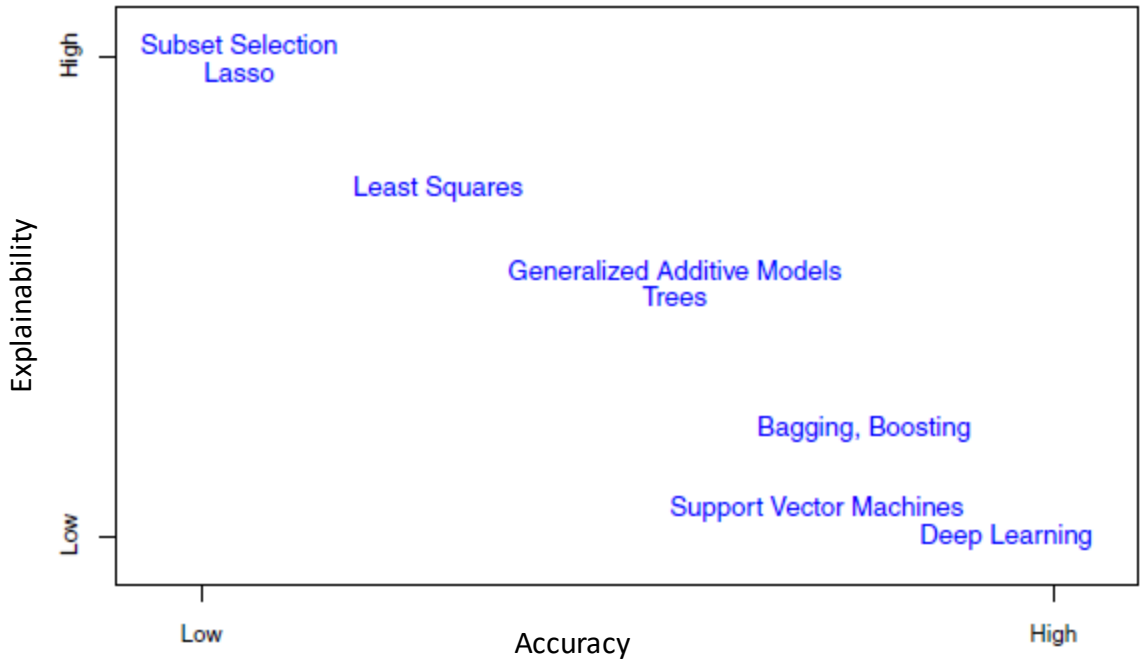
Interpretable Models for Predicting Heart Attack Incidence Using Demographic Data

7th Annual Research Roundtable - Data Analytics in Healthcare
May 8, 2024 - Rotman School of Management

Amir Rastpour*
Reidar Hagtvedt**
Armann Ingolfsson**
Padma Kaul***

*Faculty of Business and Information Technology, Ontario Tech University
**Alberta School of Business, University of Alberta
***Faculty of Medicine & Dentistry, University of Alberta

Explainability vs Accuracy: should we sacrifice one for another?



Underlying premise: The most accurate prediction method is "a 'black box' that is impossible to interpret"

A potential solution: We must develop a separate method to explain how the prediction method works. It is called explainable AI.

James G, Witten D, Hastie T, Tibshirani R (2021) An Introduction to Statistical Learning: With applications in R (Cham, Switzerland: Springer Nature).
NOTE: I have changed the axes labels to be consistent with the terminology in this presentation.

