It's NOT all about Deep Learning: The Case for Simpler Models

SYSTOPIA

UBC

Margo Seltzer
Co-Head, Computer Science
Canada 150 Research Chair in Computer Systems
The University of British Columbia

What do Simple Models have to do with Systems?

Types of ML Models/Frameworks

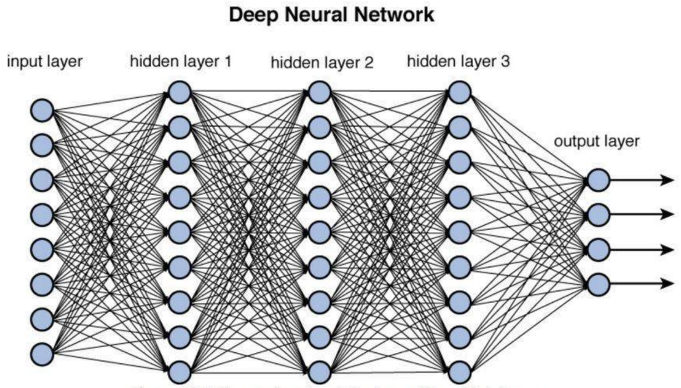
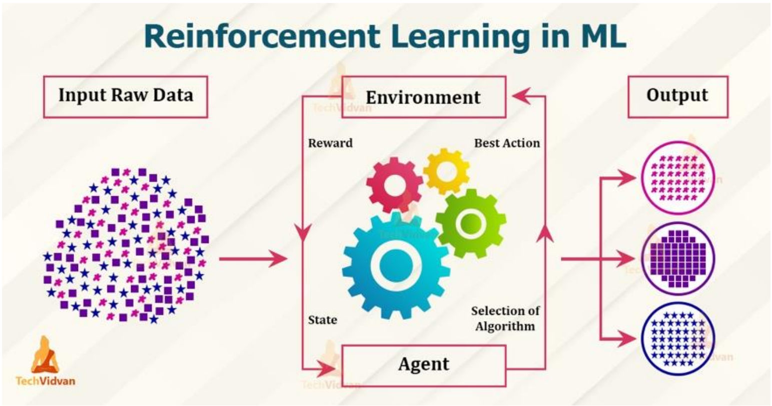


Figure 12.2 Deep network architecture with multiple layers.

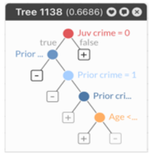


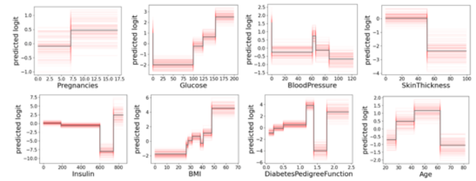
Interpretable Models

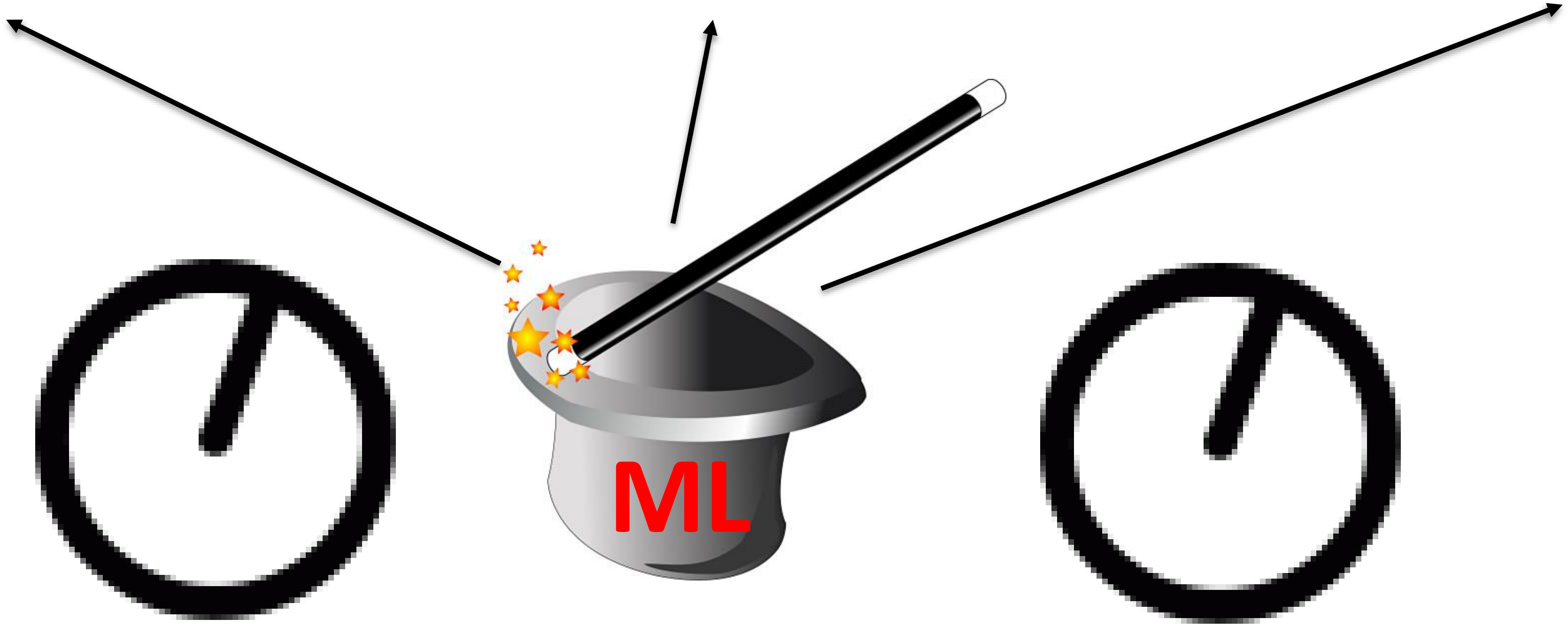
1.	Oval Shape	-2 points		...
2.	Irregular Shape	4 points	+	...
3.	Circumscribed Margin	-5 points	+	...
4.	Spiculated Margin	2 points	+	...
5.	Age ≥ 60	3 points	+	...
		SCORE	=	

