01/26/23

data

50% cancer

50% not

baseline accuracy: 50%

data:

5% cancer

95% not

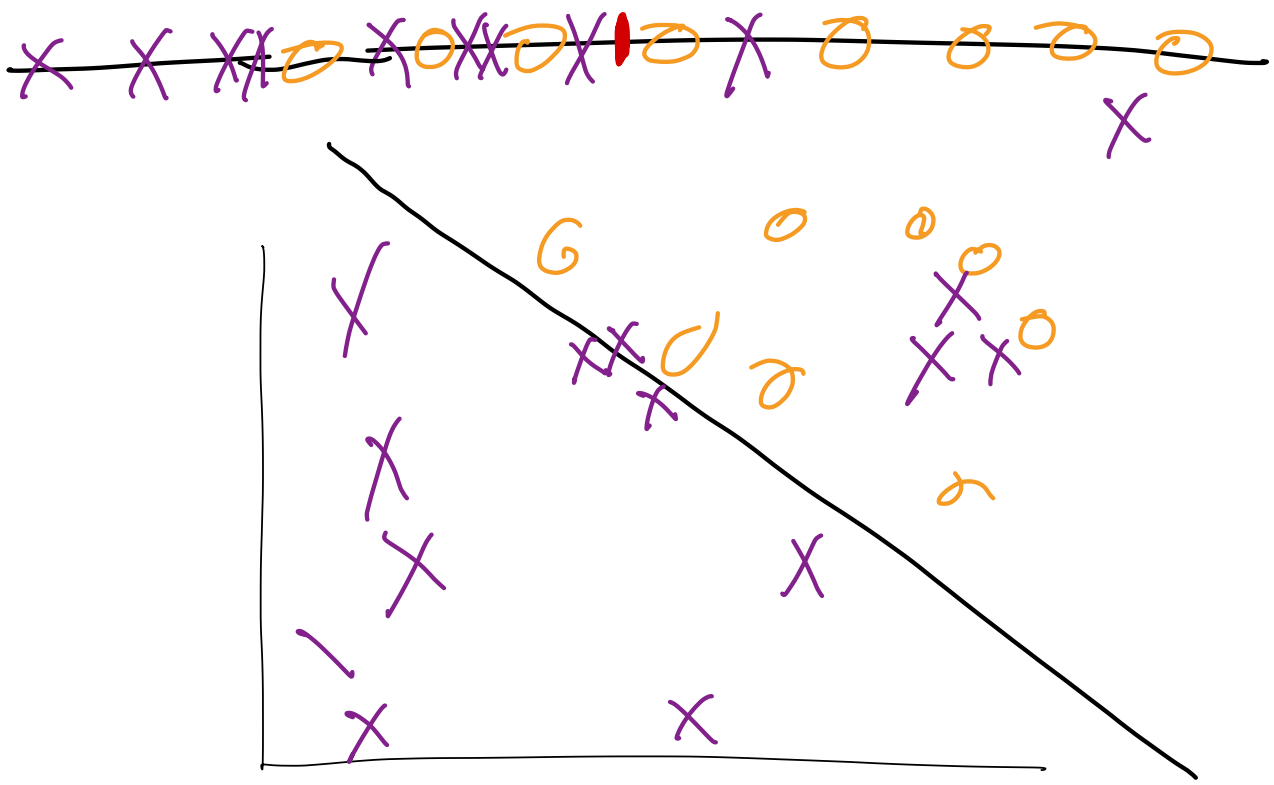
baseline accuracy: 95%

SVM: support vector machines

snake
crocodile

$$\chi : \{(\vec{x}_i, y_i)\}$$

$+1 \quad -1$



$\text{sign}\left(\langle \vec{a}, \vec{x} \rangle + b\right)$

