Assignment 8
Due Nov. 18, 2023
Torin White
UIN: 657467127

Q1. [100 pt.] Logistic regression model w/ parameters w_0, w_1, w_2

$$P(y=1|x_1,x_2) = \frac{2^{w_2x_2+w_1x_1+w_0}}{1+2^{w_2x_2+w_1x_1+w_0}}$$

a. [20 pt] For weights $w_0 = -.5$, $w_1 = 1$, $w_2 = 1$ draw the decision boundary

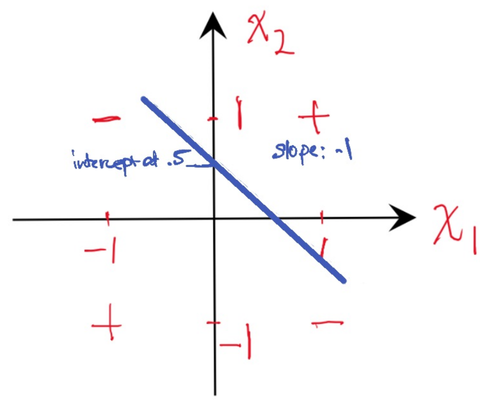
decision boundary is found where:

$0 = w_2x_2 + w_1x_1 + w_0$

$0 = x_1 + x_2 - .5$

$x_2 = -x_1 + .5$

line w/ slope 1
& intercept .5 on
x_2 axis



b. [20 pt] What is the log likelihood of the negative data point $(x_1 = -1, x_2 = 1)$ i.e. The value of $\log_2 P(Y=0|x_1=-1, x_2=1)$ from a.?

$P(Y=0|x_1,x_2) = 1 - P(Y=1|x_1,x_2)$

with weights and datapoints: we get

$w_0 = -.5$
$w_1 = 1$
$w_2 = 1$
$x_1 = -1$
$x_2 = 1$

$P(Y=0|x_1=-1, x_2=1) = \log_2 \frac{1}{1+2^{(1\cdot 1)+(1\cdot -1)+(-.5)}}$

* the $w_2x_2 + w_1x_1$ terms cancel, leaving $-.5$

$= \log_2 \frac{1}{1+2^{(-.5)}} = -\log_2(1+2^{(-.5)})$

$= -\log_2(1.707) = -.771$

c. what is the gradient of this point?

$\frac{\partial}{\partial w_0} \log_2 P(Y=0 \mid x_1 = -1, x_2 = 1) = ?$

$\frac{\partial}{\partial w_1} \log_2 P(Y=0 \mid x_1 = -1, x_2 = 1) = ?$

$\frac{\partial}{\partial w_2} \log_2 P(Y=0 \mid x_1 = -1, x_2 = 1) = ?$

$$\frac{\partial}{\partial w_i} \log_2 P(Y=0 \mid x_1, x_2) = -\frac{\partial}{\partial w_i} \log_2 (1 + 2^{w_2x_2+w_1x_1+w_0})$$

$$= -\frac{1}{1+2^{w_2x_2+w_1x_1+w_0}} \cdot 2^{w_2x_2+w_1x_1+w_0} \frac{\partial}{\partial w_i}(w_2x_2 + w_1x_1 + w_0)$$

$$2^{w_2x_2+w_1x_1+w_0} = 2^{\cancel{(1 \cdot 1)+(-1 \cdot 1)}+.5} = 2^{.5} = .707$$

$$= -\frac{1}{1+.707} \cdot .707 = \frac{-.707}{1.707} = .414$$

$w_0 = -.414 \cdot \frac{\partial}{\partial w_0} = -.414$

$w_1 = -.414 \cdot \frac{\partial}{\partial w_1} = .414$

$w_2 = -.414 \cdot \frac{\partial}{\partial w_2} = -.414$

d. $$P(y=1 \mid x_1, x_2) = \frac{2^{w_3(x_1x_2)+w_2x_2+w_1x_1+w_0}}{1+2^{w_3(x_1x_2)+w_2x_2+w_1x_1+w_0}}$$

with additional feature $x_1 x_2$ what weights provide good fit?

As far as the properties that need to be satisfied, the probability must be $> .5$ for positive examples, & that can be satisfied by $w_3(x_1x_2) + w_2x_2 + w_1x_1 + w_0 > 0$. For negative examples, it should be the opposite, summing to < 0.

for each example:

(negative examples)	(positive examples)
$- (-1,1)$ $(-1)(1)w_3 + (1)w_2 + (-1)w_1 + w_0$	$+ (-1,-1$ $(-1)(-1)w_3 + (-1)w_2 + (-1)w_1 + w_0 > 0$
$- (1,-1)$ $(1)(-1)w_3 + (-1)w_2 + (1$ w_0	$+ (1,1$ $(1)(1)w_3 + w_2 + w_1 + w_0 > 0$

if $w_0 = w_1 = w_2 = 0$, any positive value of w_3 will suffice such as $w_3 = 2$

e. the decision boundary is

- $x_1 = 0$, $x_2 = 0$

and on the 2-dimensional axis falls on the x_1, x_2 axis.

positive zone in yellow

negative zone pink