New frontiers in machine learning XAI:
Climbing the “Discovery” Ladder

Mihaela van der Schaar

John Humphrey Plummer Professor of Machine Learning, Artificial Intelligence and Medicine, University of Cambridge //
Director, Cambridge Center for AI in Medicine // Turing Faculty Fellow, The Alan Turing Institute

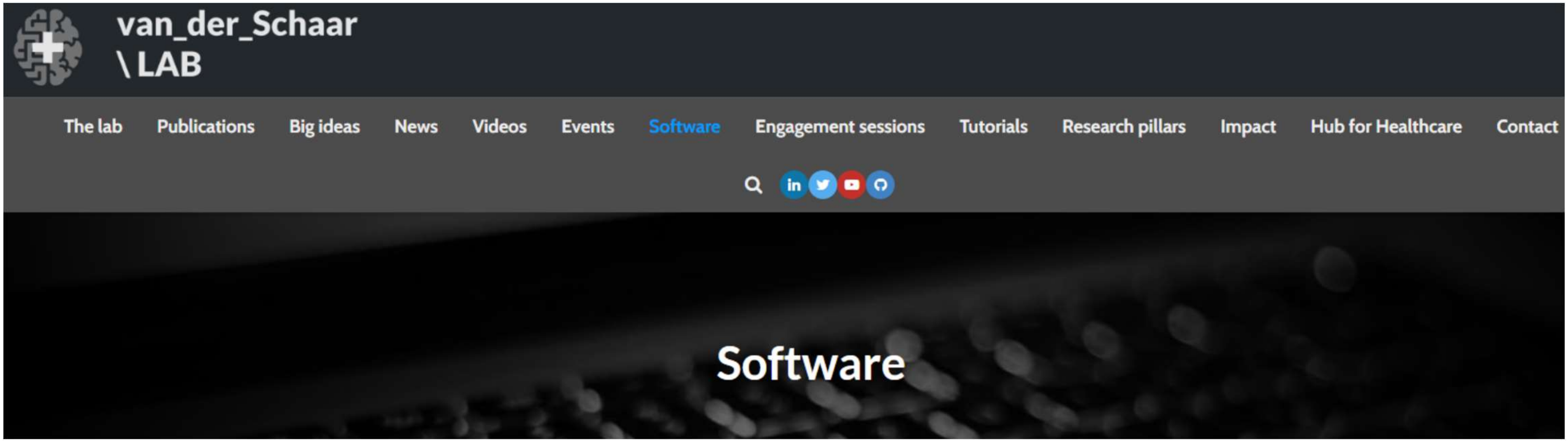
van_der_Schaar \ LAB
vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

