

Causal Inference, Interpretability, Summary

Supervised Learning



Online Learning

perceptron
winnow
multi-armed bandits

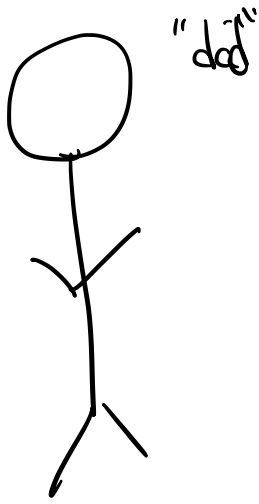
Evaluation

ROC, AUC, ...
cross-val

Unsupervised Learning

K-means, EM & GMM
hierarchical clustering

Is ML useless for making decisions? Possibly.



male

68 yrs

history of congenital heart failure

atrial fib

taken aspirin

Use ML to predict stroke within 1 year

$f(x)$ = fcn of (gender, race, age, medical history, drugs, etc.)

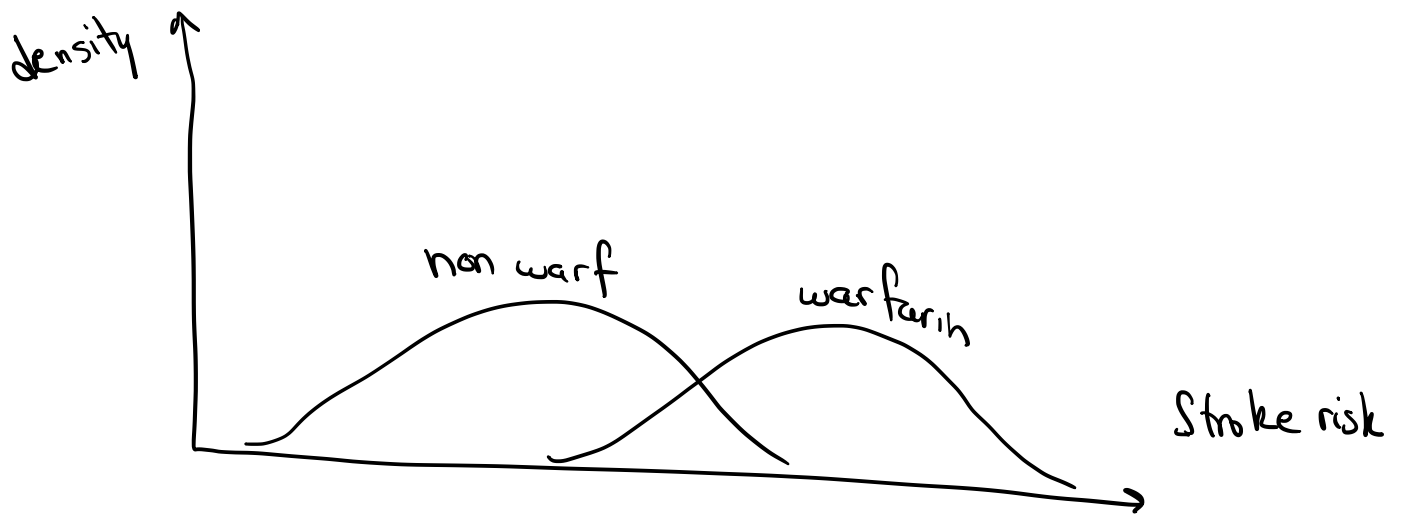
$$f(x) = 3.4 \cdot \mathbb{1}_{[\text{congenital} \dots]} + 4.2 \cdot \mathbb{1}_{[\text{a fib}]} + \dots \\ + 1.1 \cdot \mathbb{1}_{[\text{female}]} + 2.4 \cdot \mathbb{1}_{[\text{smoking}]}$$

If dad stops smoking, does it reduce his risk of stroke?

$$f(x) = 2 \cdot \mathbb{1}_{[\text{warfarin}]} - 1 = \begin{cases} \text{predict stroke if} \\ \text{on warfarin} \\ \text{predict no stroke if not} \end{cases}$$

correlation \nRightarrow causation

What happened?

