1. Naïve Bayes

1.1 Duplicate Feature and Decision Rules

1.

$$P(Y=T|X_1=T,X_2=F,X_3=F)=\frac{P(X_1=T,X_2=F,X_3=F|Y=T)P(Y=T)}{P(X_1=T,X_2=F,X_3=F)}$$
$$=\frac{P(X_1=T|Y)P(X_2=F,X_3=F|Y=T)P(Y=T)}{P(X_1=T,X_2=F,X_3=F)}$$
$$=\frac{P(X_1=T|Y=T)P(X_2=F|Y=T)P(X_3=F|Y=T)P(Y=T)}{P(X_1=T,X_2=F,X_3=F|Y=T)P(Y=T)+P(X_1=T,X_2=F,X_3=F|Y=F)`P(Y=F)}$$
$$=\frac{p\cdot q\cdot q\cdot 0.5}{p\cdot q\cdot q\cdot 0.5+(1-p)\cdot(1-q)\cdot(1-q)\cdot 0.5}=\frac{0.5\cdot p\cdot q^2}{0.5\cdot p\cdot q^2+0.5\cdot(1-p)(1-q)^2}$$

For a positive classification the above probability has to be greater than 0.5

$$\frac{0.5\cdot p\cdot q^2}{0.5\cdot p\cdot q^2+0.5\cdot(1-p)(1-q)^2}\geq 0.5$$
$$\Leftrightarrow \frac{p\cdot q^2}{p\cdot q^2+(1-p)(1-q)^2}\geq 0.5$$
$$\Leftrightarrow pq^2\geq 0.5\cdot(pq^2+(1-p)(1-q)^2)$$
$$\Leftrightarrow\ 0.5pq^2\geq 0.5(1-p)(1-q)^2$$
$$\Leftrightarrow pq^2\geq 1-2q+q^2-p+2pq-pq^2$$
$$\Leftrightarrow 2pq^2+p-2pq\geq 1-2q+q^2$$
$$\Leftrightarrow p(2q^2+1-2q)\geq(1-q)^2$$
$$\Leftrightarrow p\geq\frac{(1-q)^2}{q^2+q^2+1-2q}$$
$$\Leftrightarrow p\geq\frac{(1-q)^2}{q^2+(1-q)^2}$$

2.

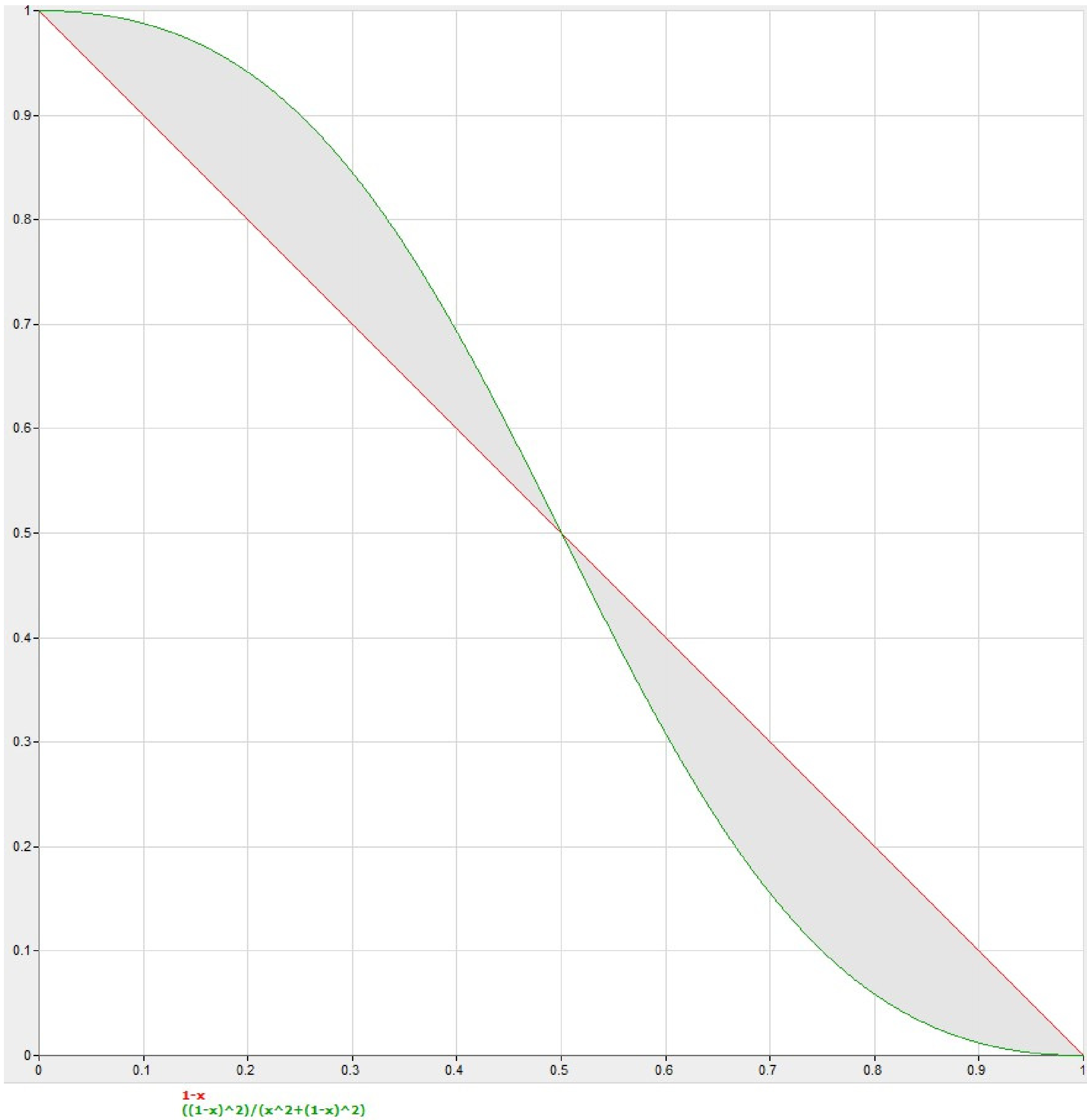
$$P(Y=T|X_1=T,X_2=F,X_3=F)=\frac{P(X_1=T,X_2=F,X_3=F|Y=T)P(Y=T)}{P(X_1=T,X_2=F,X_3=F)}$$
$$=\frac{P(X_1=T|Y)P(X_2=F,X_3=F|Y=T)P(Y=T)}{P(X_1=T,X_2=F,X_3=F)}$$
$$=\frac{P(X_1=T|Y=T)P(X_2=F|Y=T)P(Y=T)}{P(X_1=T,X_2=F,X_3=F|Y=T)P(Y=T)+P(X_1=T,X_2=F,X_3=F|Y=F)`P(Y=F)}$$
$$=\frac{p\cdot q\cdot 0.5}{p\cdot q\cdot 0.5+(1-p)\cdot(1-q)\cdot 0.5}=\frac{0.5\cdot p\cdot q}{0.5\cdot p\cdot q+0.5\cdot(1-p)(1-q)}$$