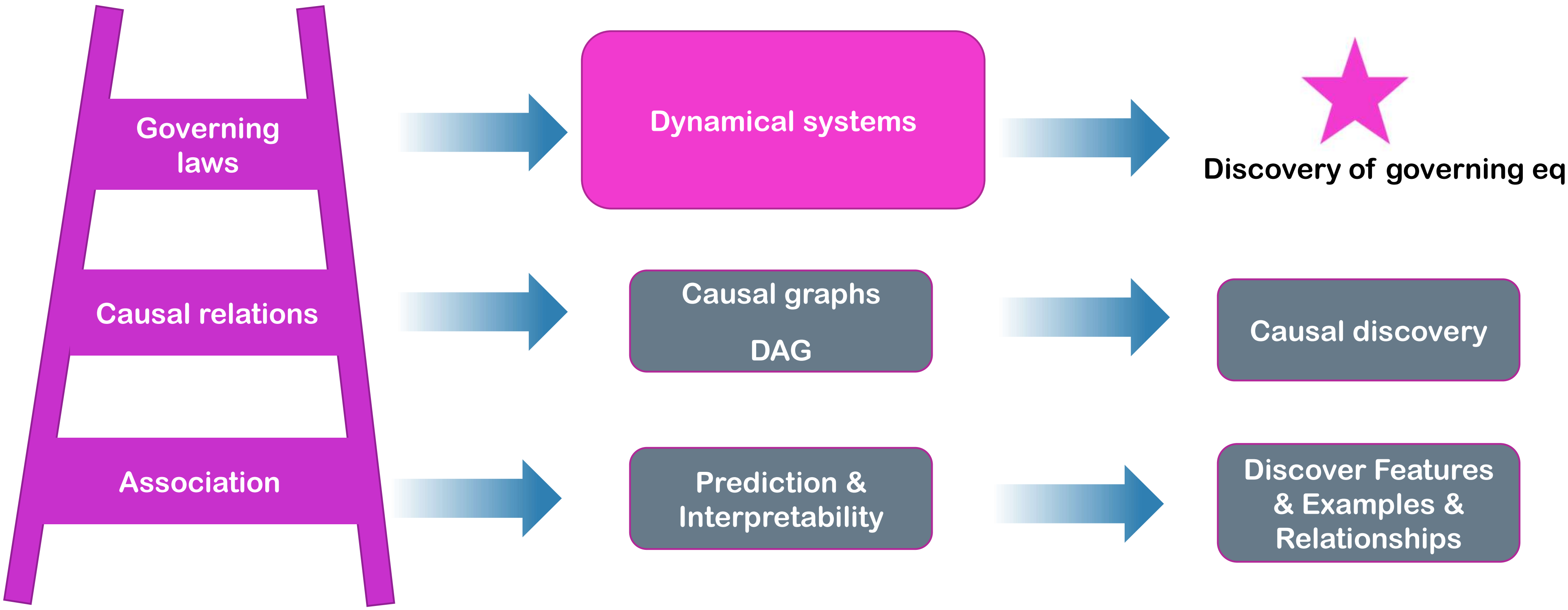
mv472@cam.ac.uk

@MihaelaVDS

linkedin.com/in/
mihaela-van-der-schaar/


van_der_Schaar
\ LAB
The lab
Publications
Big ideas
News
Videos
Events
Software
Engagement sessions
Tutorials
Research pillars
Impact
Hub for Healthcare
Contact
Software

The "Discovery" Ladder



Discover the governing models

- Discover powerful models!

- Why?

  Models are needed to

  ✓ *understand* variables, relationships, components

  ✓ *experiment*

  ✓ *act*

Discovery of governing models

Our focus: governing equations – compact and closed-form equations

The function $f: \mathbb{R}^J \rightarrow \mathbb{R}$ has a closed-form if it can be expressed as
- a finite sequence of operations $(+, \div, log, \dots)$,
- input variables $(x_1,\ x_2, \dots)$ and
- numeric constants $(1.5, 0.8, \dots)$

Discovery of governing models

Our focus: governing equations – compact and closed-form equations

Benefits:

Concise
Generalizable
Amenable to further analysis (e.g., identifying stable equilibria)
Transparent
Interpretable to human experts

Discovery of governing equations

	Explicit function	Implicit function	Ordinary differential equation	Partial differential equation
Typical form	$y = f(x)$	$f(x, y) = c$	$\frac{dx}{dt} = f(x, t)$	$\frac{\partial u}{\partial t} = f(u, x)$

Governing equations in medicine

- Risk prediction
  - Models for disease-specific mortality
- Pharmacology and physiology
  - Dosing, cell potential
- Disease progression
  - Tumor growth models and treatment effect
- Epidemiology
  - Modeling spread of infection

NHS Predict Breast Cancer equations

- if ER+

$$H_c(t) = \exp[0.7424402 - 7.527762/\sqrt{t} - 1.812513 * \log(t)/\sqrt{t}]$$

- if ER-

$$H_c(t) = \exp[-1.156036 + 0.4707332/t^2 - 3.51355/t].$$

Governing equations in medicine

- Risk prediction
  - Models for disease-specific mortality
- Pharmacology and physiology
  - Dosing, cell potential
- Disease progression
  - Tumor growth models and treatment effect
- Epidemiology
  - Modeling spread of infection

Hill equation

$$\frac{E}{E_{\max}} = \frac{[A]^n}{\mathrm{EC}_{50}^n + [A]^n}$$

Governing equations in medicine

- Risk prediction
  - Models for disease-specific mortality
- Pharmacology and physiology
  - Dosing, cell potential
- Disease progression
  - Tumor growth models and treatment effect
- Epidemiology
  - Modeling spread of infection

Geng et al., 2017

$$V(t+1) = \Big(1 + \underbrace{\rho \log\big(\frac{K}{V(t)}\big)}_{\text{Tumor growth}} - \underbrace{\beta_c C(t)}_{\text{Chemotherapy}} - \underbrace{\big(\alpha_r d(t) + \beta_r d(t)^2\big)}_{\text{Radiotherapy}} + \underbrace{e_t}_{\text{Noise}}\Big) V(t)$$

Governing equations in medicine

- Risk prediction
  - Models for disease-specific mortality
- Pharmacology and physiology
  - Dosing, cell potential
- Disease progression
  - Tumor growth models and treatment effect
- Epidemiology
  - Modeling spread of infection

etc.