SCORE	-7	-5	-4	-3	-2	-1
RISK	6.0%	10.6%	13.8%	17.9%	22.8%	28.6%

SCORE	0	1	2	3	4	≥ 5
RISK	35.2%	42.4%	50.0%	57.6%	64.8%	71.4%





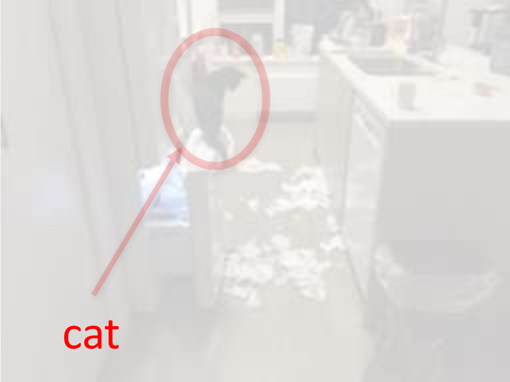


The Two Sides of Machine Learning

In French “chat chapeau”

There is a correct answer
Individual features not meaningful

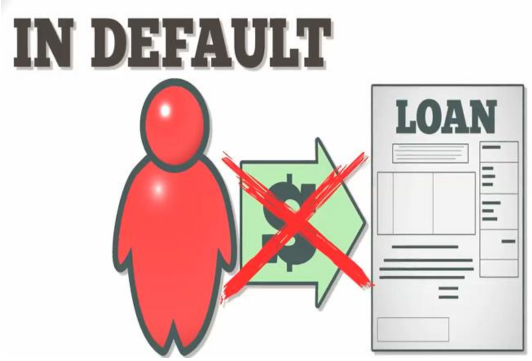
The Two Sides of Machine Learning



In French "chat chapeau"

Will this patient have a seizure?

Answers are probabilistic
Features are meaningful

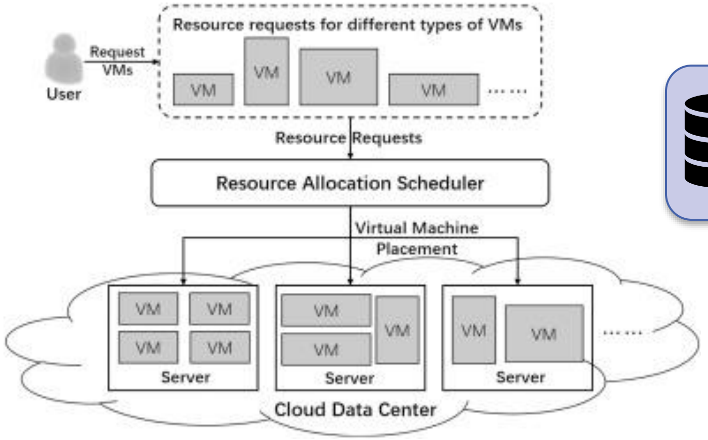


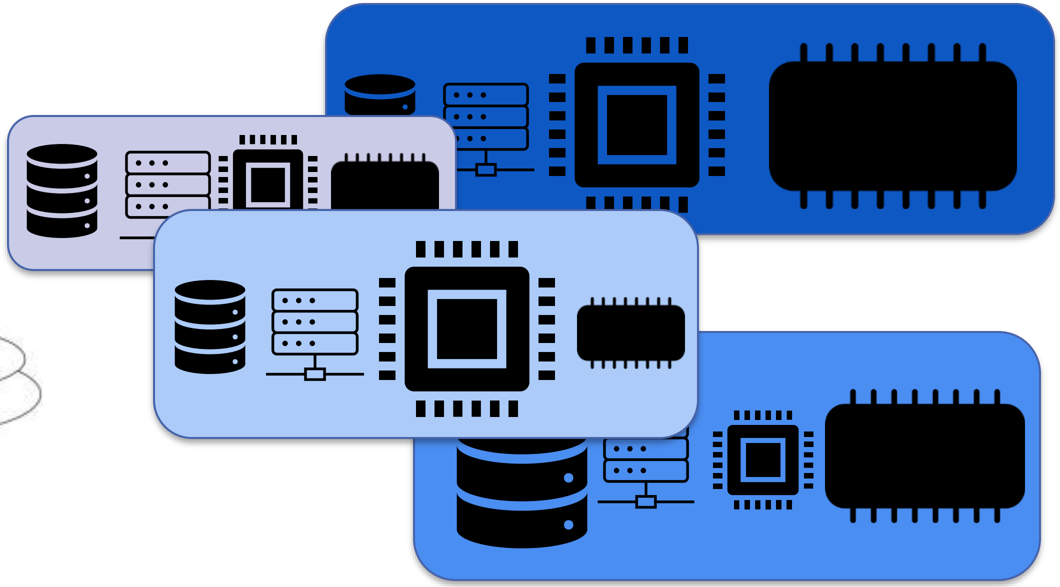
Who will default on a loan?

Is this person likely to commit another crime in the next two years?

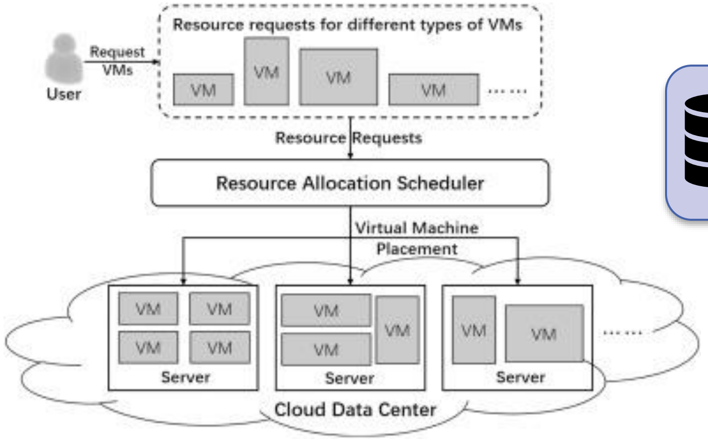
There is a correct answer
Individual features not meaningful

