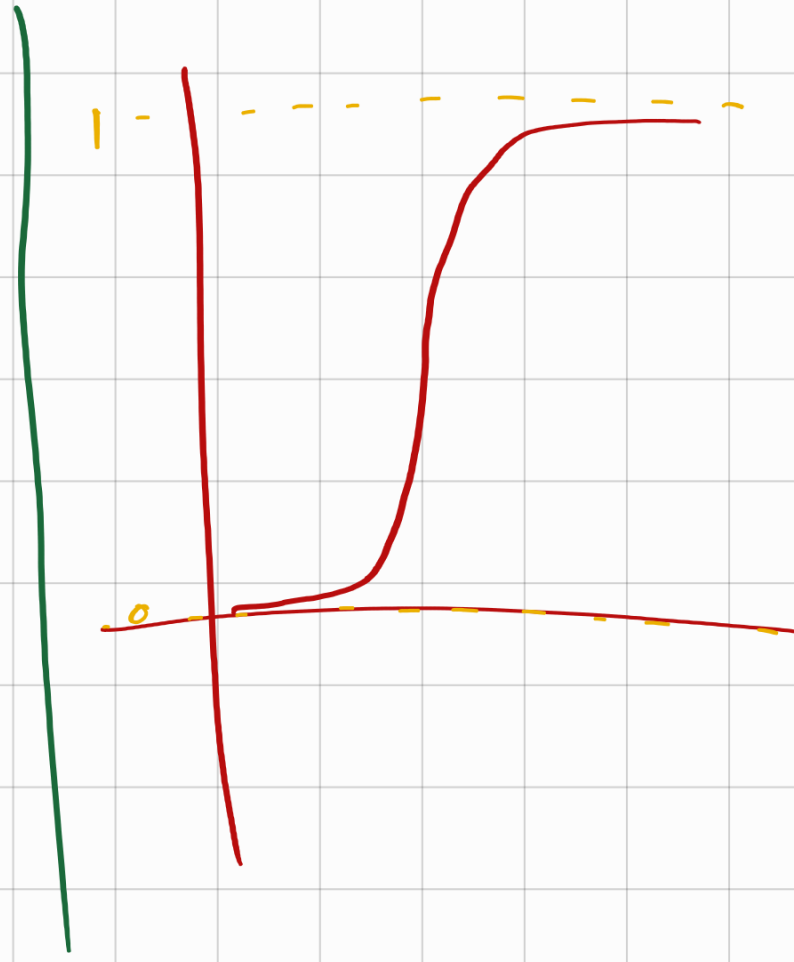
$$y = \frac{1}{1 + e^{-(a+bx)}}$$



$\Rightarrow$ We want to Minimize log loss

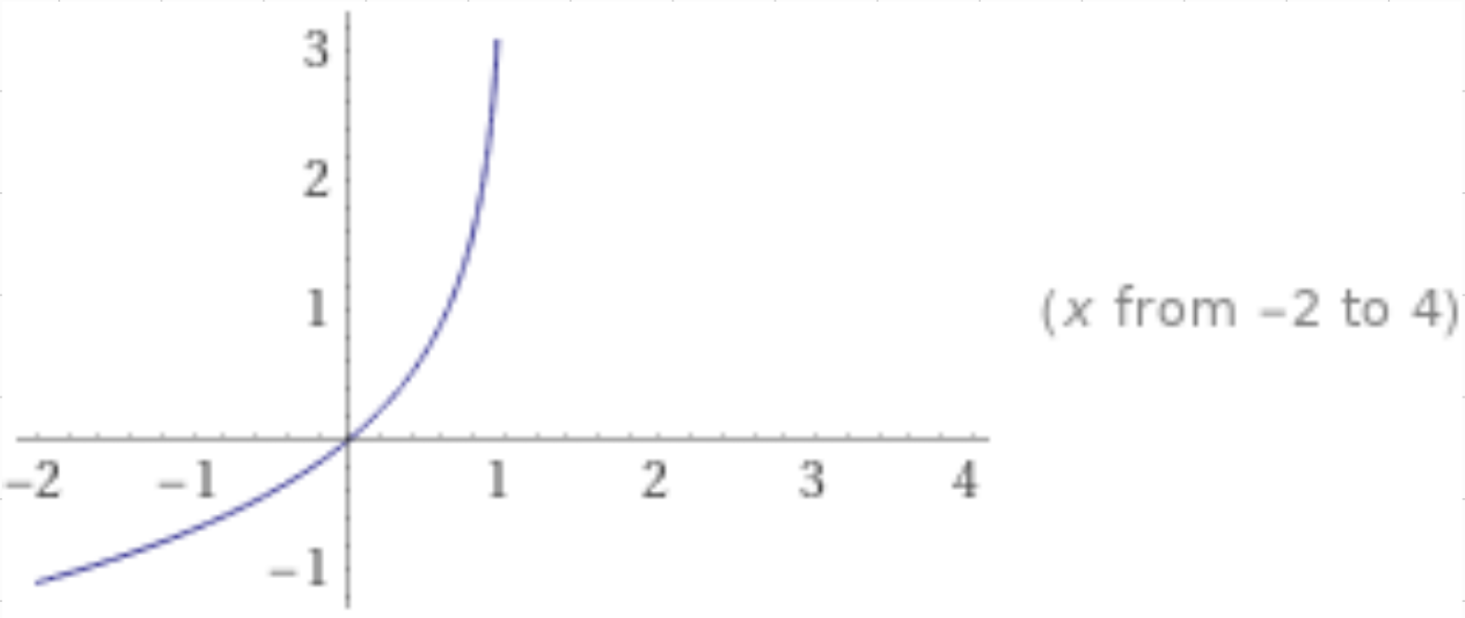
$$\text{Log loss} = \left(-y \log(\hat{y}) - (1-y) \log(1-\hat{y})\right)$$