For a positive classification, again, the above probability has to be greater than 0.5

$$\frac{0.5\cdot p\cdot q}{0.5\cdot p\cdot q+0.5\cdot(1-p)(1-q)}\geq 0.5$$
$$\frac{p\cdot q}{p\cdot q+(1-p)(1-q)}\geq 0.5$$

$pq \geq 0.5pq + 0.5(1-p)(1-q)$
$0.5pq \geq 0.5(1-p)(1-q)$
$pq \geq (1-p)(1-q)$
$pq \geq 1 - p - q + pq$
$pq - pq + p \geq 1 - q$
$p \geq 1 - q$

3.



The first rule makes mistakes in the gray areas.

1.2 Logistic Regression and Naïve Bayes – Boolean Case

$$P(Y=1|X)=\frac{P(X|Y=1)P(Y=1)}{P(X|Y=1)P(Y=1)+P(X|Y=0)P(Y=0)}=\frac{1}{1+\frac{P(X|Y=0)P(Y=0)}{P(X|Y=1)P(Y=1)}}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{P(X|Y=0)P(Y=0)}{P(X|Y=1)P(Y=1)}\right)\right]}=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i\ln\left(\frac{P(X_i|Y=0)}{P(X_i|Y=1)}\right)\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i\ln\left(\frac{(\theta_{i0})^{X_i}(1-\theta_{i0})^{(1-X_i)}}{(\theta_{i1})^{X_i}(1-\theta_{i1})^{(1-X_i)}}\right)\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i\ln((\theta_{i0})^{X_i}(1-\theta_{i0})^{(1-X_i)})-\ln((\theta_{i1})^{X_i}(1-\theta_{i1})^{(1-X_i)})\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i\ln((\theta_{i0})^{X_i})+\ln((1-\theta_{i0})^{(1-X_i)})-\ln((\theta_{i1})^{X_i})-\ln((1-\theta_{i1})^{(1-X_i)})\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i X_i\ln(\theta_{i0})+(1-X_i)\ln(1-\theta_{i0})-X_i\ln(\theta_{i1})-(1-X_i)\ln(1-\theta_{i1})\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i X_i\ln(\theta_{i0})+\ln(1-\theta_{i0})-X_i\ln(1-\theta_{i0})-X_i\ln(\theta_{i1})-\ln(1-\theta_{i1})+X_i\ln(1-\theta_{i1})\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i X_i(\ln(\theta_{i0})-\ln(1-\theta_{i0})-\ln(\theta_{i1})+\ln(1-\theta_{i1}))+\ln(1-\theta_{i0})-\ln(1-\theta_{i1})\right]}$$

$$=\frac{1}{1+\exp\left[w_0+\sum_i w_iX_i\right]}$$

where $w_0=\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i\ln\left(\frac{1-\theta_{i0}}{1-\theta_{i1}}\right)$

and $w_i=\ln(\theta_{i0})-\ln(1-\theta_{i0})-\ln(\theta_{i1})+\ln(1-\theta_{i1})=\ln\left(\frac{\theta_{i0}}{\theta_{i1}}\right)+\ln\left(\frac{1-\theta_{i1}}{1-\theta_{i0}}\right)$

