تمرین ۶۰۱ هوش مصنوعی

امیرحسین اثری 9810 ~~5~~621

سوال اول

الف

$$P(D|\theta=\hat{\theta}) = \prod_{i=1}^{N} P^{y_i}(1-P)^{1-y_i} = P^{\sum y_i}(1-P)^{N-\sum y_i}$$

$y \in \{0,1\}$

$$\Rightarrow \boxed{\log P(D|\theta=\hat{\theta}) = \sum y_i \log P + (N-\sum y_i)\log(1-P)}$$

ب

$$\log \frac{P}{1-P} = ax+b \Rightarrow \frac{P}{1-P} = e^{ax+b} \Rightarrow \frac{P+1-1}{1-P} = e^{ax+b} \Rightarrow -1+\frac{1}{1-P} = e^{ax+b}$$

$$\Rightarrow \frac{1}{1-P} = 1+e^{ax+b} \Rightarrow \boxed{P = \frac{e^{ax+b}}{1+e^{ax+b}}}$$

$$\log P(D|\theta=\hat{\theta}) = \sum y_i \underbrace{[\log P - \log 1-P]}_{\log \frac{P}{1-P}} = \sum y_i (ax+b)$$

$$N\log(1-P) = N\log(1+e^{ax+b})^{-1}$$

$$P(D|\theta=\hat{\theta}) = \sum y_i (ax+b) - N\log(1+e^{ax+b})$$

$$\Rightarrow \boxed{-P(D|\theta=\hat{\theta}) = N\log(1+e^{ax+b}) - \sum y_i (ax+b)} \Rightarrow$$

P

$$\frac{\partial P}{\partial a} = \frac{N \times a}{1+e^{ax+b}} e^{ax+b} - \sum y_i \times a = 0 \Rightarrow a(P - \sum y_i) = 0$$

$$\Rightarrow P = \sum y_i \to \max$$

$$\frac{\partial P}{\partial b} = \frac{N}{1+e^{ax+b}} e^{ax+b} - \sum y_i = 0 \Rightarrow P - \sum y_i = 0 \Rightarrow$$

$$P = \sum y_i \to \max$$

Logistic regression است.

حساب

$\hat{P} = \Sigma y_i$ و با بلاک آن logistic regression به رابطه حسشهور بدست می آید :

سوال ۲

آ) به جای داشتن یک set از Parameter a b، a و b کلاسی پارامتر a خود و b خود را دارد : و بنابراین:

$$\log \frac{P_i(x)}{1-P_i(x)} = \beta_{0i} + \beta_i X$$

$$\Rightarrow \Pr(Y=c \mid \vec{X}=x) = \frac{e^{\beta_0^{(c)} + X\beta(c)}}{\sum_c e^{\beta_0^{(c)} + X\beta(c)}}$$

+ بنابراین به ازای هر کلاس بردار w_i را در نظر بگیرید

$$P(y_{ij}) = \begin{bmatrix} X_i^T w_1 \\ \vdots \\ X_i^T w_K \end{bmatrix}_j = \frac{e^{X_i^T w_j}}{\sum_{j=1}^{n} e^{X_i^T w_j}}$$

$$Loss(\beta) = \prod_{1}^{n} P(X_i = y_j) = \prod_{i=1}^{N} \frac{e^{X_i^T w_j}}{\sum_{j=1}^{K} e^{X_i^T w_j}} \Rightarrow \log L \Rightarrow$$

$$l(\beta) = \sum_{i=1}^{N} \frac{e^{X_i^T w_i}}{\sum_{j=1}^{K} e^{X_i^T w_j}} \Rightarrow \text{Gradian}(L) \ \text{min}\ 0 \Rightarrow \nabla L = 0$$

while loss > threshold

for j from 1 to k

$$w_j = w_j - \eta \nabla l_j(w)$$

η → Learning rate

or → while loss > threshold

~~for j from 1 to k~~

~~w~~

$$w = w - \eta \nabla l$$