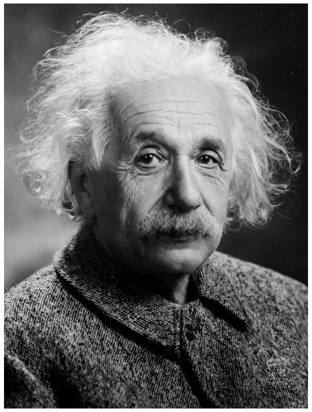
SIR Model

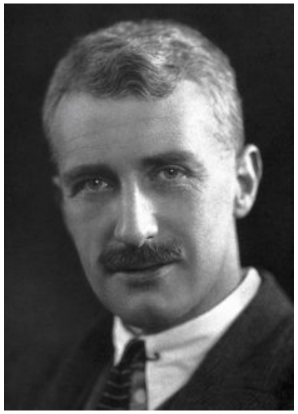
$$\frac{\mathrm{d}S}{\mathrm{d}t} = -r\beta S\frac{I}{N}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = r\beta S\frac{I}{N} - \gamma I$$
$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I$$

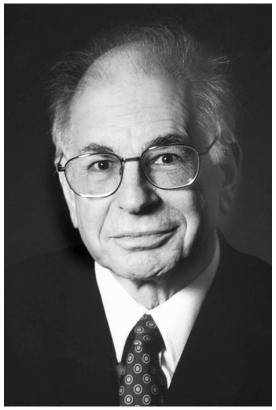
Traditional approach: expert-driven discovery

Human experts (scientists) discover governing equations

- Domain knowledge, intuition, experience
- Requires brilliant insight to tease out equations from the underlying data, which form the new hypothesis







Traditional approach: expert-driven discovery - limitations

- **Built upon current hypotheses (may not be correct)**
- **Data is often selected or even ignored (!) to support the hypothesis**
- **Only few variables considered**
- **Some variables are measurable in the lab (petri-dish) but not in the real-world (patient)**

Medicine is complex

Variabilities between individuals

- Different genetic background
- Different environmental exposures
- Different life-styles
- Different histories and interventions

etc.

Lead to

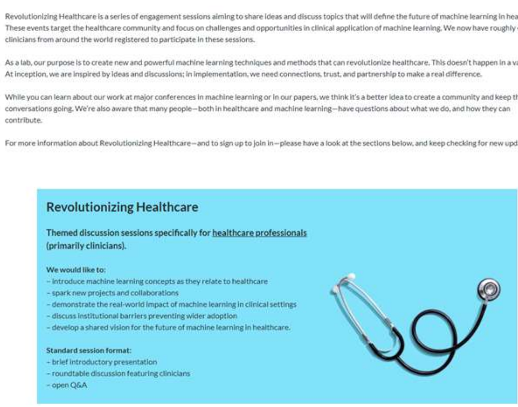
- Different risks (and different risks over time)
- Variation in symptoms
- Different health and disease trajectories
- Different responses to treatment

etc.

In medicine, judgements are made based on this complex information to manage patients

Machine learning can help improve this!

Engagement sessions: Revolutionizing Healthcare


Revolutionizing Healthcare is a series of engagement sessions aiming to share ideas and discuss topics that will define the future of machine learning in healthcare. These events target the healthcare community and focus on challenges and opportunities in clinical application of machine learning. We now have roughly 400 clinicians from around the world registered to participate in these sessions.

As a lab, our purpose is to create new and powerful machine learning techniques and methods that can revolutionize healthcare. This doesn't happen in a vacuum. At inception, we are inspired by ideas and discussions; in implementation, we need connections, trust, and partnership to make a real difference.

While you can learn about our work at major conferences in machine learning or in our papers, we think it's a better idea to create a community and keep these conversations going. We're also aware that many people—both in healthcare and machine learning—have questions about what we do, and how they can contribute.

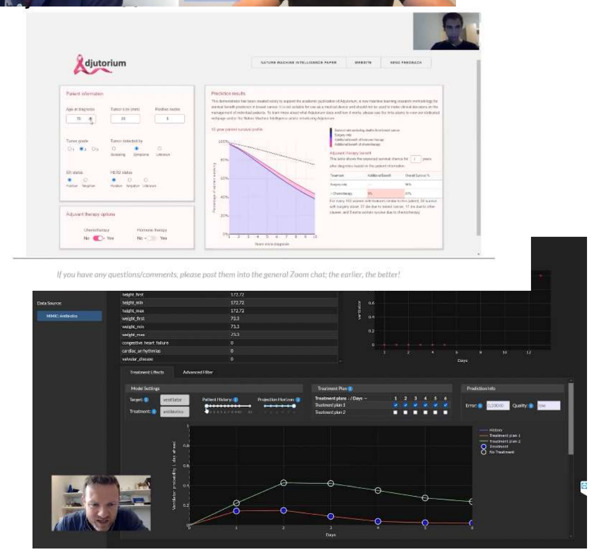
For more information about Revolutionizing Healthcare—and to sign up to join in—please have a look at the sections below, and keep checking for new updates.