Here is someone who disagrees with the premise of explainable AI

NATURE MACHINE INTELLIGENCE | VOL 1 | MAY 2019 | 206–215

Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead

Cynthia Rudin

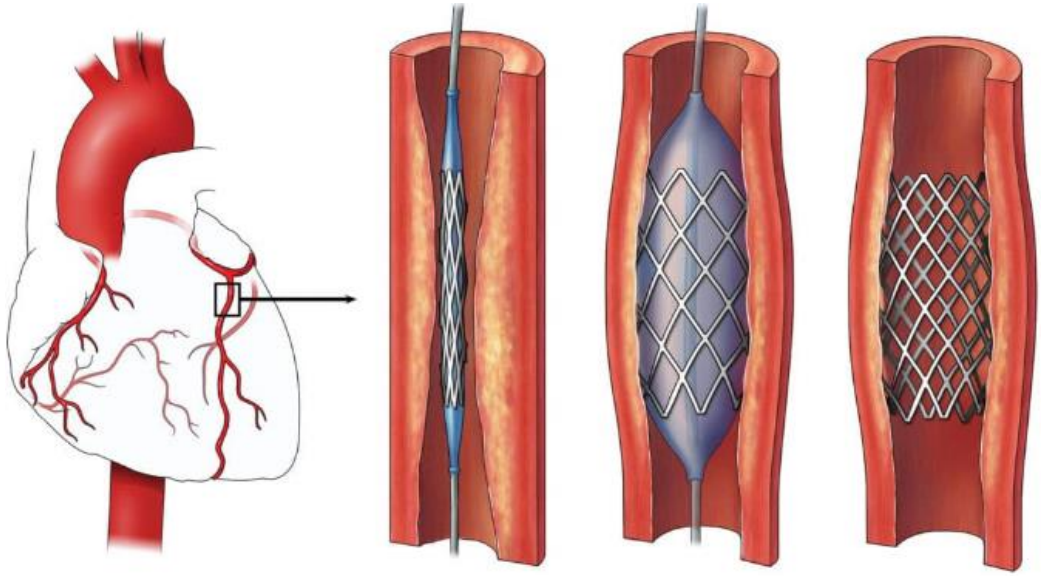
Duke Computer Science

Hypothesis: In many settings, it is possible to develop a simple and interpretable prediction method that is almost as accurate as the most accurate method

Explainable vs Interpretable AI w/ respect to explanation mechanism

- A posteriori: Explainable
  - Sometimes a second model is created to explain the first one
- A priori: Interpretable
  - The model is built so that it satisfies some constraints to help us understand it better
    - Sign restrictions
    - Monotonicity restrictions
    - etc

Demand prediction for a highly-effective heart attack treatment



- Heart attacks occur when a coronary artery is partially or completely blocked by a blood clot.
- Percutaneous Coronary Intervention (PCI): restoration of blood flow by inserting a stent
- PCI is effective if administered within 2 hours of the onset of symptoms

PCI picture from: https://centralgaheart.com/need-know-heart-stent/, Accessed on 2 June. 2022