y = observed

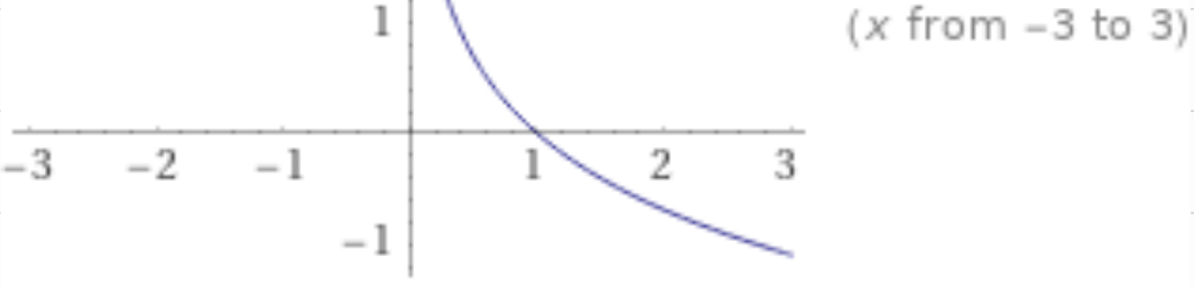
$\hat{y}$ = Predicted.

$$\begin{cases} -\log(\hat{y}) & \text{if } y=1 \\ -\log(1-\hat{y}) & \text{if } y=0 \end{cases}$$

$y = -\log(1-x)$



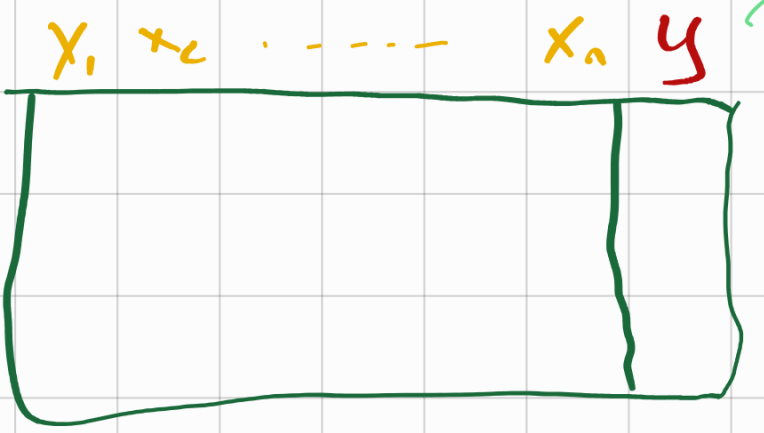
$-\log(x)$



final Logit function

$$y = \frac{1}{e^{-(a + b_1x_1 + b_2x_2 \cdots + b_nx_n)}}$$

Thresholds



→ $\hat{y}$ is a probability

=> When deciding on the threshold for logit functions,

Put it where outliers or mis-predictions will have a very low cost

Confusion Matrix

Pred \ Truth	P	N
P	True positives (TP)	false positives (FP)
N	false negatives (FN)	True negative (TN)

Measures

(1) Accuracy

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

(2) Recall (Sensitivity / True Positive rate)

$$\text{Recall} = \frac{TP}{TP + FN}$$

} out of all True positive

what fraction did we identify?

(3) Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$