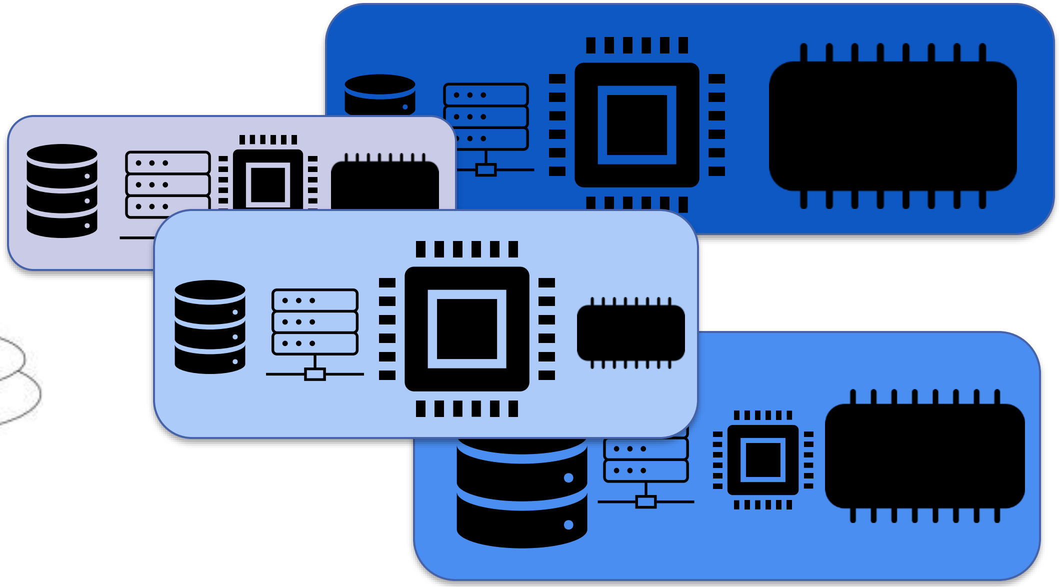
Which Kinds of Problems do we Have?

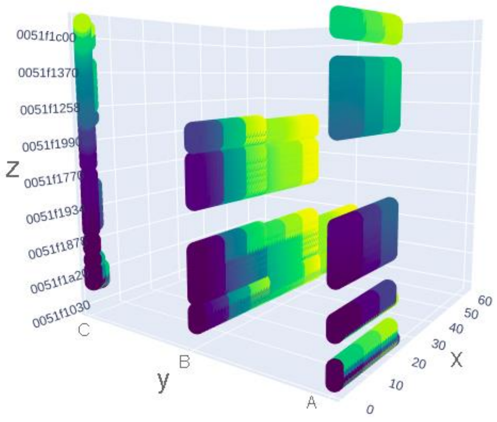




Which Kinds of Problems do we Have?





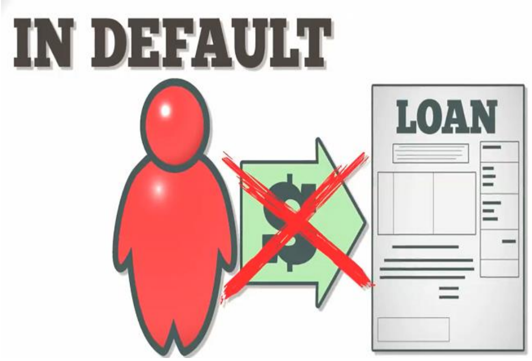


The Two Sides of Machine Learning

In French "chat chapeau"

Will this patient have a seizure?

Answers are probabilistic
Features are meaningful

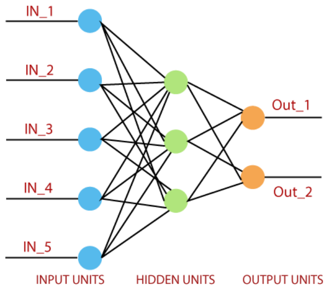


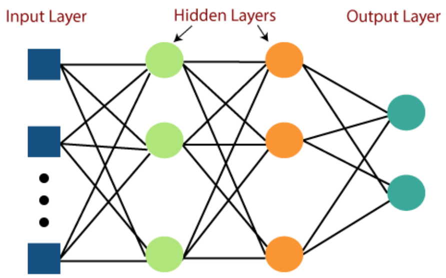
Who will default on a loan?

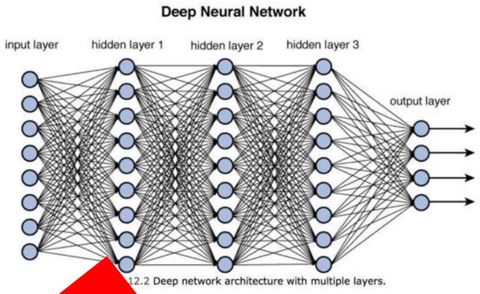
There is a correct answer
Individual features not meaningful

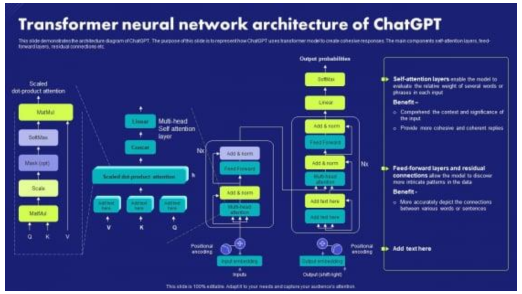
Is this person likely to commit another crime in the next two years?

Increase Model Complexity?



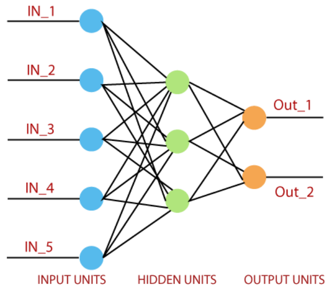


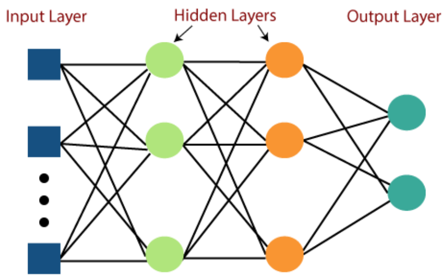


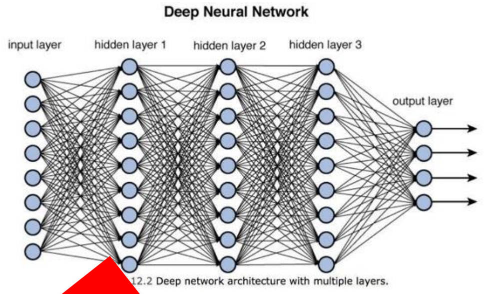


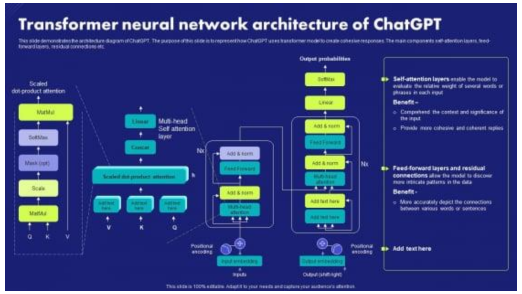
Increasi ex models

Increase Model Complexity?