$\langle \vec{a}, \vec{x} \rangle$

$\vec{a}^T \vec{x}$

$\vec{a} \cdot \vec{x}$

one feature

$\text{sign}(ax + b)$

$ax + b = 0$

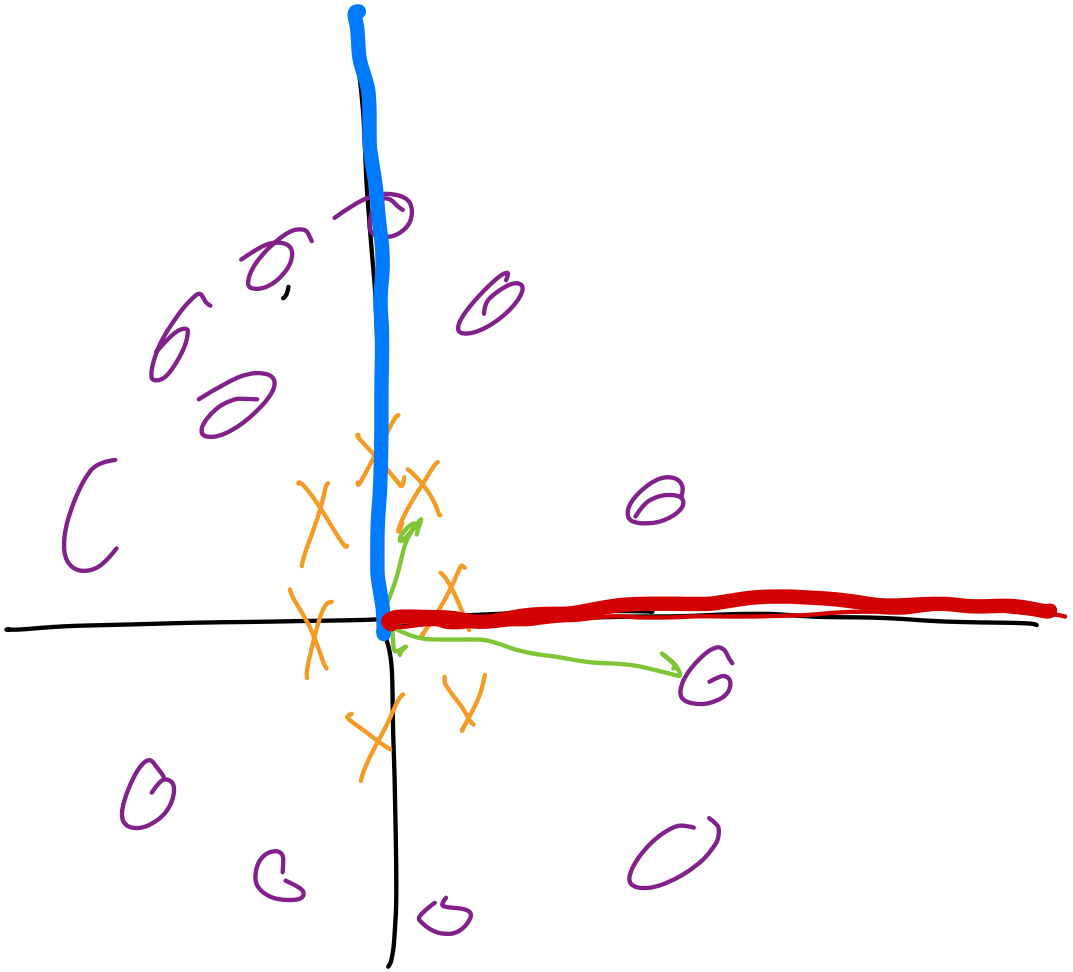
$ax = -b$

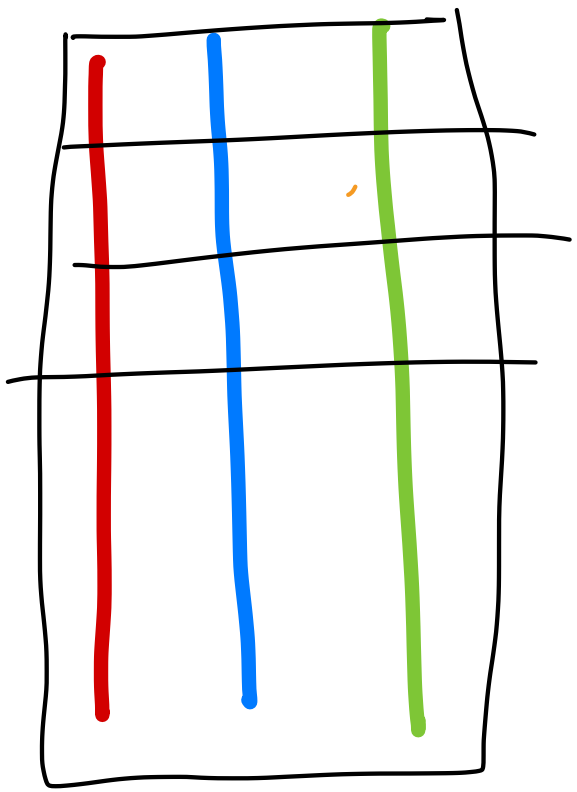
$x = -\frac{b}{a}$

two features

$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$a_1 x_1 + a_2 x_2 + b$