1.3 Relaxing the Conditional Independence Assumption

$$P(Y=1|X)=\frac{\prod_{i=3}^{n}P(X_i|Y=1)\,P(X_1,X_2|Y=1)P(Y=1)}{\prod_{i=3}^{n}P(X_i|Y=1)\,P(X_1,X_2|Y=1)P(Y=1)+\prod_{i=3}^{n}P(X_i|Y=0)\,P(X_1,X_2|Y=0)P(Y=0)}$$

$$=\frac{1}{1+\frac{\prod_{i=3}^{n}P(X_i|Y=0)\,P(X_1,X_2|Y=0)P(Y=0)}{\prod_{i=3}^{n}P(X_i|Y=1)\,P(X_1,X_2|Y=1)P(Y=1)}}=\frac{1}{1+\exp\left[\ln\left(\frac{\prod_{i=3}^{n}P(X_i|Y=0)\,P(X_1,X_2|Y=0)P(Y=0)}{\prod_{i=3}^{n}P(X_i|Y=1)\,P(X_1,X_2|Y=1)P(Y=1)}\right)\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_{i=3}^{n}\left[\ln\left(\frac{P(X_i|Y=0)}{P(X_i|Y=1)}\right)\right]+\ln\left(\frac{P(X_1,X_2|Y=0)}{P(X_1,X_2|Y=1)}\right)\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_{i=3}^{n}\left[\ln\left(\frac{(\theta_{i0})^{X_i}(1-\theta_{i0})^{(1-X_i)}}{(\theta_{i1})^{X_i}(1-\theta_{i1})^{(1-X_i)}}\right)\right]+\ln\left(\frac{(\beta_{110})^{X_1X_2}(\beta_{100})^{X_1(1-X_2)}(\beta_{010})^{(1-X_1)X_2}(\beta_{000})^{(1-X_1)(1-X_2)}}{(\beta_{111})^{X_1X_2}(\beta_{101})^{X_1(1-X_2)}(\beta_{011})^{(1-X_1)X_2}(\beta_{001})^{(1-X_1)(1-X_2)}}\right)\right]}$$

$$=\frac{1}{1+\exp\left[\ln\left(\frac{1-\pi}{\pi}\right)+\sum_{i=3}^{n}\left[X_i\left(\ln\left(\frac{\theta_{i0}}{\theta_{i1}}\right)+\ln\left(\frac{1-\theta_{i1}}{1-\theta_{i0}}\right)\right)+\ln\left(\frac{1-\theta_{i0}}{1-\theta_{i1}}\right)\right]+X_1X_2B_1+X_1B_2+X_2B_3+B_4\right]}$$

$$=\frac{1}{1+\exp\left[w_0+w_{1,2}X_1X_2+\sum_{i=1}^{n}w_iX_i\right]}$$

where $w_0=\ln\left(\frac{1-\pi}{\pi}\right)+\sum_i\ln\left(\frac{1-\theta_{i0}}{1-\theta_{i1}}\right)+\ln\left(\frac{\beta_{000}}{\beta_{001}}\right)$

and $w_1=\ln\left(\frac{\beta_{100}}{\beta_{101}}\right)+\ln\left(\frac{\beta_{001}}{\beta_{000}}\right)$

and $w_2=\ln\left(\frac{\beta_{010}}{\beta_{011}}\right)+\ln\left(\frac{\beta_{001}}{\beta_{000}}\right)$

and $w_i=\ln\left(\frac{\theta_{i0}}{\theta_{i1}}\right)+\ln\left(\frac{1-\theta_{i1}}{1-\theta_{i0}}\right)$ for $i\in\{3,\dots,n\}$

and $w_{1,2}=\ln\left(\frac{\beta_{110}}{\beta_{111}}\right)+\ln\left(\frac{\beta_{000}}{\beta_{001}}\right)+\ln\left(\frac{\beta_{101}}{\beta_{100}}\right)+\ln\left(\frac{\beta_{011}}{\beta_{010}}\right)$

$$B_1=\ln\left(\frac{\beta_{110}}{\beta_{111}}\right)+\ln\left(\frac{\beta_{000}}{\beta_{001}}\right)+\ln\left(\frac{\beta_{101}}{\beta_{100}}\right)+\ln\left(\frac{\beta_{011}}{\beta_{010}}\right)$$

$$B_2=\ln\left(\frac{\beta_{100}}{\beta_{101}}\right)+\ln\left(\frac{\beta_{001}}{\beta_{000}}\right)$$

$$B_3=\ln\left(\frac{\beta_{010}}{\beta_{011}}\right)+\ln\left(\frac{\beta_{001}}{\beta_{000}}\right)$$

$$B_4=\ln\left(\frac{\beta_{000}}{\beta_{001}}\right)$$