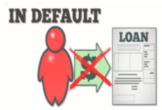








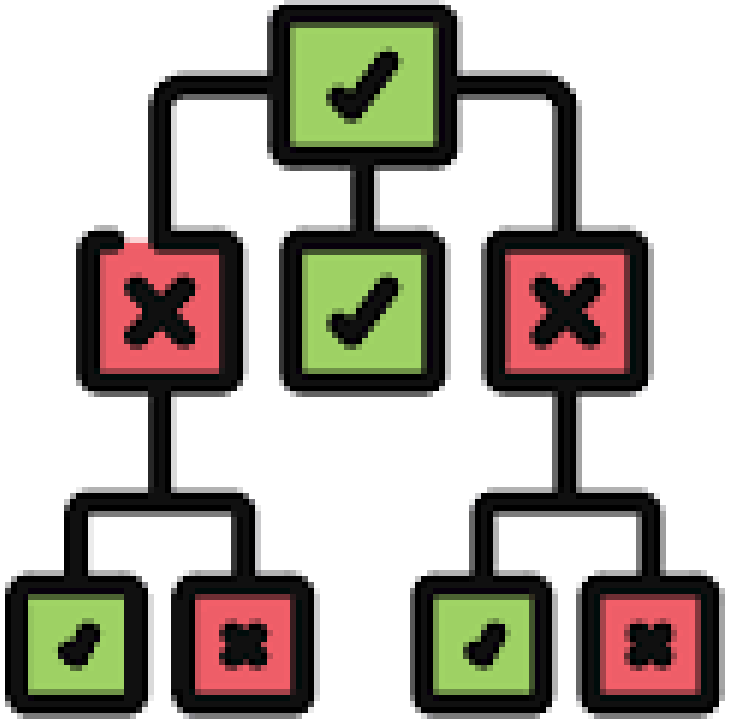
Increasi ex models



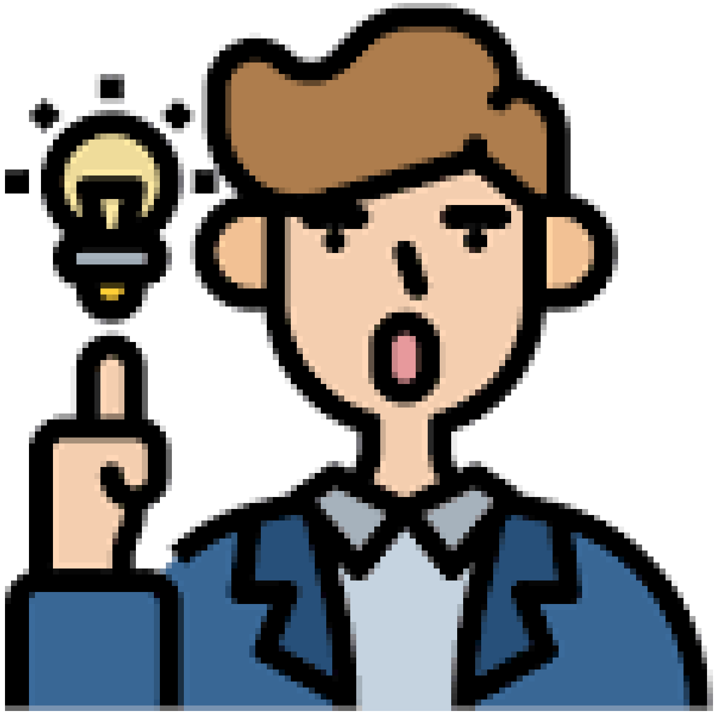
In cases where 1) Answers are probabilistic, and 2) Features are meaningful

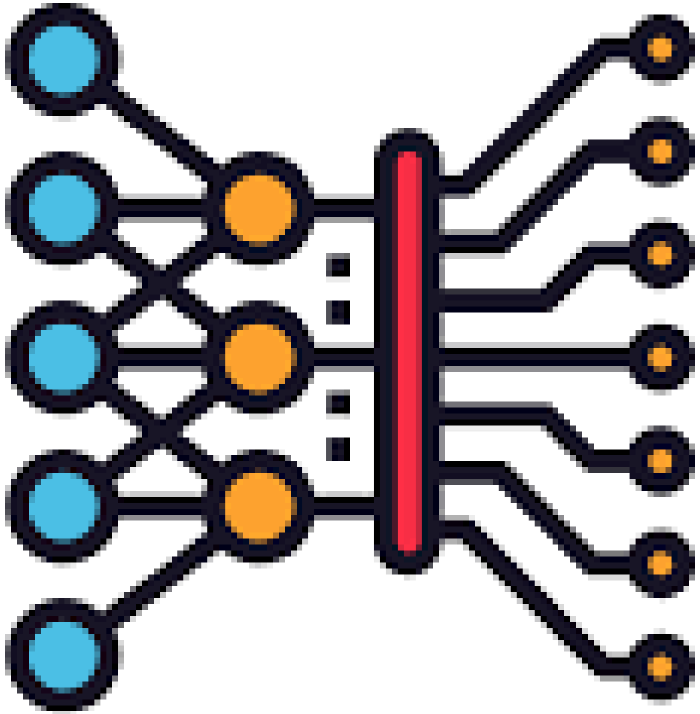
1. Deep learning does **NOT** help
2. You do **NOT** have to sacrifice accuracy to get interpretability.
3. If you discover that a bunch of different model classes perform equally well, there is an excellent chance that, you can find a simple model that is as accurate as any other.