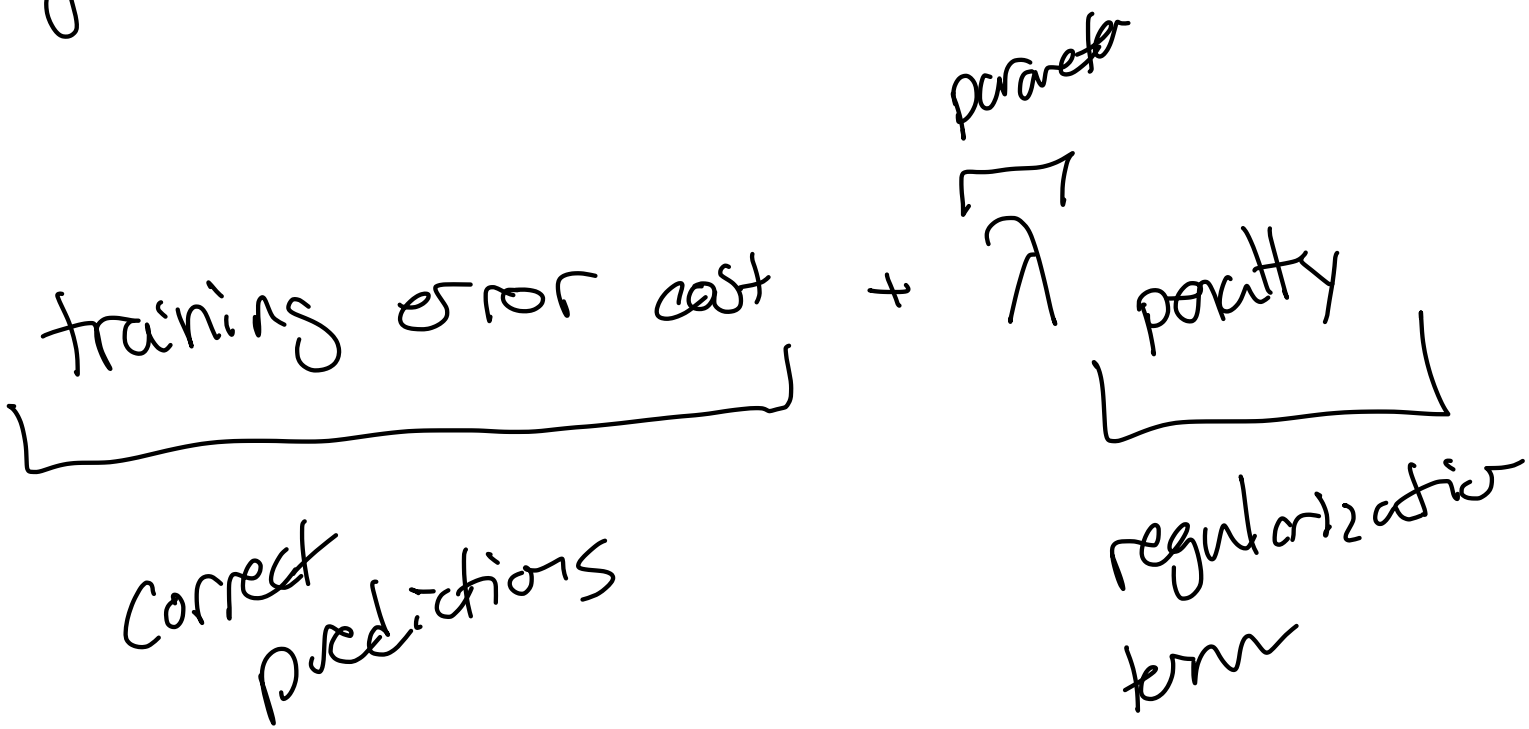




$$\langle \vec{a}, \vec{x} \rangle + b$$

1) ensure predicted correctly
training samples on correct side

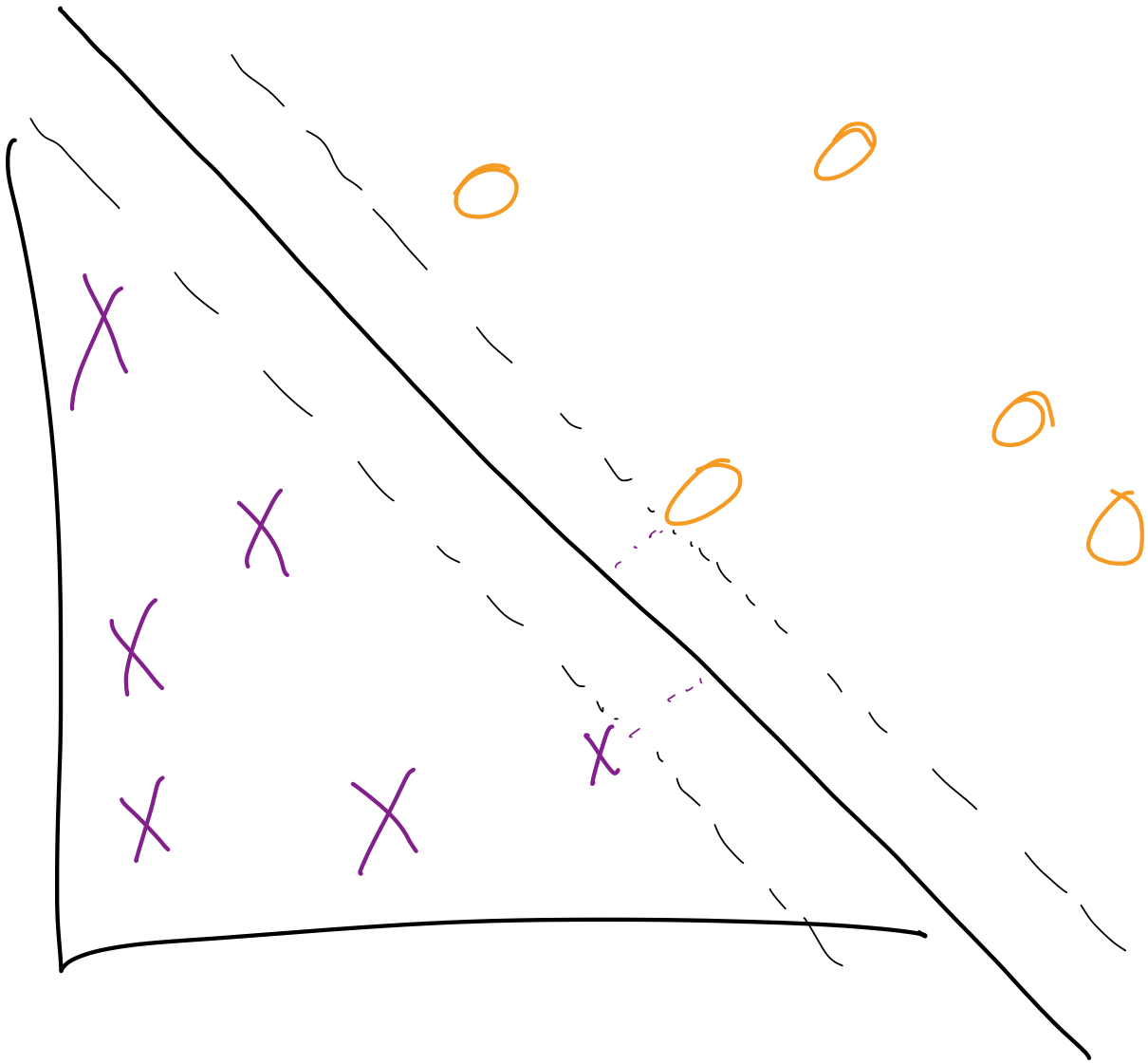
2) generalize to unseen data



$\ell\left(y_i, \langle \vec{a}, \vec{x} \rangle + b\right)$

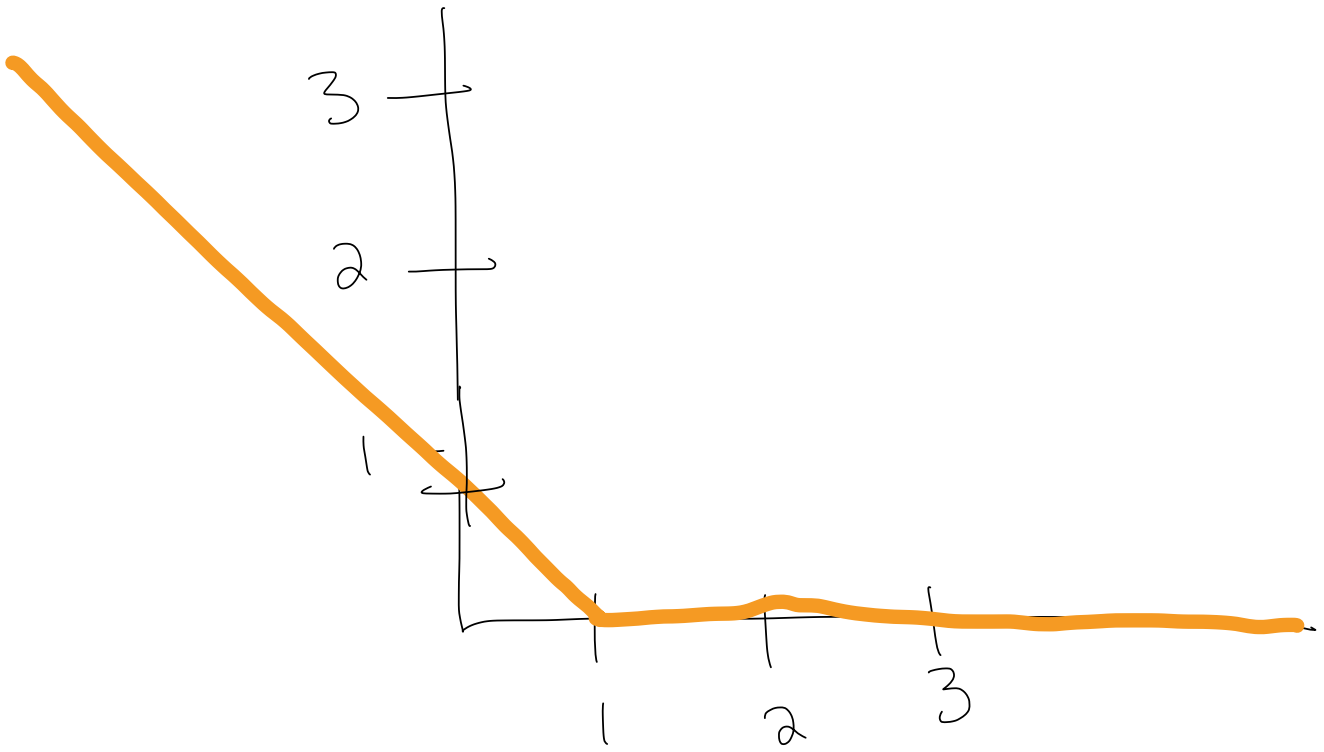