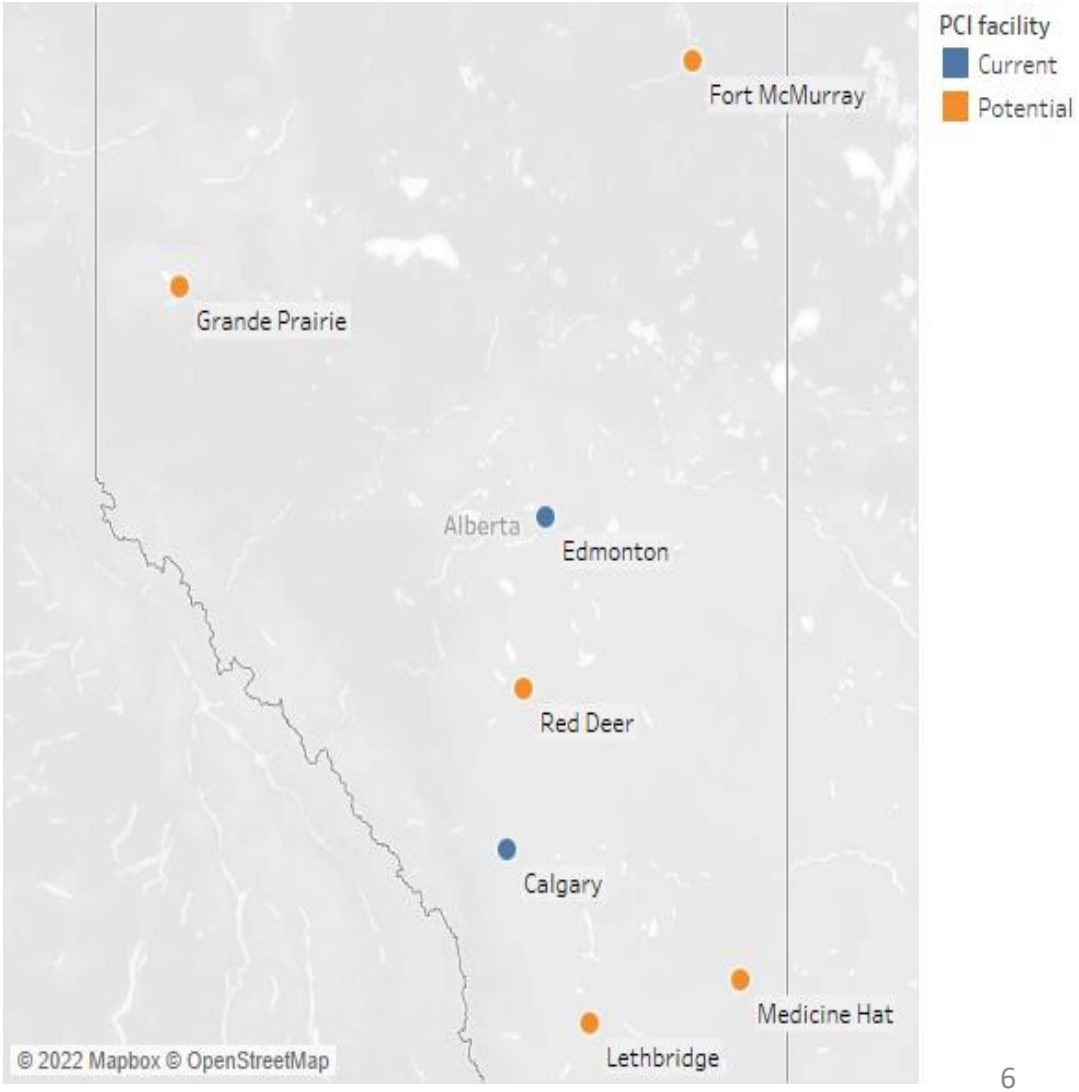
Research question

Motivation: Where should the next PCI facility in Alberta be?

Criterion 1: Maximum coverage of heart attacks, within 120 minutes of travel

Criterion 2: Minimum volume constraints – 400 PCI operations per year per facility