A Brief Digression



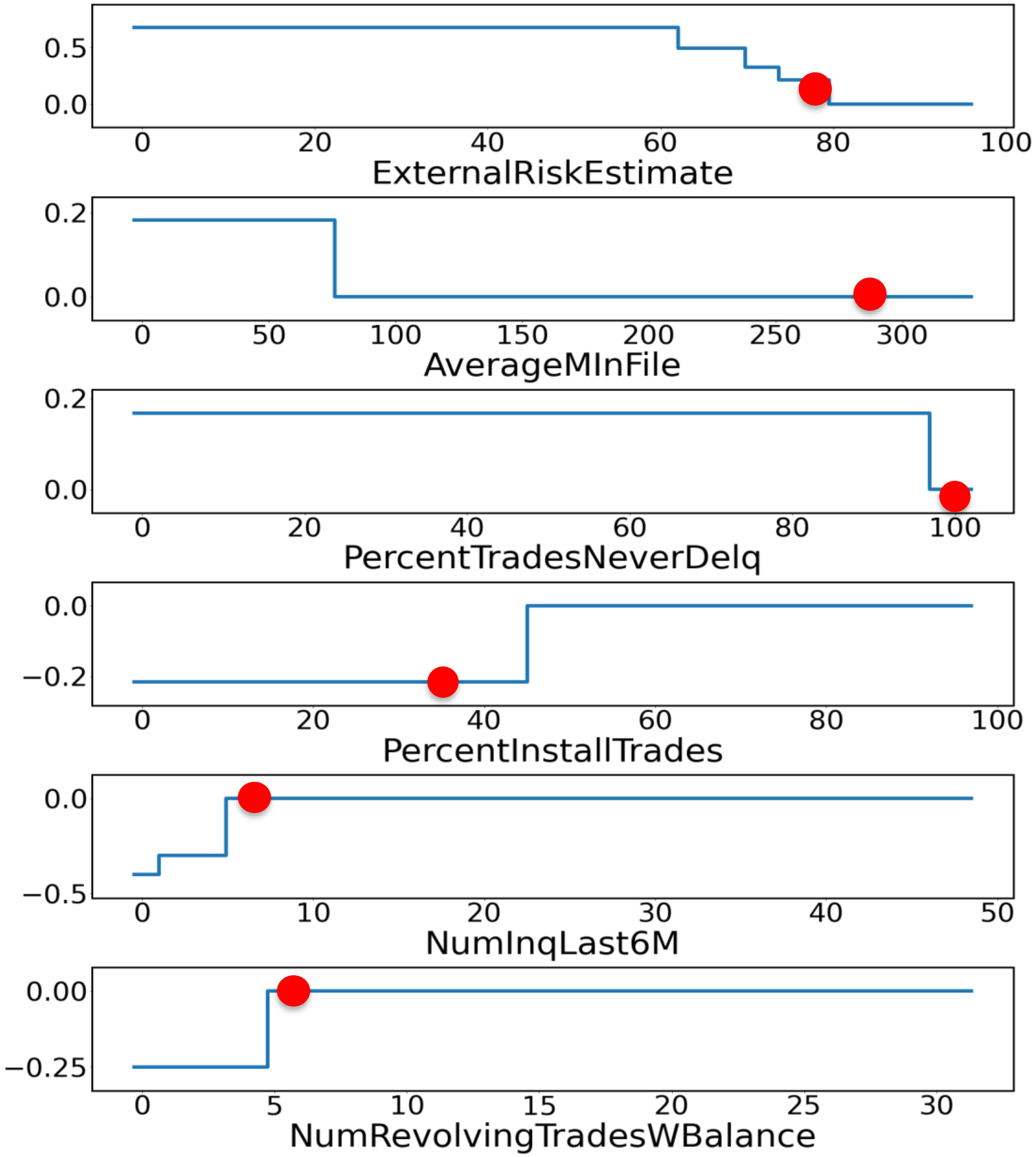
Interpretable

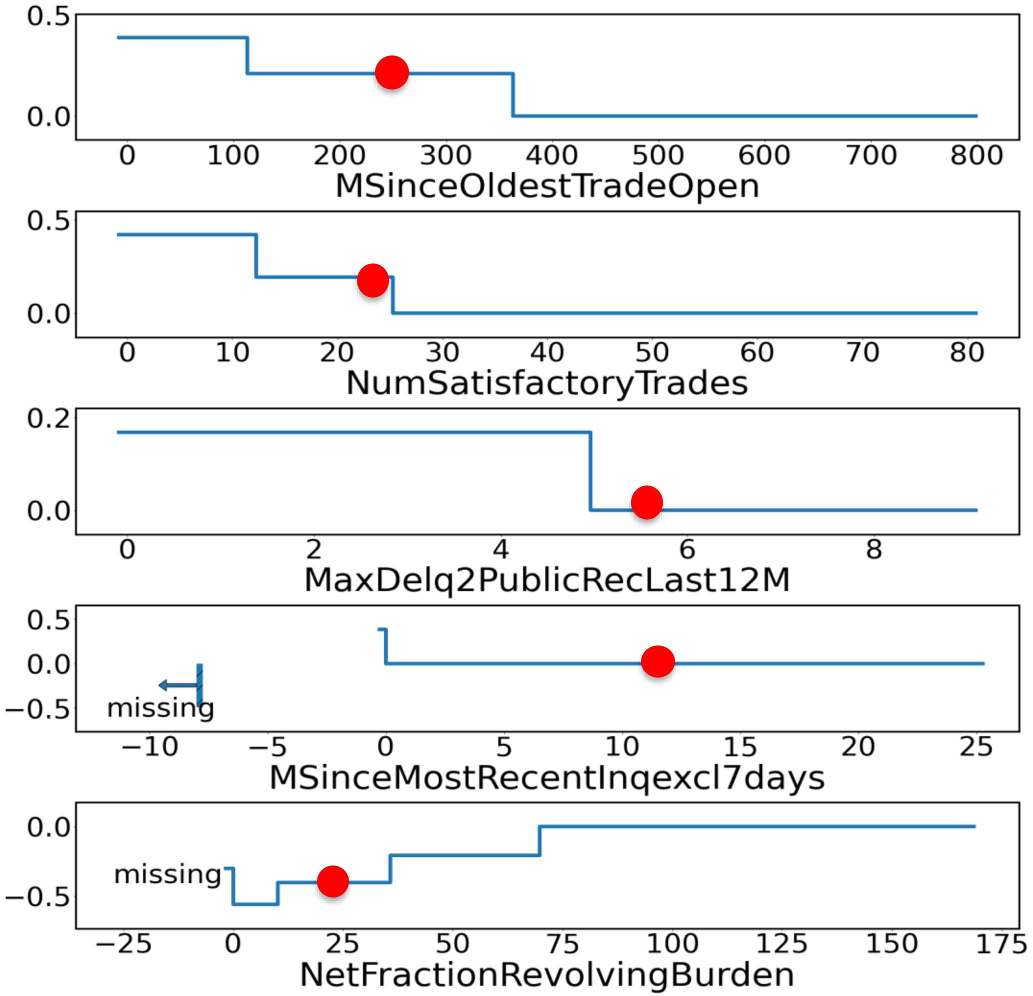




Explainable

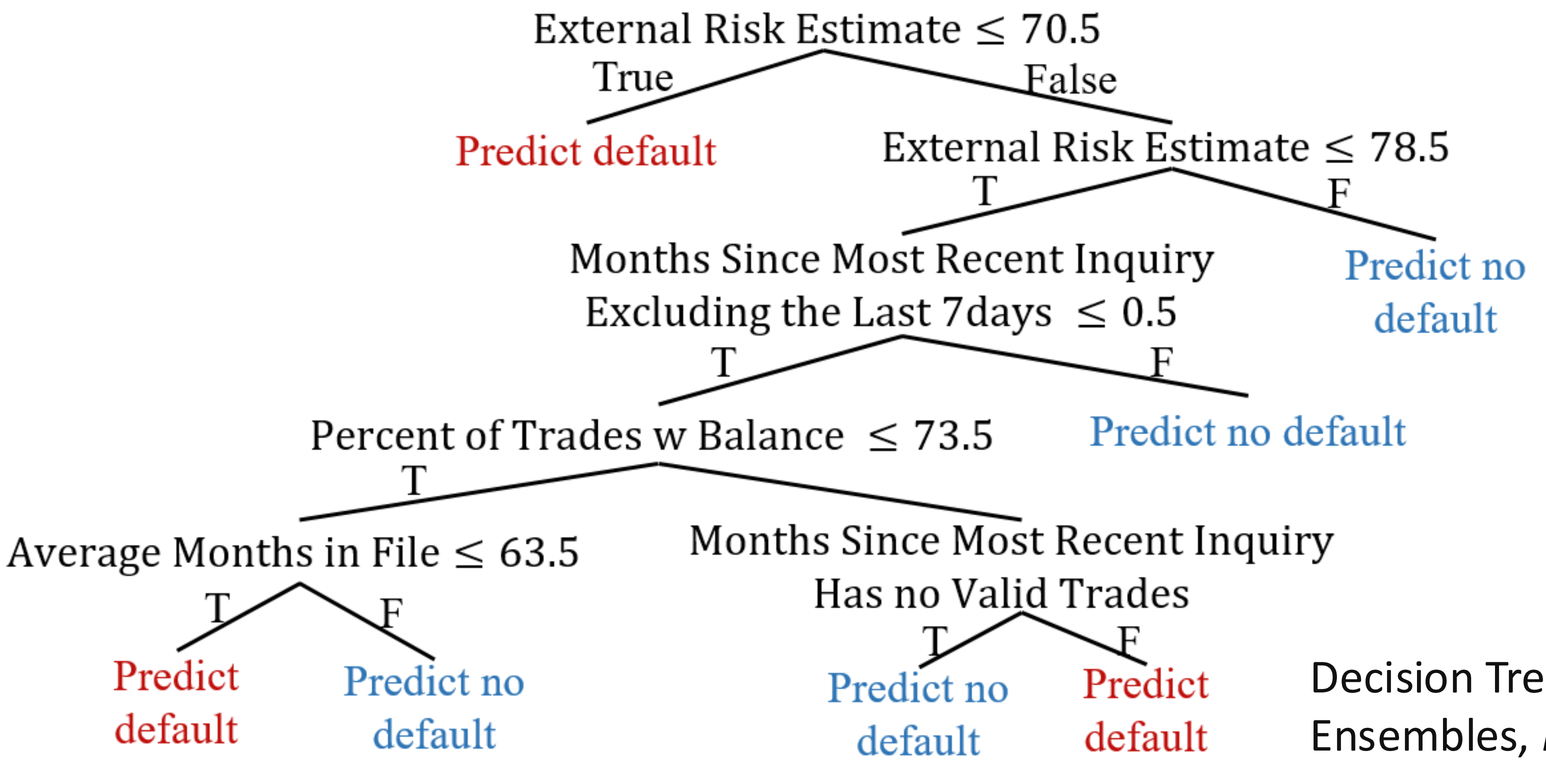
The FICO *Explainable* ML Challenge





Fast Sparse Classification for Generalized Linear and Additive Models, *Liu, J., Zhong, C., Seltzer, M., Rudin, C.,* AISTATS 2020

