FoDA L28

Classification: KNNs, Decision Trees, and Neural Nets ... oh my!

Dec 6, 2022

Classification

$(x, y) \underset{iid}{\sim} \mathcal{U}$

Input $(X, y) \subset \mathbb{R}^d \times \{-1, +1\}$
$\hookrightarrow (x_i, y_i) \in \mathbb{R}^d \quad \{-1, +1\}$
$\qquad\qquad\quad x_i \qquad\quad y_i$

$\rightarrow$ Train
$\rightarrow$ Test

Goal Build function $f : \mathbb{R}^d \to \{-1, +1\}$

$$f(x) = \text{sign}(g(x)) \qquad \underline{g(x) : \mathbb{R}^d \to \mathbb{R}}$$

So on new data $(x', y') \sim \mathcal{U}$

that $f(x') = y'$ w/ high probability

K - Nearest Neighbor Classifier (KNN)

$k = 5$

$X \subset \mathbb{R}^d$

$S_k(x') \subset X$

vote on signs in $S_k(x')$

x'

going

4 + 1 -

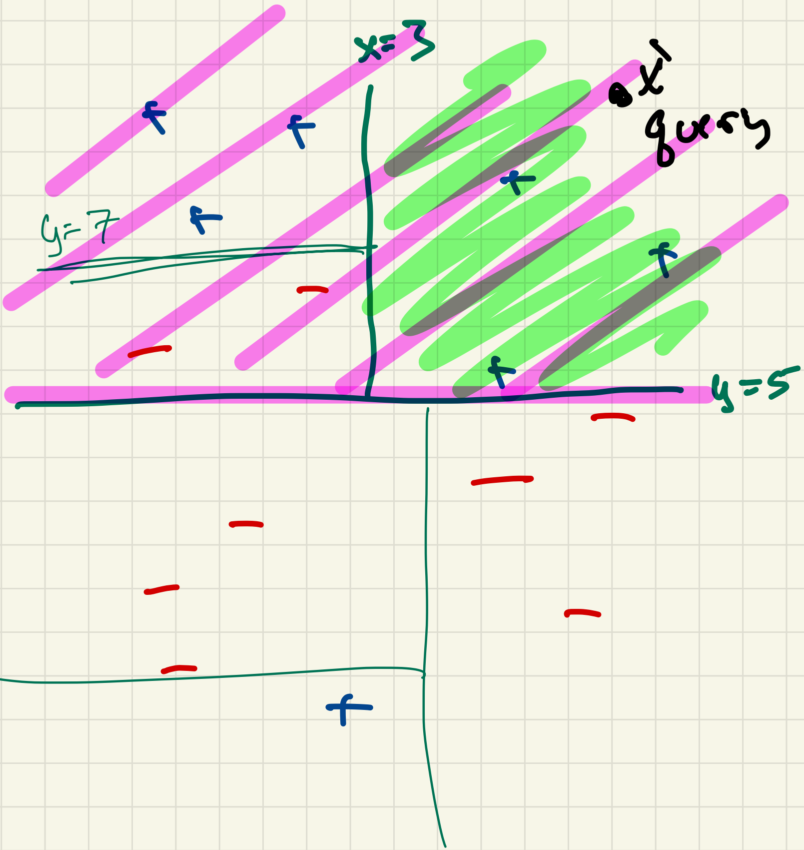
$f(x') = +$

x_i

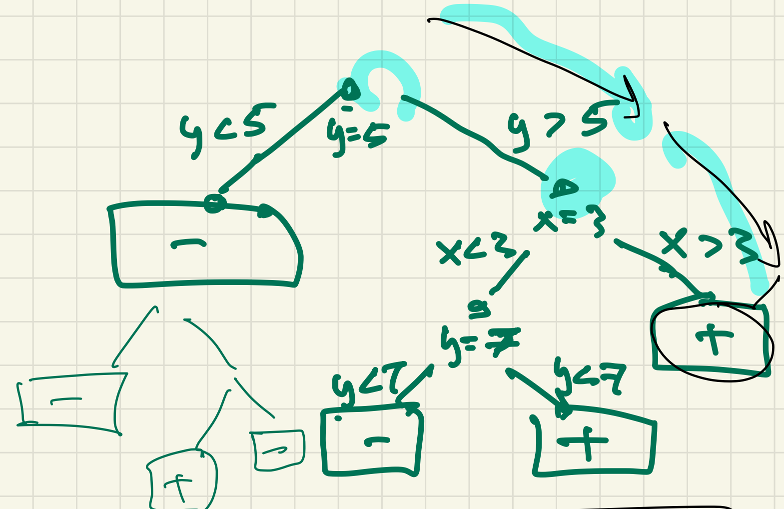
choice of distance

Decision Trees

$X \subset \mathbb{R}^d$



$f(x') = +$



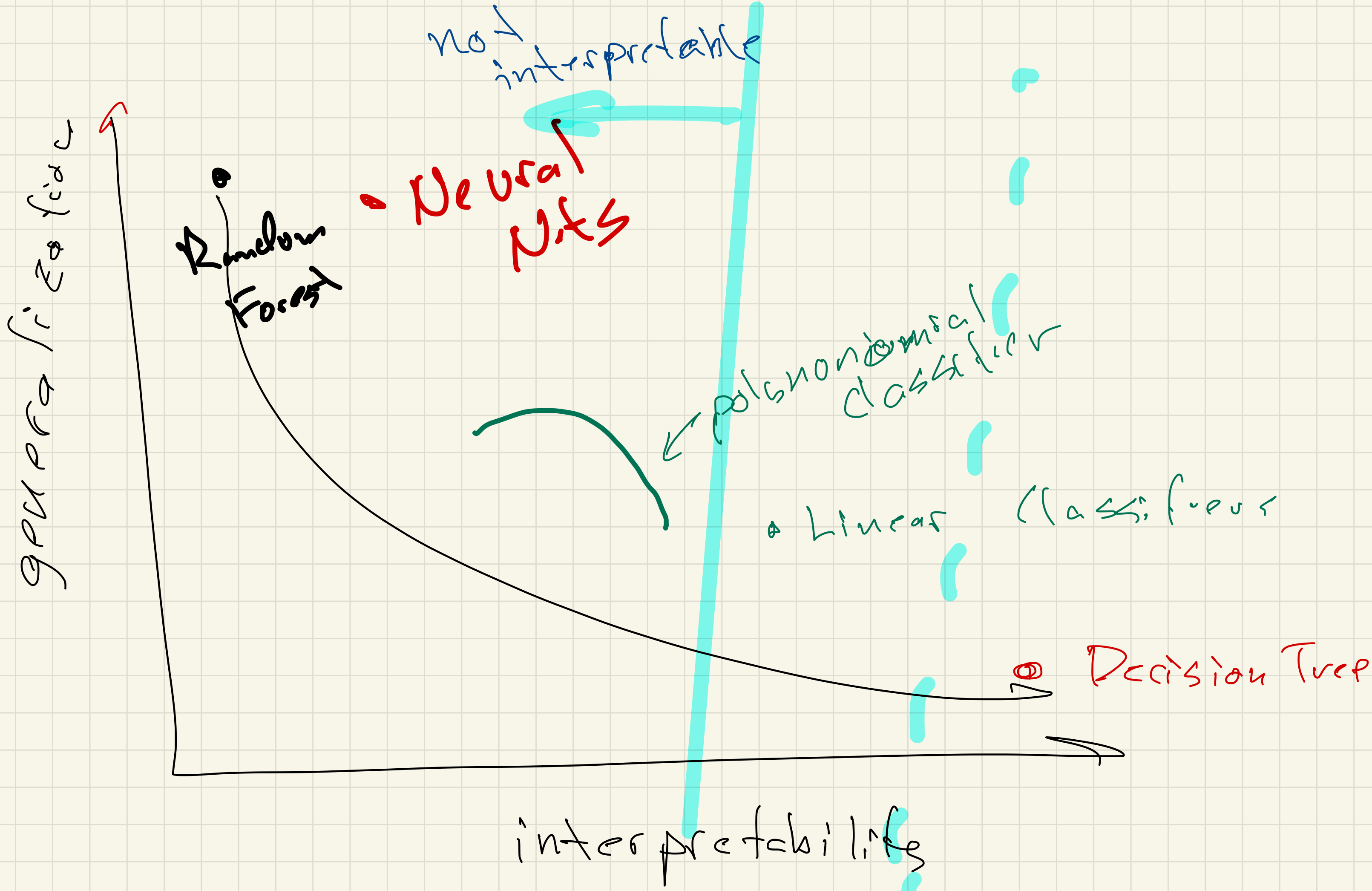
How to split?

- choose best

Gini Impurity Index

$$P_+ = \frac{\#\,+\text{pts}}{\text{total}\ \#\,\text{pts}} = \frac{6}{8} \qquad P_- = \frac{2}{8}$$

