ML WEEK5

phuong nguyen

October 2021

1 BIEN DOI HAM LOGISTIC: 2-CLASSES PROBLEM

Let C1 be event that new data belongs to class1
And let C2 be event that new data belongs to class2

$P(C1/x) = \frac{P(x/C1).P(C1)}{P(x/C1)P(C1)+P(x/C2)P(C2)}$
Divided up and down side for P(x/C1)P(C1), we get:

$P(C1/x) = \frac{1}{1+\frac{P(x/C2)P(C2)}{P(x/C1)P(C1)}}$

P(C1/x)=$\frac{1}{1+e^{ln}(\frac{P(x/C2)P(C2)}{P(x/C1)P(C1)})}$

Let $a = ln(\frac{P(x/C2)P(C2)}{P(x/C1)P(C1)}) => P(C1/x) = \sigma(a) = \frac{1}{1+e^{-a}}$

From this, we get the logistic model:
$P(C1/\Theta) = y(\Theta) = \sigma(w^T\Theta)$
$P(C2/\Theta) = 1 - P(C1/\Theta)$

For data set=Θ_n, target t_n with $t_n \in$ (0,1), Θ_n=$\Theta(x_n)$ with n=1,2,...N, and w are cofficients the likelihood function is:

$P(t/w) = \prod y^t_n n(1-y_n)^1 - t_n$
Taking the negative log of this function, we get

L=$-ln(P(t/w)) = -\sum[t_n log(y_n) + (1-t_n)log(a-y_n)]$
Where y_n=$\sigma(a_n) = u$ and $a_n = w^T\Theta_n = z$, using chain rule to take derivative of L, we get

$S = \frac{d(L)}{d(u)}.\frac{d(u)}{d(z)}.\frac{d(z)}{d(w_i)}$ with(i=0,1,...N)
$\frac{d(L)}{d(u)} = -[\frac{t_n}{u} - \frac{1-t_n}{1-u}]$
$\frac{d(u)}{d(z)} = u(z).(1-u(z))$

$=> S = \sum (y_n - t_n).\Theta_n$

Using gradient descent method to find w:
b1: random an initial w
b2: update w=w-learning rate.$\frac{d(L)}{d(w_i)}$ with i=0,1,..N, learning rate>0
b3: if Loss at new w is small enough then stop or else continue step2

2 FIND F(x) given F'(x)=F(x).(1-F(x))

f'(x)=f(x).(1-f(x))

<=> f(x)-f'(x)=$f^2(x)$

<=> $\frac{f(x)-f'(x)}{f^2(x)} = 1$ suppose(f(x)¿0)

<=> $\frac{e^x.f(x)-e^x.f'(x)}{f^2(x)} = e^x$ $(e^x > 0)$
<=> $[\frac{e^x}{f(x)}]' = e^x$

<=> $\frac{e^x}{f(x)} = e^x + C$ with C= constant

<=> $f(x) = \frac{e^x}{e^x+C}$

<=> $f(x) = \frac{1}{1+C.e^-x}$

If f(0)=1/2 then C=1 => f(x)=$\frac{1}{1+e^-x}$