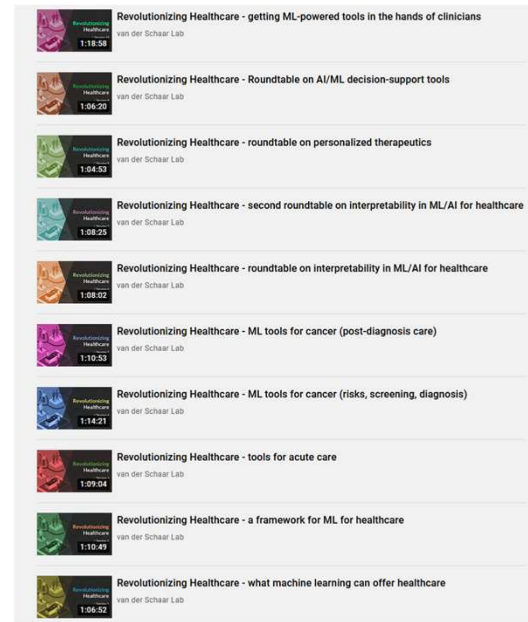
Revolutionizing Healthcare

Themed discussion sessions specifically for healthcare professionals (primarily clinicians).

We would like to:
– introduce machine learning concepts as they relate to healthcare
– spark new projects and collaborations
– demonstrate the real-world impact of machine learning in clinical settings
– discuss institutional barriers preventing wider adoption
– develop a shared vision for the future of machine learning in healthcare.

Standard session format:
– brief introductory presentation
– roundtable discussion featuring clinicians
– open Q&A





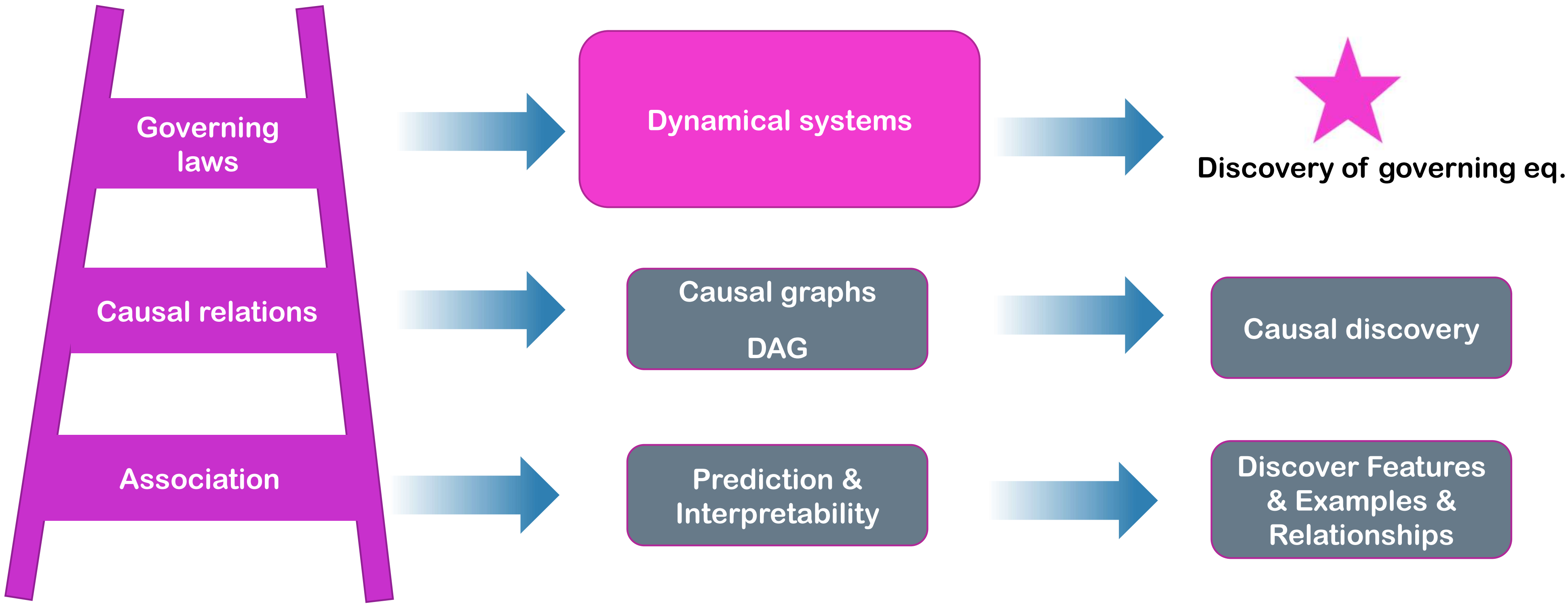
https://www.vanderschaar-lab.com/
→ Engagement sessions
→ Revolutionizing Healthcare