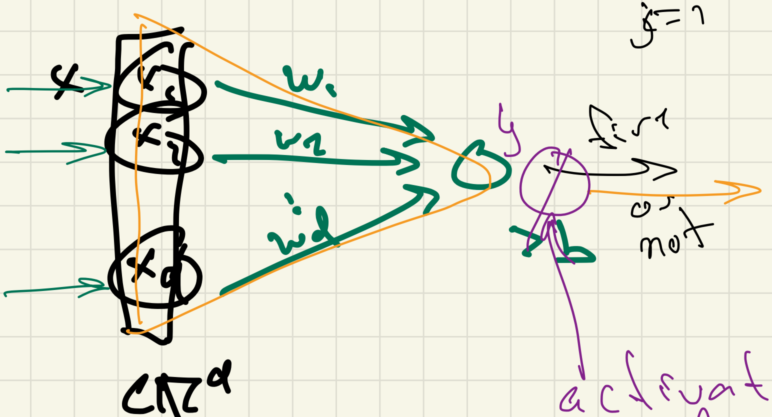
$$G = \overset{(0.75)}{P_+}\overset{(1-0.75)}{(1-P_+)} + \overset{(0.25)}{P_-}\overset{(1-0.25)}{(1-P_-)} = \frac{6}{16}$$



generalization
not interpretable
Random Forest
Neural Nets
polynomial classifier
Linear Classifiers
Decision Tree
interpretability

Neural Nets

Neuron



$$\sum_{j=1}^{d} x_j w_j \geq b$$

$$\iff$$

$$\langle w, x\rangle - b \geq 0$$

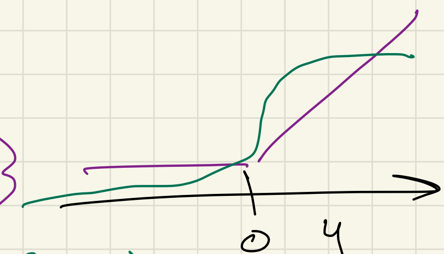
$y = g_{w,b}(x)$

linear classifier

activation function

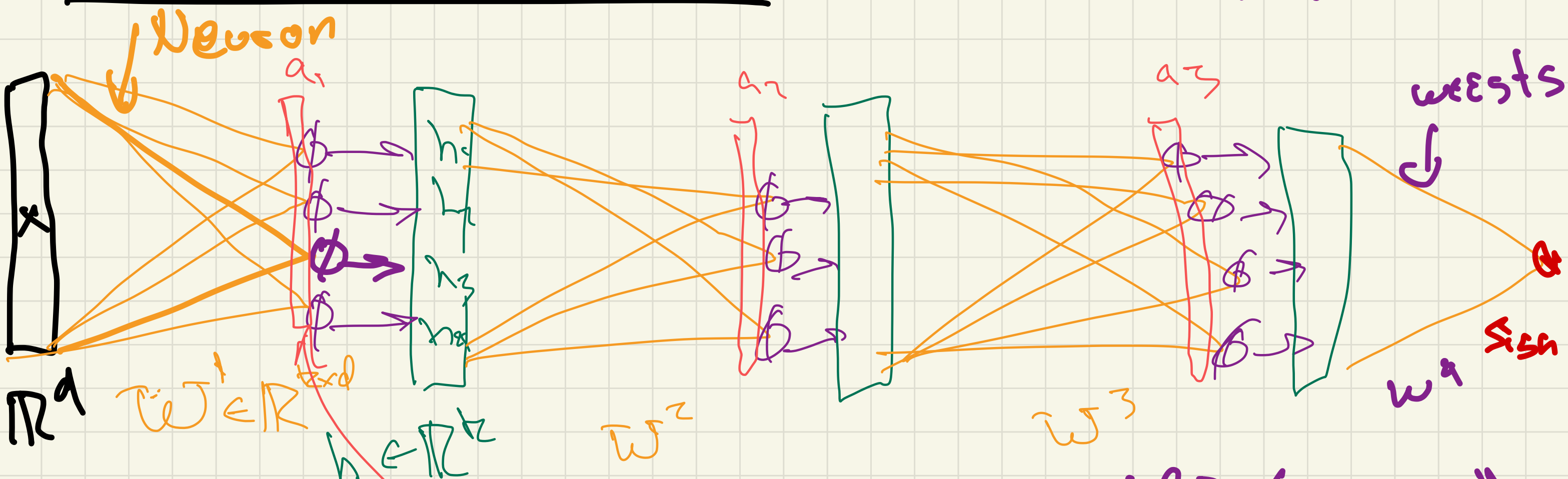
$$\phi(y) = \mathrm{ReLU}(y) = \max\{0, y\}$$

$$\phi(y) = \mathrm{Sigmoid}(y) = \frac{1}{1+\exp(-y)} \in (0,1)$$



Neural Network

backpropagation



$$g(x) = w^4 \phi(W^3 \phi(W^2 \phi(W^1 x)))$$

$$W = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_k \end{bmatrix}$$

$w_i \in \mathbb{R}^d$

$a = W^1 x$

w/ no activation

$a_2 = W^2 W^1 x$

$a_3 = \underbrace{W^3 W^2 W^1}_{} x$

$B = W^3 W^2 W^1$