The FICO *Explainable* AI Challenge



Decision Tree Optimization via Reference Ensembles, *McTavish, H., Zhong, C., Achermann, R., Karimalis, I., Chen, J., Rudin, C., Seltzer, M. AAAI-2022.*

Why Does this Work?

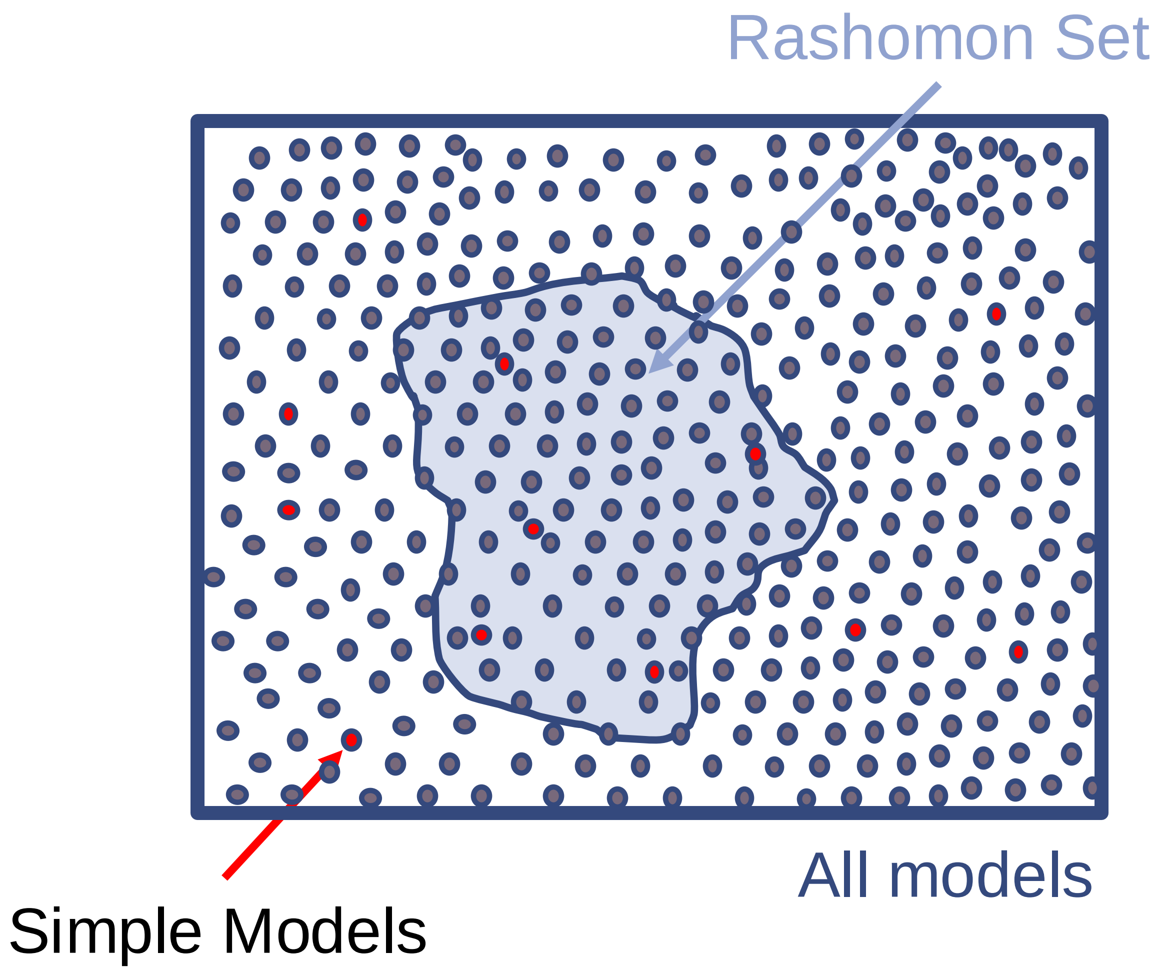
1. The Rashomon Effect (Breiman, 2001)

