閭昏總鍥炻綀鎬荤粙

2017-9-28

漢逛篛绾拧 y 渓褰掑ā 鍥y$=\omega^T+b$权撖嚚鑽勮 娴嬪yr$-*x^{\vee}$#w\$$n^{\vee}y_i \in$ $\{0,1\}$鉐備负浜唕曉瀹炲y4¨$Jigmoid˜-Gy = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-(\omega^T+b)}}$,漢$\omega$鍜bpιB,$^{\vee}$$\beta = (\omega;b)$,$\hat{x} = (x;1)$,鍒y$=\omega^T+b = \beta^T\hat{x}$,甯〜疁 sigmoid 鍑芥暟鉐們傸偢偁備傳偘傸偑傔侾偀偍$\frac{}{1+e^{-\beta^T\hat{x}}(0)}$

漢筞igmoid 鍑芥暟锛z$=\ln\frac{y}{1-y}$,鍒欐纺(**??**)鍙　互鍙樹负

$$\ln\frac{y}{1-y} = \beta^T\hat{x} \tag{1}$$

鑻ュ晓y瑙嗕负绫诲悗楠屾　鏊困及璁P(y=1—x)ØZUu)Øȷ\$[db(**??**)Γ$probability p''sK$ØΓ$equation*\ln\frac{1}{1-...}$逛篛鐭锋溄閫$\{x_i,y_i\}_{i=1}^{m}$锛屾柯鏈　殙姒傊昆鍙　互鍐欍负锛毘洗沛沫杅汓沒沪汢気汣氊泆$_i|\beta;\hat{x}_i) = P(y_i = 1|\beta;\hat{x}_i)^{y_i}(1 - P(y_i = 1|\beta;\hat{x}_i))^{1-y_i}$鍏朵腑锛毘杺氝泆汣氫湽沴汓乒氿沋$^{\beta^Tx}\frac{}{1+e^{\beta^Tx}P(y=0|x)=\frac{1}{1+e^{\beta^Tx}}}$

n涓　�села嬬殑鏍峰溄鍑虹幇鐨勁技鍒跺噣鏌颁负(鍥狆负姣忎釜鏍峰溄閮芥槸鐙珛鐨勫彇渚侾涓　柯鏈　嚣鐅扮殑姒傜瑡灏辨槸浠州滑鍚勫劘嚁鍑虹幇鐨勩　鏊固浉涔:

$$L(\theta) = \prod P(y_i|\beta;\hat{x}_i) \tag{2}$$

漢逛筼寮忔彌漢规暟锛嬪褰寬寤寃寢寪寊

涓轰簡涗樺寲姹傝В B 锛嵓　鐖　熛鍑芥暟鍙槶弽锛呬互枃浼　鏈濈忓间綊　鍖噱

寮(**??**)儴siran 鋈栻渫錫庯紅漢β妵傚 鑅帮細

$$dL(\theta) = \sum \left(\frac{e^{\beta^T \hat{x_i}} d\beta^T \hat{x_i}}{1 + e^{\beta^T \hat{x_i}}} - y_i d\beta^T \hat{x_i} \right)$$

$$= \text{tr} \sum \left(\frac{\hat{x_i} e^{\beta^T \hat{x_i}} d\beta^T}{1 + e^{\beta^T \hat{x_i}}} - \hat{x_i} y_i d\beta^T \right) \qquad (3)$$

$$= \sum \hat{x_i}(P(y_i = 1|\beta; \hat{x_i}) - y_i) d\beta^T$$

$$\frac{\partial dL(\theta)}{\partial \beta^T} = \sum \hat{x_i}^T (P(y_i = 1|\beta; \hat{x_i}) - y_i)$$

$$\frac{\partial dL(\theta)}{\partial \beta} = \sum \hat{x_i}(P(y_i = 1|\beta; \hat{x_i}) - y_i) \qquad (4)$$

杩慅鞍姊 害涓壓欄娉疊繘琛屼紭鍖枕睉瑙 o 細

$$\beta_{n+1} = \beta_n - \alpha \frac{\partial dL(\theta)}{\partial \beta}$$