Necessary input: Heart attack incidence predictions



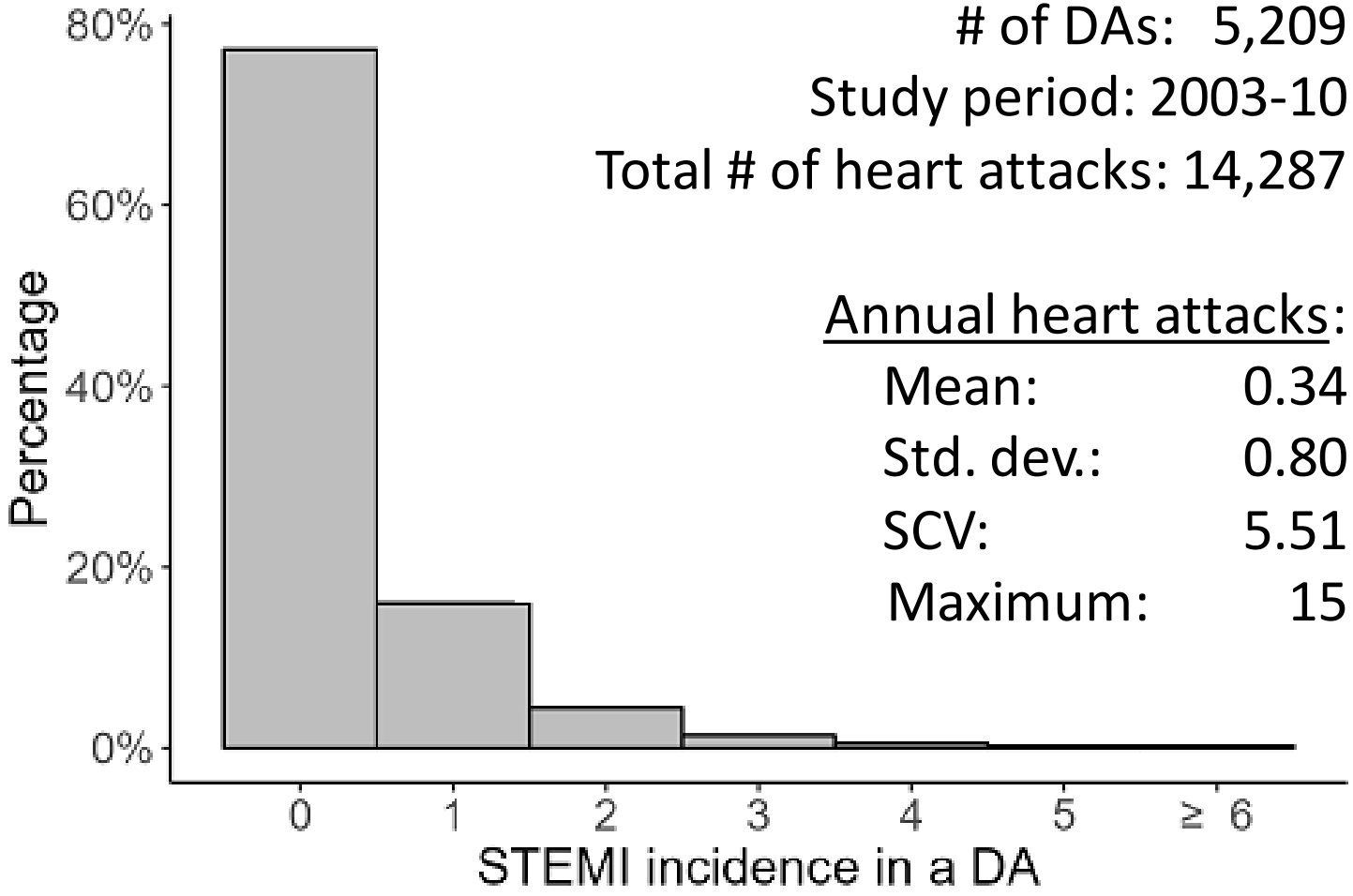
Demand for health care

- Individual level:
  - Models are available in the medical literature
  - Patient's diet
  - Family history
  - etc.
- Jurisdiction level
  - Models from the medical literature may not be applied directly
  - Lack of detailed information about the population

Spatial units

- Postal Codes
  - > 100,000 in Alberta
  - We have 2003-10 heart attack incidence by postal code
- Dissemination Areas (DA)
  - 5,209 in Alberta
  - Smallest standard spatial unit for which all census data are disseminated
  - Population: 400-700, with some exceptions
- We map our postal code data to the DA level

Heart attacks: Descriptive statistics



# of DAs: 5,209
Study period: 2003-10
Total # of heart attacks: 14,287

Annual heart attacks:

Mean:	0.34
Std. dev.:	0.80
SCV:	5.51
Maximum:	15

DA-level Canada census 2006

- 5-year age groups from 0-85 across genders
- Educated people across genders
- Low-income people

Interpretability features

- Functional forr $g(\mathbf{E}[Y]) = \beta_0 + f_1(x_1) + \cdots + f_n(x_n)$
- Per-person rate of hear attack is equal for two different DAs with identical cohort proportions
- DAs with zero population have a prediction of zero heart attack incidence
- All coefficients are non-negative
- Predictions are non-negative.
- Within each gender group (male and female), older individuals have a higher chance of experiencing a heart attack.