van_der_Schaar \LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

Climbing the "Discovery" Ladder



Climbing the “Discovery” Ladder

1. Explanatory features and examples - Beyond interpreting static predictions
   - Interpreting time-series models
   - Unsupervised models
   - Heterogeneous treatment effect models

2. Personalized explanations - Beyond “one-size-fits-all” example explanations
   - Give explanations based on specific patients selected by each clinician

3. Discover governing equations of medicine - Going beyond association and causal relations

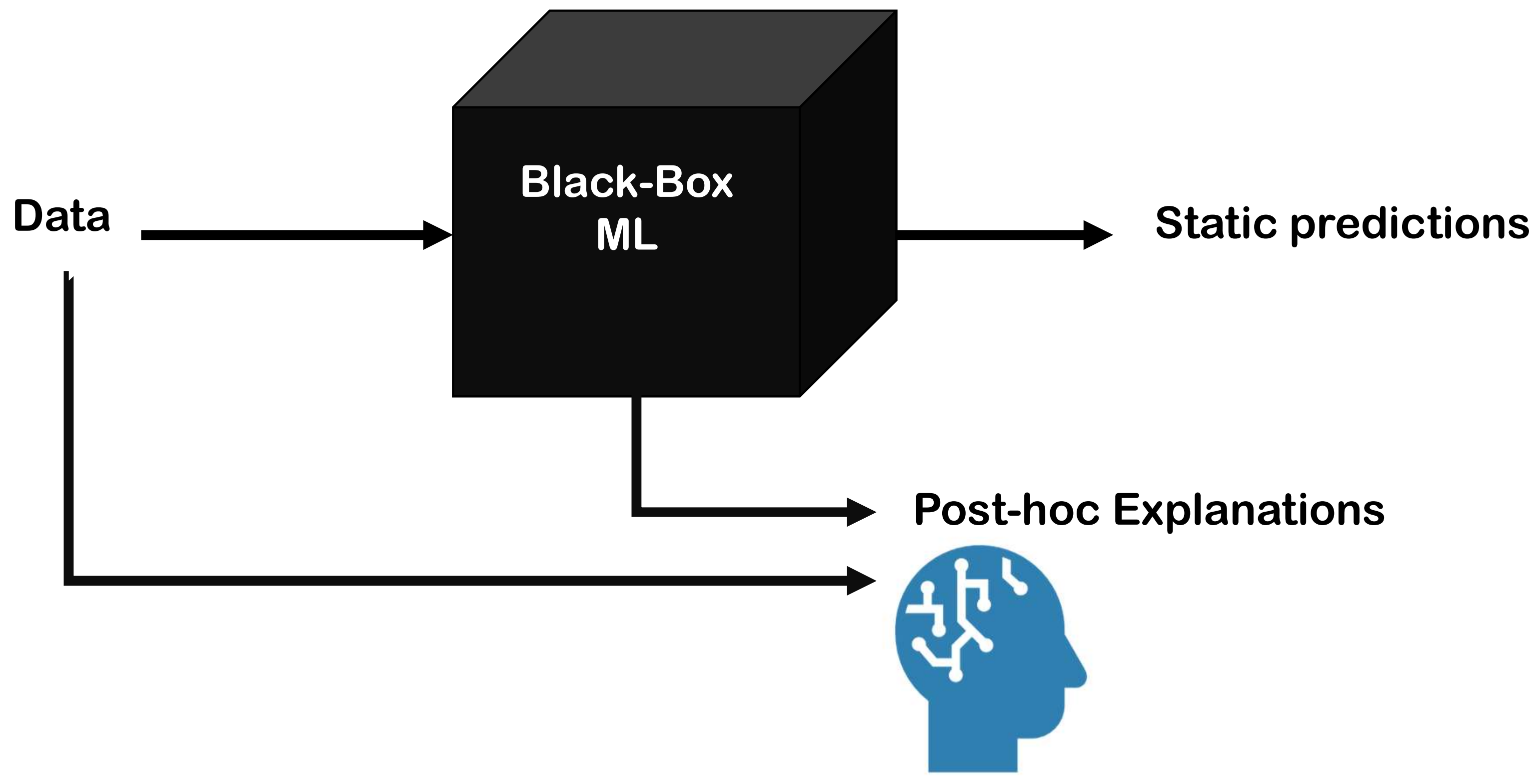
Climbing the "Discovery" Ladder

1. Explanatory features and examples - Beyond interpreting static predictions
