Subject: Hw2 Deep learning
Date

97521036 سلطان حری سکر

① $\hat{y} = sigmoid(w_1x_1 + w_2x_2 + b)$, $\eta = 0.05$, N = batch size = 2

$$log\ loss = -\frac{1}{N}\sum_{i=1}^{N}(y\log(\hat{y}) + (1-y)\log(1-\hat{y}))$$ weights = weights − $\eta \nabla w L$

$\nabla w L$: $\frac{\partial\, log loss}{\partial \hat{y}} = \frac{-1}{N}\sum_{i=1}^{N}\left(\frac{\hat{y}-y}{\ln(10)(\hat{y}-1)\hat{y}}\right)$, $\left\{\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{w}\right.$, $\left\{\frac{\partial \hat{y}}{\partial w_1} = x_1,\ \frac{\partial \hat{y}}{\partial w_2} = x_2\right.$

$\ln(10) \to 2.3$

Calculation from Internet

$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial b}$ $\left\{\frac{\partial \hat{y}}{\partial b} = 1\right.$

Epoch-1: دارد = 1 , ندارد = −1

- Batch 1: $\hat{y}_1 = (1)(.22) + (1)(1) + 1 = 2.22$ Sig = 0.902
  $\hat{y}_2 = (1)(.25) + (1)(-1) + 1 = 0.25$ Sig = 0.562

→ نکته: برای انجام محاسبات منطقی تر feature سن را تقسیم بر 100 کردم Normalize

$w_1 = 1 - (0.05)(-\frac{1}{2}(-2.747)) = 0.932$

$w_2 = 1 - (0.05)(-\frac{1}{2}(-6.596)) = 0.836$, $b = 1 - 0.05(1)(6.057) = 0.698$

- Batch 2: $\hat{y}_1 = (0.932)(.47) + (0.836)(1) + 0.698 = 1.972$ Sig = 0.877
  $\hat{y}_2 = (0.932)(.52) + (0.836)(-1) + 0.698 = 0.346$ Sig = 0.585

$w_1 = 0.932 - 0.05(0.621) = 0.901$, $w_2 = 0.836 - 0.05(-1.667) = 0.919$

$b = 0.698 - 0.05(1.131) = 0.642$

- Batch 3: $\hat{y}_1 = (0.901)(.46) + 0.919(1) + 0.642 = 1.975 \to 0.878$
  $\hat{y}_2 = (0.901)(.56) + 0.919(1) + 0.642 = 2.065 \to 0.887$

$w_1 = 0.901 - 0.05(-0.251) = 0.913$, $w_2 = 0.919 - 0.05(-0.492) = 0.943$

Subject:
Date

$b = 0.642 - 0.05(-0.492) = 0.664$

- Batch 4:
$\hat{y}_1 = (0.913)(0.55) + (0.943)(-1) + .664 = 0.223 \rightarrow 0.555$
$\hat{y}_2 = (0.913)(0.6) + (0.943)(-1) + 0.664 = 0.268 \rightarrow 0.566$

$w_1 = 0.913 - 0.05(+0.522) = 0.887$

$w_2 = 0.943 - 0.05(-0.984) = 0.992$, $b = 0.664 - 0.05(0.984) = 0.615$ ✓

Epoch 2:

- Batch 1:
$\hat{y}_1 = 0.887(.22) + 0.992(1) + 0.615 = 1.802 \rightarrow 0.858$
$\hat{y}_2 = 0.887(.25) + 0.992(-1) + 0.615 = -0.155 \rightarrow 0.461$

$w_1 = 0.887 - 0.05(-1.048) = 0.835$

$w_2 = 0.992 - 0.05(2.037) = 0.891$, $b = 0.615 - 0.05(4.593) = 0.386$

- Batch 2:
$\hat{y}_1 = 0.835(0.47) + 0.891(1) + 0.386 = 1.669 \rightarrow 0.841$
$\hat{y}_2 = 0.835(0.52) + 0.891(-1) + 0.386 = -0.07 \rightarrow 0.482$

$w_1 = 0.835 - 0.05(+0.549) = 0.808$, $w_2 = 0.891 - 0.05(-1.548) = 0.968$

$b = 0.386 - 0.05(-1.031) = 0.437$

- Batch 3:
$\hat{y}_1 = 0.808(0.46) + 0.968(1) + 0.437 = 1.776 \rightarrow 0.855$
$\hat{y}_2 = 0.808(0.56) + 0.968(1) + 0.437 = 1.857 \rightarrow 0.864$

$w_1 = 0.808 - 0.05(-0.257) = 0.82$, $w_2 = 0.968 - 0.05(-0.505) = 0.993$

$b = 0.437 - 0.05(-0.501) = 0.462$

- Batch 4:
$\hat{y}_1 = 0.82(0.55) + 0.993(-1) + 0.462 = -0.08 \rightarrow 0.48$
$\hat{y}_2 = 0.82(0.55) + 0.993(-1) + 0.462 = -0.039 \rightarrow 0.49$

$w_1 = 0.82 - 0.05(0.442) = 0.798$, $w_2 = 0.993 - 0.05(-0.845) = 1.035$

$b = 0.462 - 0.05(0.845) = 0.42$

ادامه صفحه بعدی ←

PqPCO

Subject:
Date

← حال loss مرحله قبل آموزش (قدم اول) و مرحله آخر را (بعد از آموزش) مقایسه می‌کنیم.

1) $-\frac{1}{2}\left((-1)\log(0.902) + \frac{(1-(-1))}{2}\log(0.098) + (-1)\log(0.562) + \frac{(1-(-1))}{2}\log(0.438)\right)$

$= 2.80$

2) $\begin{cases} 0.798(0.22) + 1.035(1) + 0.42 = 1.63 \rightarrow 0.836 \\ 0.798(0.25) + 1.035(-1) + 0.42 = -0.415 \rightarrow 0.397 \end{cases}$

$\rightarrow$ loss $\Rightarrow$ 1.762

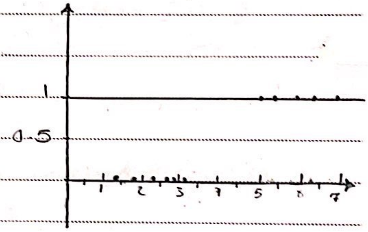
← همان طور مشاهده می‌کنیم loss کاهش یافته است. مراحل به طور کامل توضیح داده شده است ولی محاسبه دستی به صورت آنلاین انجام شد (بود) و شرح حاصل در کاغذ نیامده به علت طولانی بودن. (توضیحات بیشتر در منبع داده)

② linear Regression: $\hat{y} = w_1 x_1 + b$

$\hat{y} = 0.5 \rightarrow$ decision boundary

$w_1 x_1 + b = 0.5$

for $x < 3.1$ $y < 0.5$, $x \geq 4.9$ $y \geq 0.5$



logistic Regression: $\hat{y} = \text{Sigmoid}(w_1 x_1 + b) = 0.5$, $\frac{1}{1+e^{-z}} = 0.5$

$e^{-z} = 1$, $-z = 0 \rightarrow z = wx + b = 0$, $x = -b/w$

← حل عملی انجام شده و توضیح تصمیم انجام شده در آنجا، اینجا آورده شده، محاسبات بالا. مقداری از توضیحات هم در داک است.