Mathematical representation

- If every coefficient is interpreted as a 'rate per person per year,' then the following nonlinear program meets the requirements for interpretability

The linear function without intercept

Sum of squared errors (SSE)

$$\min_{\beta_1,\ldots,\beta_p} \quad \left(\sum_{i=1}^{m}\sum_{j=1}^{n}\beta_j x_{ij} - y_i\right)^2$$

$$\text{s.t.} \quad \beta_j \geq 0, \qquad \forall j \in \{1,\ldots,n\}$$

$$\beta_k \geq \beta_\ell, \qquad \text{if } k \text{ and } \ell \text{ are associated with two different age cohorts of the same gender, and the age cohort associated with } k \text{ is older than the age group associated with } \ell.$$

Model comparison

- Neural Network
- Single point
- Inclusion and exclusion of the interpretability features
- Poisson Regression or zero-inflated Poisson Regression models with similar constraints to our model
- In total, we compare $(2^4 \times 3) + 2 = 50$ different models using Root Mean Square Error from the test data

Results

Model	RMSE
Single point	0.8030
Neural Network	0.7830

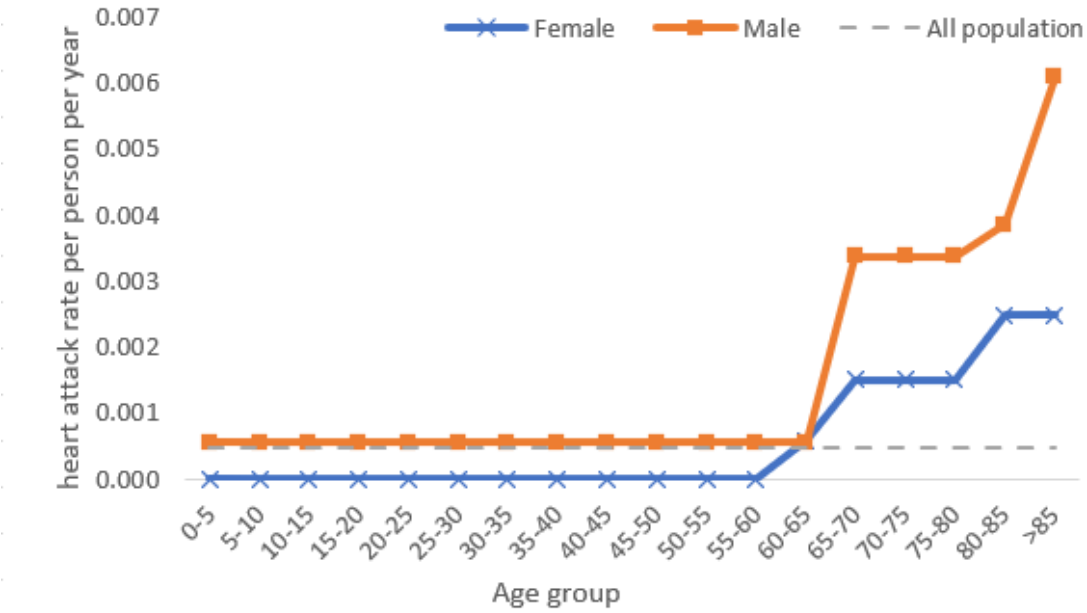
Some of the interpretable models:

Coeffs >= 0	Age coeffs ordered	Without intercept	Without Year variables	Fitting Method	RMSE
Yes	Yes	Yes	No	SSE	0.7659
Yes	Yes	Yes	Yes	SSE	0.7669

Coefficients for the most interpretable model

Ave. # of heart attack= $\sum_{i=1}^{\#\ of\ variables} coefficient_i * variable_i$

Variable	Coefficient
Male Educated	2.08E-09
Female Educated	1.89E-09
Low Income Before Tax	3.74E-08



Coefficients are consistent with medical findings that heart attacks are twice as common in men than women throughout life (Harvard Health Publishing)

Our findings so far and our next steps

- We built interpretable models to predict the number of heart attacks, which are
  - Simple
  - Easy to understand
  - Easy to troubleshoot
- Our interpretable models outperform black box models
- Next step:
  - Use simulation and calculate the probability that each of the candidate locations in Alberta satisfy the medical requirements to open a PCI center