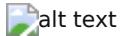


机器学习

监督学习 Supervised learning

输入向量 x 和输出值 y 的训练数据集

模型



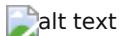
- 线性模型
- 梯度下降法(梯度)的损失函数
- 逻辑回归模型

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

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

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

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

softmax

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

1

- ဗိသုကရာဇ်အတွက် မြန်မာနိုင်ငံ(မြန်မာ)
 - ဗိသုကရာဇ်အတွက် မြန်မာနိုင်ငံ(မြန်မာနိုင်ငံ)
 - မြန်မာနှင့်
 1. မြန်မာနိုင်ငံ
 2. မြန်မာနိုင်ငံ
 3. မြန်မာ

□□□(□□□)

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

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

□□□(□□□)

```

\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

မြန်မာ

မြန်မာစာတင်ဆက်မှု မြန်မာစာတင်ဆက်မှု

မြန်မာUnsupervised Learning

မြန်မာxမြန်မာy

မြန်မာ

- မြန်မာပြန်ဆောင်ရွက်ခြင်း 

K-meansမြန်မာ

မြန်မာပြန်ဆောင်ရွက်ခြင်း

- မြန်မာပြန်ဆောင်ရွက်ခြင်း
- မြန်မာပြန်ဆောင်ရွက်ခြင်း
- မြန်မာပြန်ဆောင်ရွက်ခြင်း 

မြန်မာ

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

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

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

- မြန်မာပြန်ဆောင်ရွက်ခြင်း
- မြန်မာ
- မြန်မာပြန်ဆောင်ရွက်ခြင်း c_i
- မြန်မာပြန်ဆောင်ရွက်ခြင်း μ_i

မြန်မာ

မြန်မာပြန်ဆောင်ရွက်ခြင်း မြန်မာပြန်ဆောင်ရွက်ခြင်း

မြန်မာ

- မြန်မာပြန်ဆောင်ရွက်ခြင်း

မြန်မာ

- မြန်မာပြန်ဆောင်ရွက်ခြင်း
- မြန်မာပြန်ဆောင်ရွက်ခြင်း

မြန်မာ

မြန်မာပြန်ဆောင်ရွက်ခြင်း မြန်မာပြန်ဆောင်ရွက်ခြင်း

မြန်မာ

- မြန်မာပြန်ဆောင်ရွက်ခြင်း
- မြန်မာပြန်ဆောင်ရွက်ခြင်း

မြန်မာ

- မြန်မာပြန်ဆောင်ရွက်ခြင်း

-  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. Preprocess

- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$ (the first word is the padding symbol)

2. Tokenization

将文本转换为词向量  alt text

- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \frac{\text{vocab}}{\text{vocab} + \text{vocab}}$
- $\text{vocab} = \frac{\text{vocab}}{\text{vocab} + \text{vocab}}$ 为每个词分配一个唯一的索引
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \text{F1}(\text{vocab})(\text{vocab}) \text{vocab} \quad \text{F1_SCORE} = \frac{1}{\frac{1}{2}(\frac{1}{\text{vocab}} + \frac{1}{\text{vocab}})} = 2 \frac{\text{vocab}}{\text{vocab} + \text{vocab}}$

3. Encoding

- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$
- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$

计算信息熵  alt text 计算信息熵 $\text{H}(\text{vocab})$ $\text{H}(\text{vocab}) = - \sum p_i \log_2(p_i)$

$p_0 = 1 - p_1 \quad H(p_1) = - p_1 \log_2(p_1) - (1-p_1) \log_2(1 - p_1)$

计算信息增益  alt text 计算信息增益 $\text{G}(A) = \text{H}(\text{vocab}) - \sum p_i \text{H}(\text{vocab}_i)$

p 表示纯类别的权重 **w** 表示混合类别的权重 $\text{H}(\text{vocab}) - (w^{\text{left}} \text{H}(\text{vocab}_1) + w^{\text{right}} \text{H}(\text{vocab}_2))$

纯类别 **k** 表示纯类别的权重 **k** 表示纯类别的权重 **01** 表示纯类别的权重 **01** 表示

- $\text{vocab} = \{\text{word}_1, \text{word}_2, \dots, \text{word}_n\}$

4. Decoding

从词向量中解码出文本  alt text 从词向量中解码出文本 (词典)  alt text

Conclusion (Summary)

从文本中提取特征并将其表示为向量

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

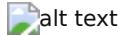
ပုဂ္ဂနိုင်  alt text

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

ပုဂ္ဂနိုင်  alt text

ဗိုလ်ချုပ်

- ဗိုလ်ချုပ်
- ဗိုလ်ချုပ်sigmoid
- ဗိုလ်ချုပ်



ဗိုလ်ချုပ်

ဗိုလ်ချုပ်

- $w_0x_0 + b_0 = 0$

ဗိုလ်ချုပ်

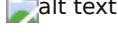
- ဗိုလ်ချုပ်
- ဗိုလ်ချုပ်
- $wx + b + \mu$



ဗိုလ်ချုပ်

ဗိုလ်ချုပ်

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



Momentum ဗိုလ်ချုပ်

- ဗိုလ်SGD

ဗိုလ်

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

α