Logistic-Regression-2

Date: 2020-07-07 Tuesday

Agenda :

By Picture.

① Linear fn $\rightarrow$ $y^{(i)} = w^T x^{(i)} + b$

② $z^{(i)} \in (-\infty, \infty)$ Squashing $\hat{y}^{(i)} = \sigma(z^{(i)})$ $(0,1)$: Probabilities $P(y^{(i)}=1|x^{(i)})$

③ Loss fnc $y^{(i)}$ vs $\hat{y}^{(i)}$ log-loss

④ Cost / Evaluat $\frac{1}{m}\sum_{i=1}^{m} \underbrace{\text{Loss}(y^{(i)}, \hat{y}^{(i)})}_{\text{Loss}} = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(y^{(i)}, \hat{y}^{(i)})$

$\mathcal{L}(y^{(i)}, \hat{y}^{(i)})$ $\leftarrow$ Reason.

⑤ Algo - parameter Minimize : Gradient or Derivative

"σ"

* Definitions

Accuracy

Act Class / Pred Class

	$y^{(i)}$ = 1	$y^{(i)}$ = 0
$\hat{y}^{(i)}$ = 1	0	0
$\hat{y}^{(i)}$ = 0	1	0
$\hat{y}^{(i)}$ = 1	1	0
$\hat{y}^{(i)}$ = 0	0	0
$\vdots$	$\vdots$	$\vdots$

Accuracy

$y^{(i)}$ | $\hat{y}^{(i)}$ | $|y^{(i)} - \hat{y}^{(i)}|$: if ≥ 0.5 else 0

0 | 0.99 | 1 ✓

0 | 0.97 | 0 ✓

1 | 0.8 | 0 ✗

1 | 0.9 | 1 ✓

$\vdots$ | $\vdots$ | $\vdots$

Accuracy

let's say $\frac{95 \text{ numbers}}{100}$ $\Rightarrow$ 95% $\Rightarrow$ 0.95

* Log-odds/Interpretation of log-reg

$\frac{\text{success}}{1-\text{success}}$ $\frac{1-p}{p}$

p: $P(\text{success}) = \frac{1}{2}$

$\text{odds} = \frac{1-p}{p} = \frac{1}{2}$

$1 - p = P(\text{fail}) = \frac{1}{2}$

$\frac{3}{2}$

$\frac{1}{2}$

$\frac{3}{2}$

$\frac{1}{2}$

$p \Rightarrow P(y=1)$

$1-p \Rightarrow P(y=0)$

$\text{odds} = \frac{p}{1-p} \Rightarrow$ odds of class label 1

$p \Rightarrow \sigma(w^T x + b)$ $\Rightarrow$ $\frac{1}{1+e^{-(w^T x+b)}}$

$\Rightarrow \frac{1}{1+\frac{1}{e^{w^T x+b}}}$

$\Rightarrow \frac{e^{w^T x+b}}{e^{w^T x+b}+1}$

$1-p \Rightarrow 1 - \frac{e^{w^T x+b}}{e^{w^T x+b}+1}$

$\Rightarrow \frac{e^{w^T x+b}+1 - e^{w^T x+b}}{e^{w^T x+b}+1}$

$1-p \Rightarrow \frac{1}{e^{w^T x+b}+1}$

$\text{odds} \Rightarrow \frac{p}{1-p} \Rightarrow \frac{\frac{e^{w^T x+b}}{e^{w^T x+b}+1}}{\frac{1}{e^{w^T x+b}+1}} \Rightarrow e^{w^T x+b}$