   "… there is often a multitude of different descriptions [equations f(x)] in a class of functions giving about the same minimum error rate."

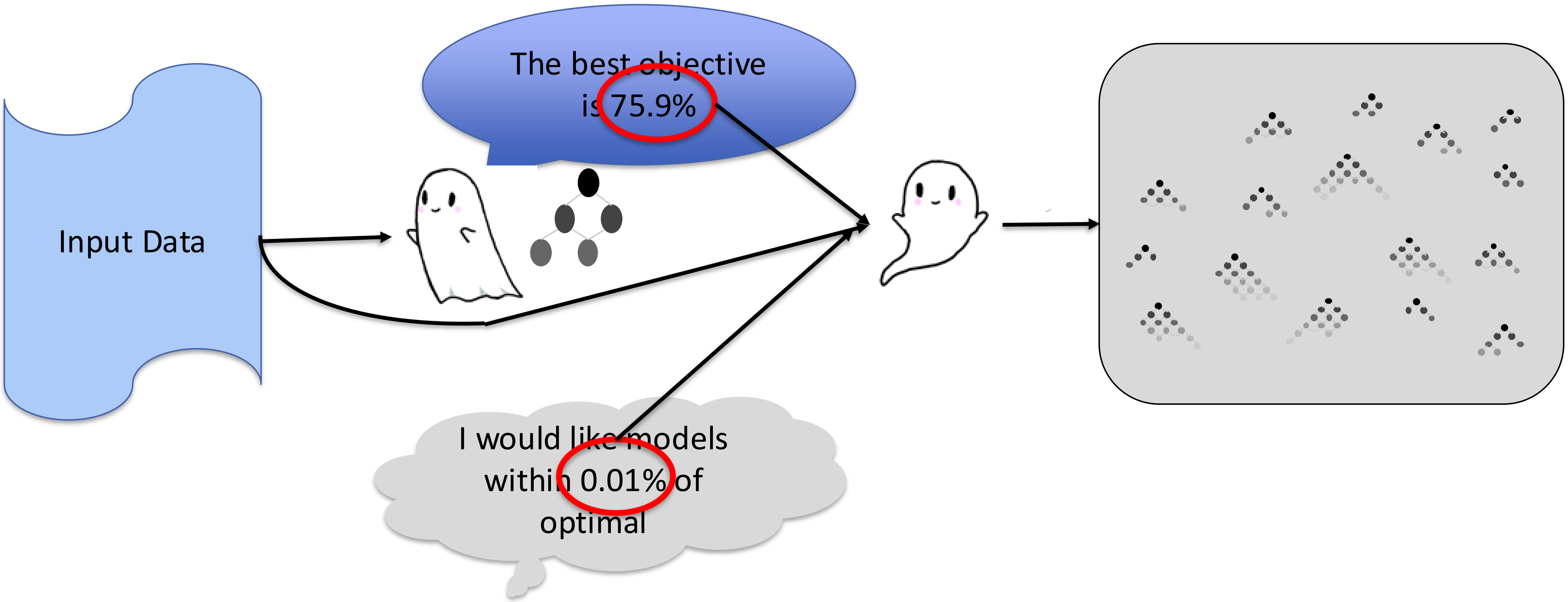
2. If many models produce similar accuracy, the Rashomon set is likely to be large.
3. Large Rashomon ratios make it likely that a simple model exists.

On the existence of Simpler Machine Learning Models, *Semenova, L., Rudin, C., Parr, R.*, FAccT-2022
12-minute discussion: https://www.youtube.com/watch?v=VFAKfIVrnWY
hour-long lecture: https://www.youtube.com/watch?v=xZSRN_kSJUs