   - Interpreting time-series models
   - Unsupervised models
   - Heterogeneous treatment effect models

2. Personalized explanations - Beyond "one-size-fits-all" example explanations
   - Give explanations based on specific patients selected by each clinician

3. Discover governing equations of medicine - Going beyond association and causal relations

Medicine – We need to go beyond interpretability of static predictions



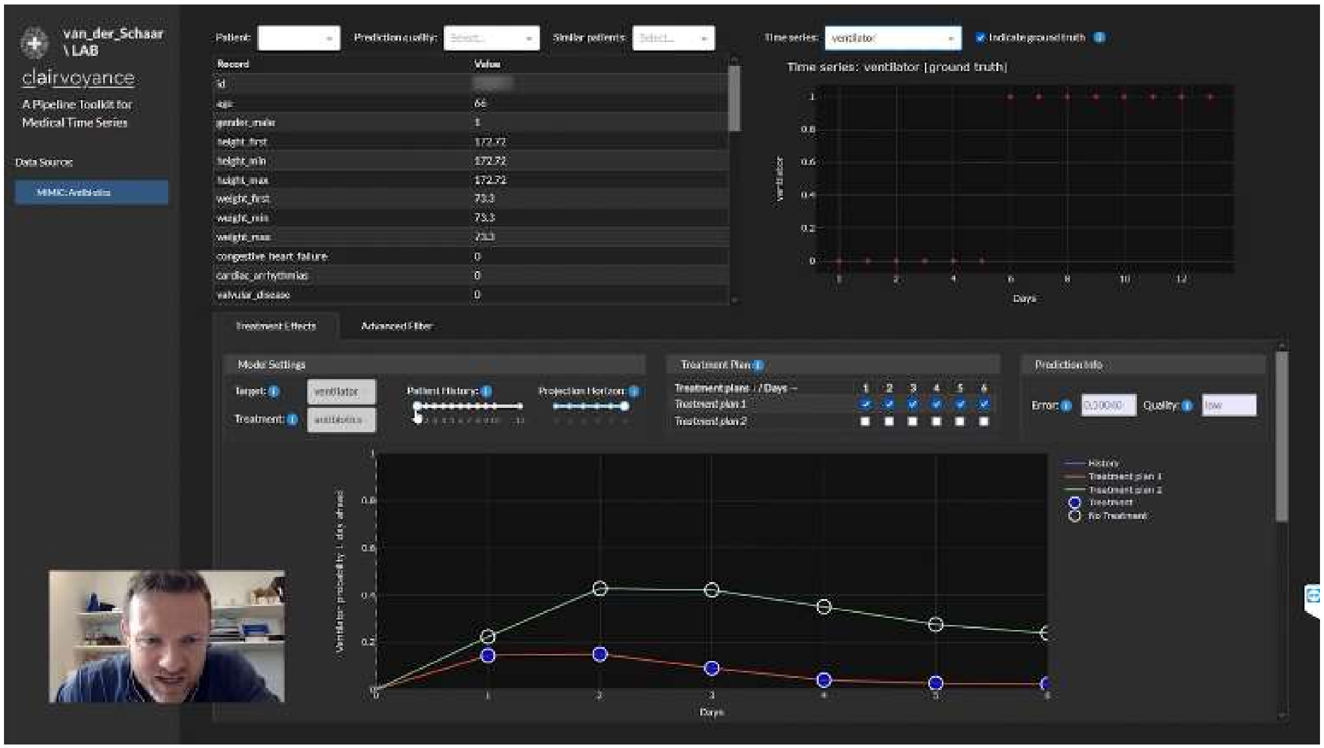
Medicine – We need to go beyond interpretability of static predictions

Patient trajectories
Time-series forecasting - Dynamask [ICML 2021]

Patient phenotyping
Unsupervised learning methods – Label-free explainability [ICML 2022]