$\text{odds} = e^{w^T x+b}$

$\log\left(\frac{p}{1-p}\right) = w^T x + b$

Linear fn. in x

Actual meth $\rightarrow$ Logit

$z^{(i)} \Rightarrow \log\left(\frac{p}{1-p}\right)$

$\log(x)$

$z^{(i)} \Rightarrow \log\left(\frac{p}{1-p}\right)$

$\log(x)$

$(0,1)$

$P(y=1) \in (0,1)$

$\rightarrow$

$\frac{p}{1-p}$

$\log\left(\frac{p}{1-p}\right) = e^{-z}$

$\frac{1-p}{p} = \frac{e^{-z}}{1}$

$p = \frac{1}{1+e^{-z}}$

$w^T x + b = 0$

$\frac{p}{1-p} = 1 \Rightarrow \log\left(\frac{p}{1-p}\right) = 0$

$P(y=1) = 0.5$

$\frac{p}{1-p} = \frac{1}{1-p} = 1$

$p = 1-p$

$2p = 1$

$p = \frac{1}{2}$

$P(y=1) = \frac{1}{2}$

$\frac{p}{1-p} = \frac{1}{1-p} = 1$

$\frac{p}{1-p} = \frac{0.9}{0.1} = 9$

$\frac{p}{1-p} = \frac{0.1}{0.9} = \frac{1}{9}$

$\frac{p}{1-p} = \frac{0.5}{0.5} = 1$

$\frac{p}{1-p} = \frac{0.99}{0.01} = 99$

$\frac{p}{1-p} = \frac{0.01}{0.99} = \frac{1}{99}$

$\frac{p}{1-p} = \frac{0.9999}{0.0001} = 9999$

$\frac{p}{1-p} = \frac{0.0001}{0.9999} = \frac{1}{9999}$

* Multi class classifications. (One vs Rest / One vs All) OVR

x_1	x_2	x_3	y
			Shark
			Tuna
			whale

class 1 : Shark

class 2 : Tuna

class 3 : whale

Create 3 log-reg

1st Model : Shark or Not

2nd Model : Tuna or Not

3rd Model : whale or Not

Step 1

x	y
	1
	0
	0
	1
	0
	0
	0

if y = Shark $\Rightarrow$ 1

else $\Rightarrow$ 0

M1 $\rightarrow$ (Shark, Others)

Step 2

x	y
	0
	1
	0
	1
	0
	0
	0

if y = Tuna $\Rightarrow$ 1

else $\Rightarrow$ 0

M2 $\rightarrow$ (Tuna vs Others)

Step 3

x	y
	0
	0
	1
	0
	0
	1
	0

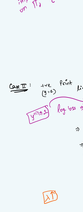
if y = whale $\Rightarrow$ 1

else $\Rightarrow$ 0

M3 $\rightarrow$ (Tuna vs Others)

Multi class classification $\Rightarrow$ log-reg + "OVR"

* Impact of Outliers



$y^{(i)} = 0$: $\hat{y}^{(i)} \approx 0$

$y^{(i)} = 1$: $\hat{y}^{(i)} \approx 1$

$\text{Loss}(y^{(i)}, \hat{y}^{(i)}) \approx 0 + 0$

$y = 0 \Rightarrow -\log(1-\hat{y}^{(i)})$

$y = 1 \Rightarrow -\log(\hat{y}^{(i)})$

Actual $\frac{y=0}{y=1}$

Pred $\frac{y=0}{y=1}$

Log-loss $\begin{cases} -\log(1-\hat{y}^{(i)}) \\ -\log(\hat{y}^{(i)}) \end{cases}$

$\Rightarrow$ convex

$\text{loss} = -\log(1-\hat{y}^{(i)})$

$\hat{y}^{(i)} \approx 0$

$\text{loss} = -\log(\hat{y}^{(i)})$

$\hat{y}^{(i)} \approx 1$

$\text{loss}(\hat{y}^{(i)}) \approx 0$

$\text{loss}(\hat{y}^{(i)}) \approx \infty$

Case 1 : yes, first plot on Linear scale

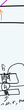
$y^{(i)} = 0$ $\quad -\log(1-\hat{y}^{(i)})$

$= -\log(-\infty)$

$\hat{y}^{(i)} = -\infty$

$\Rightarrow \infty$

(v. large value)

if $y^{(i)} = 0$

$\hat{y}^{(i)} \approx -\infty$