Visually: Rashomon Theory Hypothesis

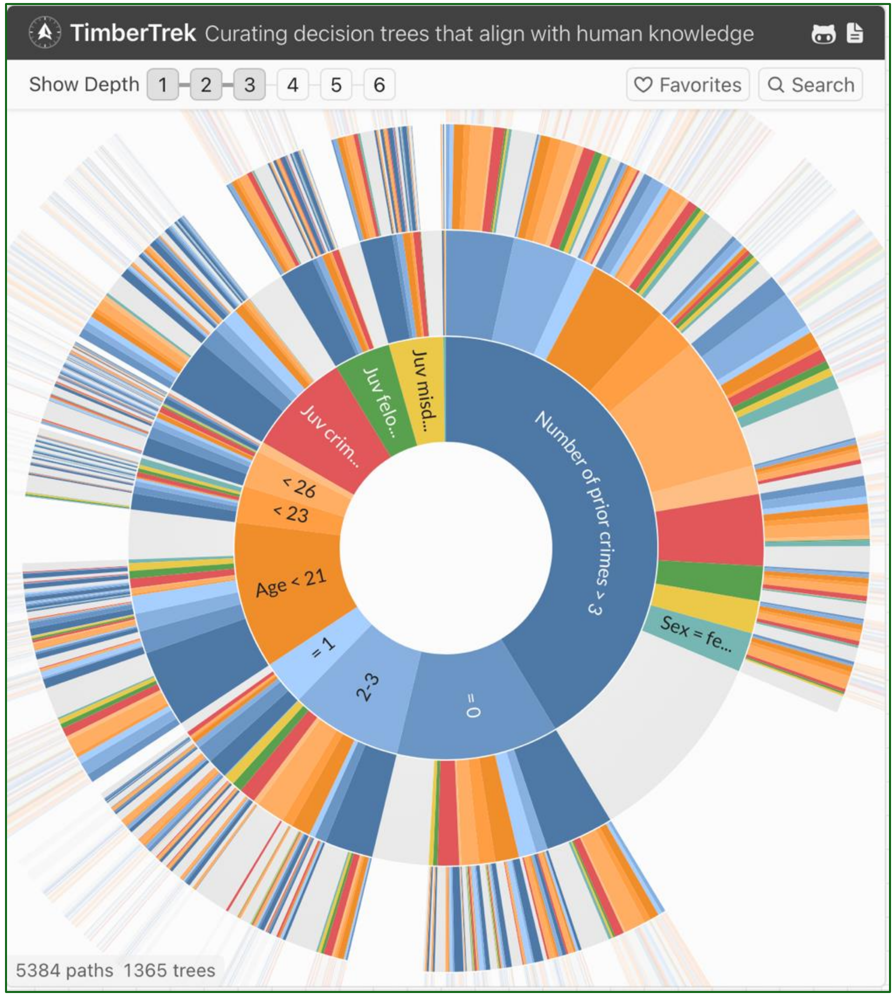


Finding the Rashomon Set of Decision Trees



Finding the Whole Rashomon Set of Spare Decision Trees,
Xin, R., Zhong, C., Li, B., Seltzer, M., Rudin, C., NeurIPS-2022.

Visualizing the Rashomon Set: TimberTrek

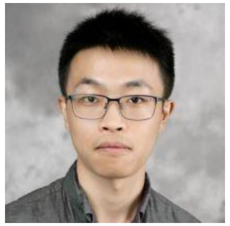


Take aways

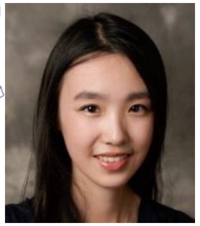
- Many prediction and classification problems can be solved with simple models
  - The modern tools to build these models are fundamentally more powerful than the most commonly used heuristic algorithms.
  - The first model you produce is rarely the one you want; demand many or **all** the good models, so you can pick one that makes sense for your problem.

Thank You!

NSERC
CRSNG



Xiyang Hu
CMU



Chudi Zhong
Duke

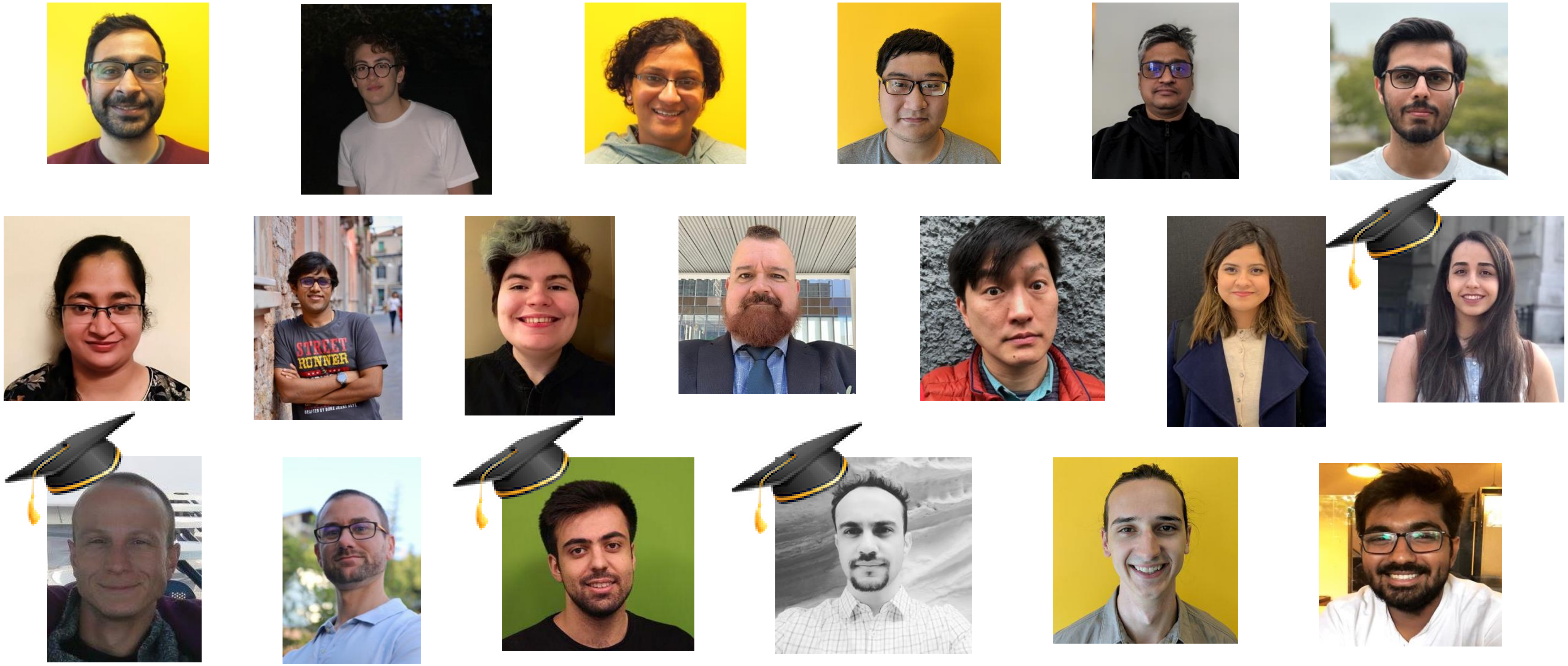
Jimmy Lin
UBC/UofT

Hayden McTavish
UBC/Duke

Cynthia Rudin
Duke

Rui Xin
Duke/UW

My Team



... and many, many undergraduates!