

GDA

On Dataset 1 GDA perform worse than logistic regression.

Because $p(x|y)$ may be not Gaussian distribution.

(h)

Box-Cox transformation.

2.

(a)

$$P(y = 1|t = 1, x)P(t = 1|x)P(x) = P(y = 1, t = 1, x) = P(t = 1|y = 1, x)P(y = 1|x)P(x)$$

$$P(t = 1|x) = P(y = 1|x) \frac{P(t = 1|y = 1, x)}{P(y = 1|t = 1, x)}$$

$$P(t = 1|y = 1, x) = 1, P(y = 1|t = 1, x) = P(y = 1|t = 1)$$

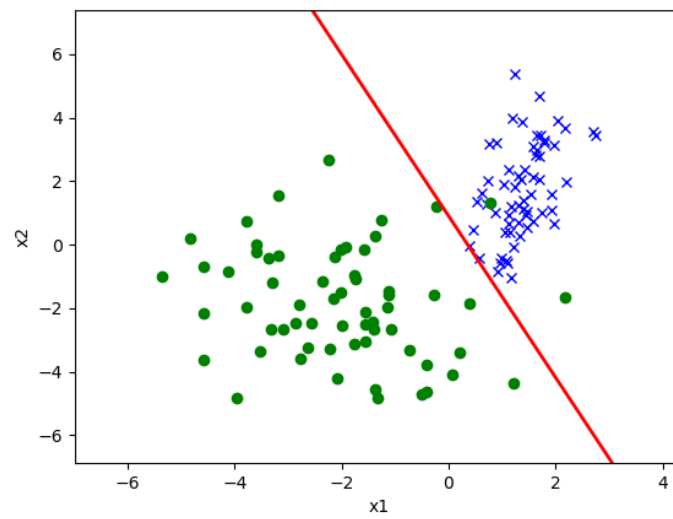
$$P(t = 1|x) = \frac{P(y = 1|x)}{P(y = 1|t = 1)}$$

$$P(y = 1|t = 1) = \alpha$$

(b)

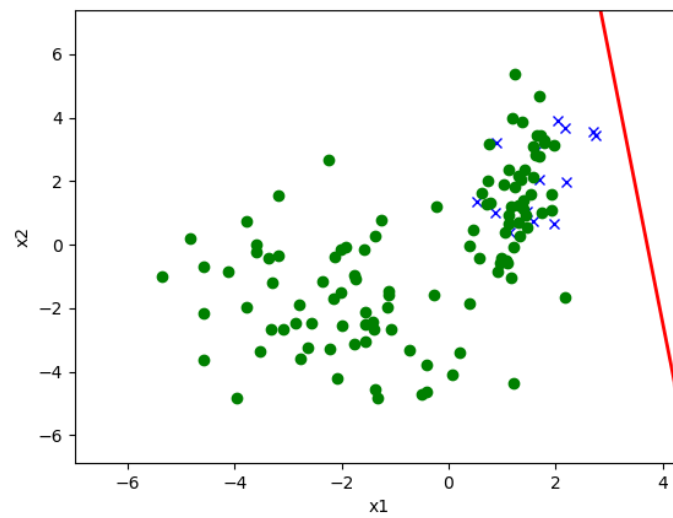
$$h(x) \approx p(y = 1|x) = p(t = 1|x)\alpha \approx \alpha \quad \text{for all } x \in V_+$$

(c)



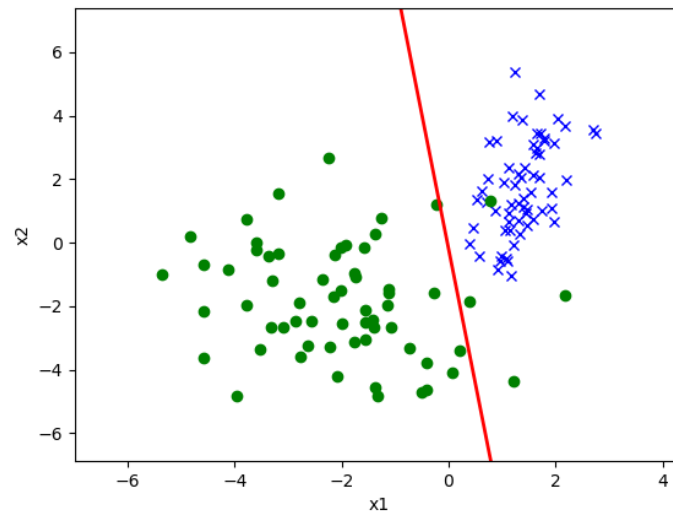
train use t-label

(d)



train use y-label

(e)



train use y-label, rescale by α

3.