^{\{i\}}\\$ \cdot i_{\{x_i\}}
- \$m^{\{j\}}\\$

ပုဂ္ဂနိုင် \$w^{\{j\}} \cdot X^{\{i\}} + b^{\{j\}} \\$

ပုဂ္ဂနိုင်(ပုဂ္ဂနိုင်) ပုဂ္ဂနိုင်  alt text

ပုဂ္ဂနိုင်

- အမြတ်ဆင့်သွေးစွဲ
- အမြတ်ဆင့်သွေးစွဲများ
- အမြတ်ဆင့်သွေးစွဲများ

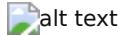
ပုဂ္ဂနိုင်  alt text

ပုဂ္ဂနိုင် \$w \cdot b \cdot x \\$  alt text

ပုဂ္ဂနိုင်  alt text

ဗိုလ်ချုပ်

- ဗိုလ်ချုပ်
- ဗိုလ်ချုပ်sigmoid
- ဗိုလ်ချုပ်



ဗိုလ်ချုပ်

ဗိုလ်ချုပ်

- $w_0x_0 + b_0 = 0$

ဗိုလ်ချုပ်

- ဗိုလ်ချုပ်
- ဗိုလ်ချုပ်
- $wx + b + \mu$



ဗိုလ်ချုပ်

ဗိုလ်ချုပ်

- ဗိုလ်ချုပ်
- $v_u^{(j)} \cdot v_m^{(j)}$
- ဗိုလ်ချုပ်



Momentum ဗိုလ်ချုပ်

- ဗိုလ်ချုပ်SGD

ဗိုလ်ချုပ်

- $\beta = 0.9$
- $v_t = \beta v_{t-1} + (1 - \beta) \nabla J(w)$
- $w_t = w - \alpha v_t$

α