y_i and $\langle \vec{a}, \vec{x} \rangle + b$ diff signs

↳ want ℓ large



hinge loss

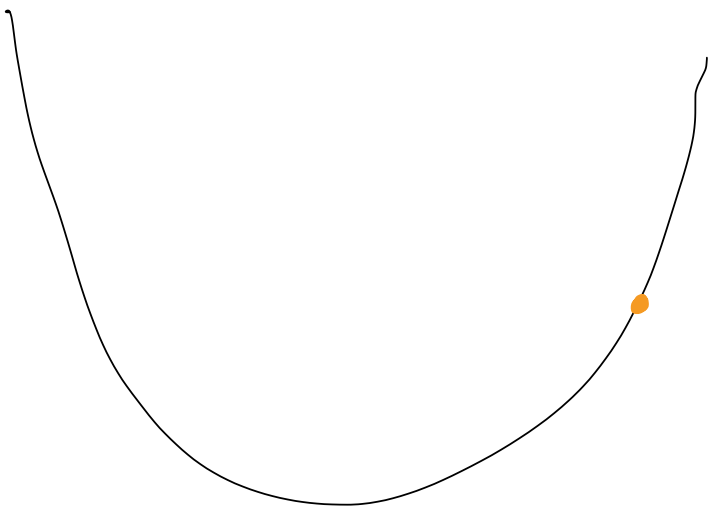
$$\ell\left(y_i, \langle \vec{a}, \vec{x}_i \rangle + b\right) = \max\left(0, 1 - y_i\left(\langle \vec{a}, \vec{x}_i \rangle + b\right)\right)$$



$$f = \left[\frac{1}{N} \sum_{i=1}^{N} \max\left(0, 1 - y_i(\langle \vec{a}, \vec{x}_i \rangle + b)\right) \right] + \lambda \frac{\|a\|_2^2}{2}$$

$$f(\vec{v}) = v_1^2 + 2v_2 + 6v_1 v_2$$

$$\nabla f = \begin{bmatrix} \partial f / \partial v_1 \\ \partial f / \partial v_2 \end{bmatrix}$$



norms

$$\|a\|_2 = \sqrt{\sum (a_i^2)} = \sqrt{\langle a, a \rangle}$$

$$\|a\|_1 = \sum |a_i|$$