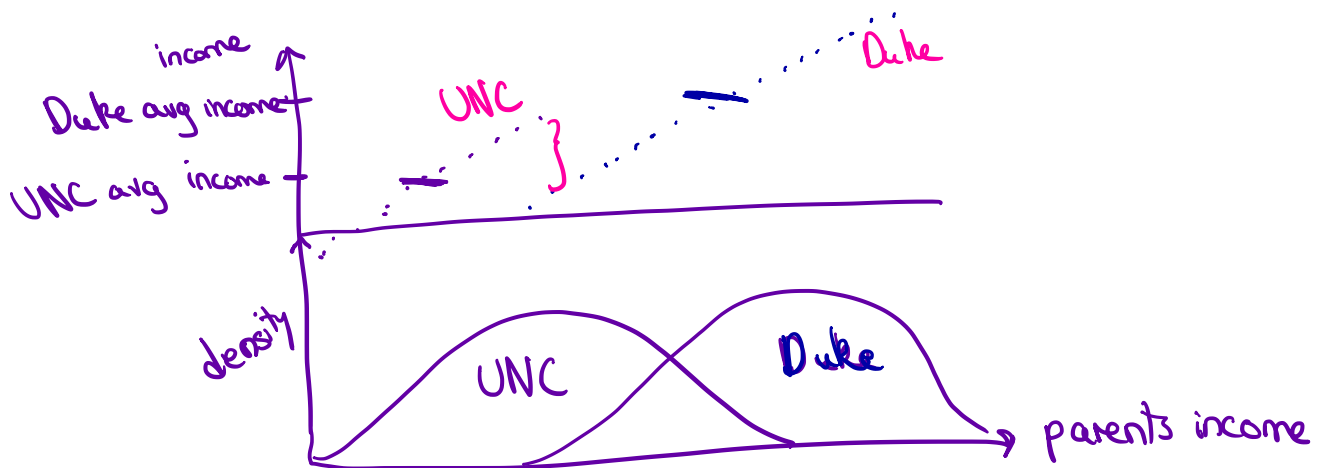


Regular ML: model $Y | X$

"Causal inference" ML: model $Y | X, \underbrace{T=1}_{\text{treated}} - Y | X, \underbrace{T=0}_{\text{control}}$

treatment effect
conditional difference

Example: predict income, given age, parent income, major, ...
treatment is education at Duke vs. education at UNC.



What data do you get?

X, T, Y

parent income	UNC/Duke	income
200K	Duke	80K
⋮	UNC	30K
⋮	⋮	⋮

To do supervised learning, need X, treatment effect $Y|T=1 - Y|T=0$

↑
almost never have that!

problem is exactly half-supervised.

How to use ML to solve it:

① Make assumptions

typical assumptions: strong ignorability - no unmeasured confounding

SUTVA - no unit's outcome affects another's treatment

$$0 < P(T=0 | X) < 1 \quad 0 < P(T=1 | X) < 1$$

② regress separately on treatment & control
put it all into a big ML alg (BART)

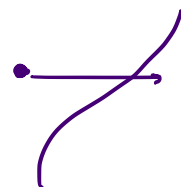


matching or nearest neighbor

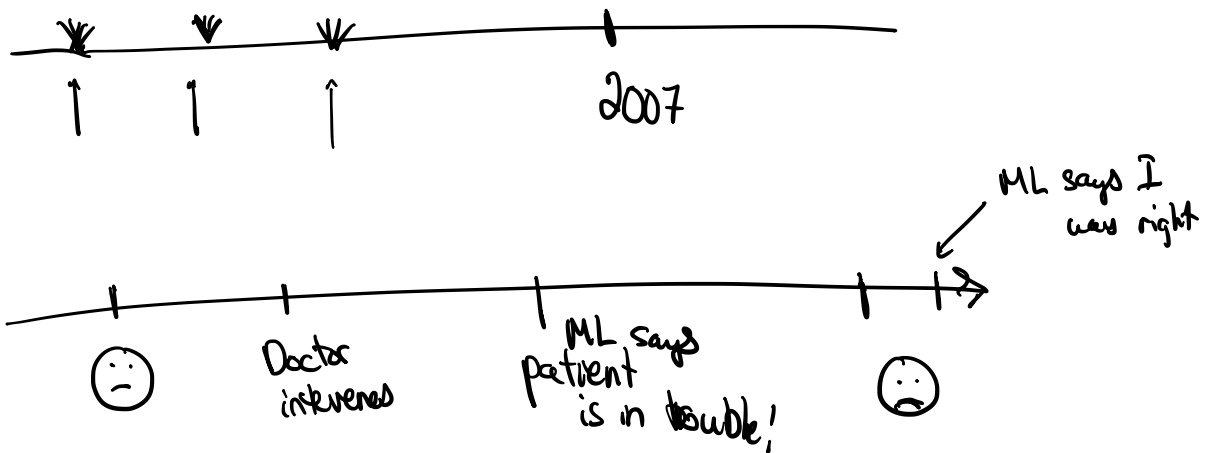
*beware" of MLers who don't know CI!



DON'T try to alter features in order to get better outcomes unless those features are causal!



"data leakage"



- You can generally find an interpretable model that is as accurate as any black box model.

- It has become clear that people don't know this:

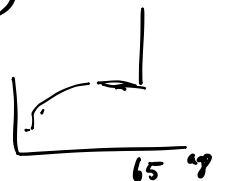
- parole hearings

- google Breezometer

- credit risk

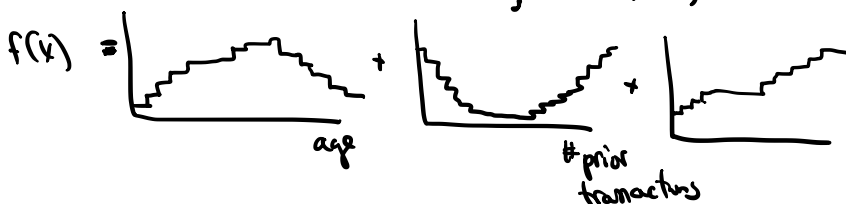
- What can you actually use? Least sq, log reg, dec trees (see some of our lab's work...)

KNN, additive models



additive models - credit risk

$$x_j > \theta_1, \dots, x_j > \theta_2, \dots$$



$$f(x) = \sum_j \alpha_j h_j(x_j)$$