Personalized therapeutics for each patient
Heterogeneous effects estimation – ITErpretability [arxiv]

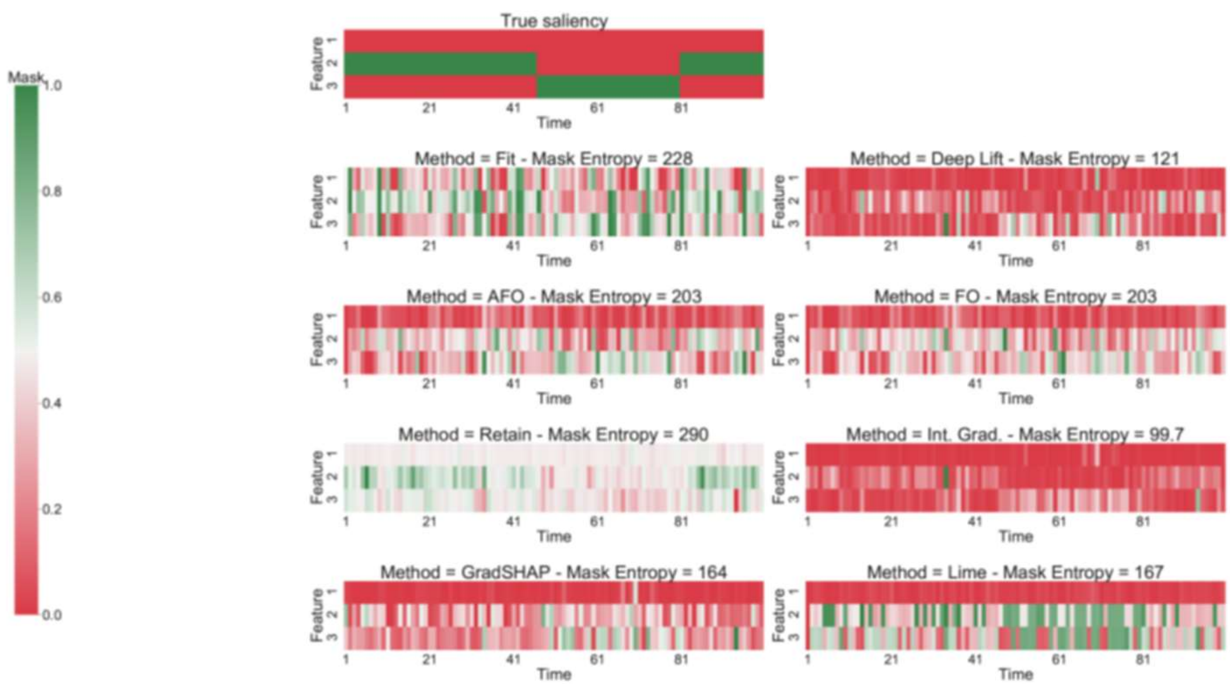
All problems in medicine should be considered time-series problems



"All problems in medicine should be considered time-series problems – all our patients have histories, all diseases have natural histories. Clinical experience tries to capture these trajectories but almost no formal methods have so far been able to do so."
Dr. Eoin McKinney

Time-series forecasting –
Do standard interpretability methods work?

NO! [Ismail et al., NeurIPS 2020]



How to take the time context into account? [Crabbé, vdS, ICML 2021]

Challenge: Time context matters!

Standard methods treat each input $x_{t,i}$ as a feature

$\Longrightarrow$ Time dependency is ignored

Dynamic Perturbation Operator

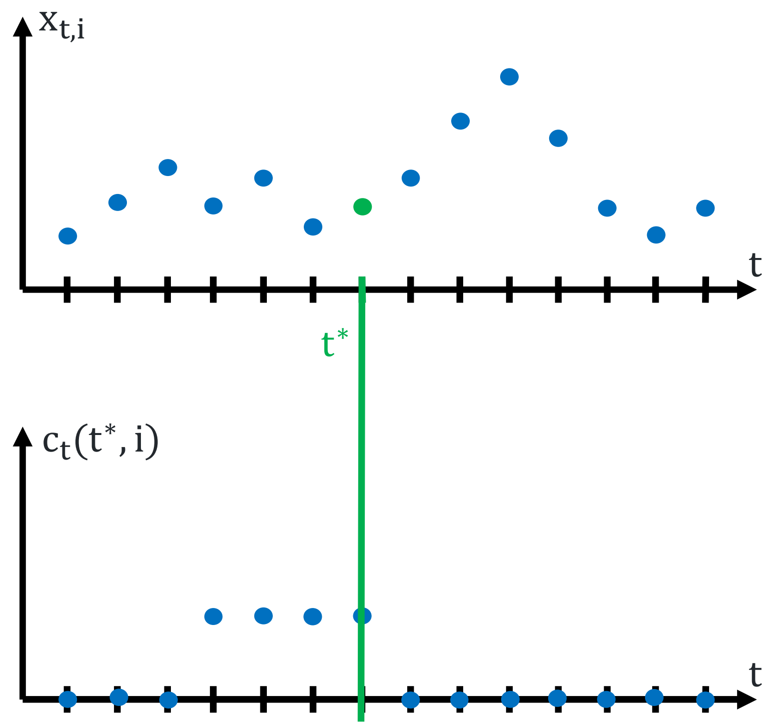
Idea: perturb each $x_{t^*,i}$ by using neighbouring times:

Perturbed input | Linear combination

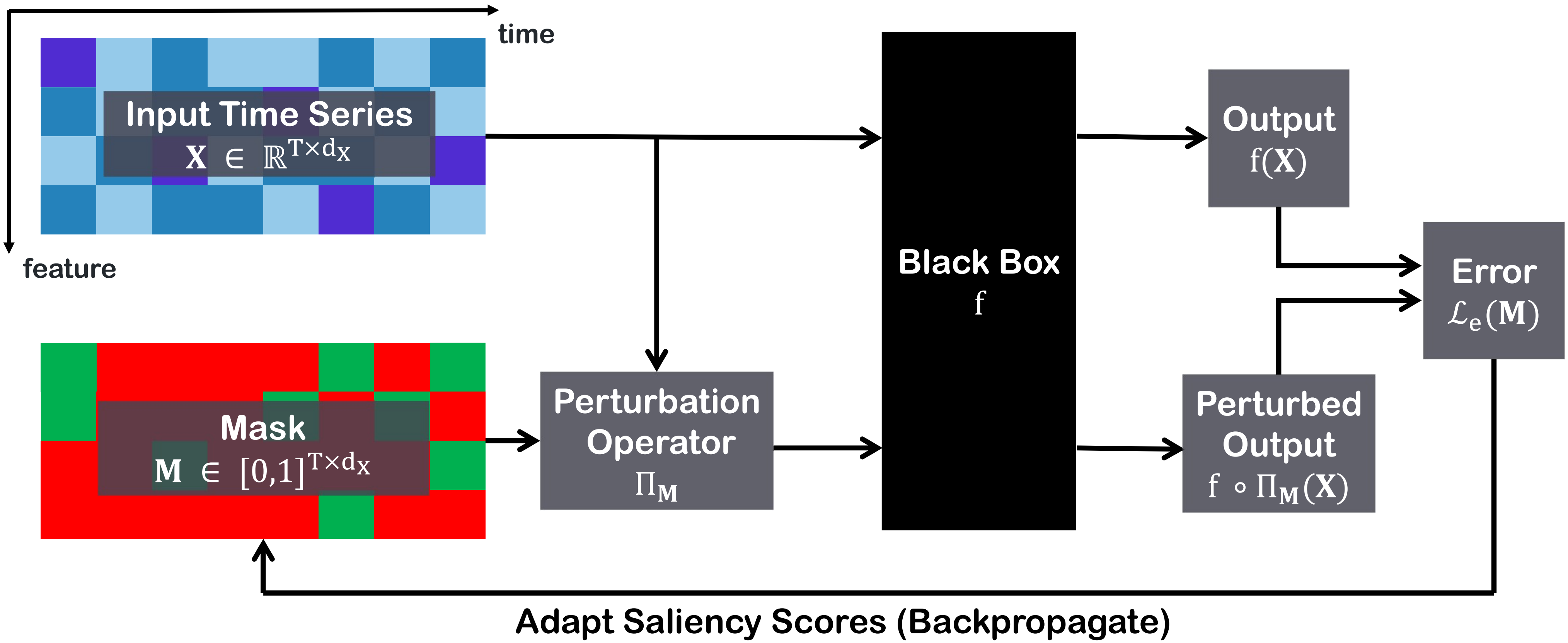
$$\pi(x_{t^*,i}\,; t^*, i) = \sum_{t=t^*-W_1}^{t^*+W_2} c_t(t^*, i) \times x_{t,i}$$

$\Longrightarrow$ Time dependency is integrated in perturbation

Past window perturbation:



Dynamask [Crabbé, vdS, ICML 2021]



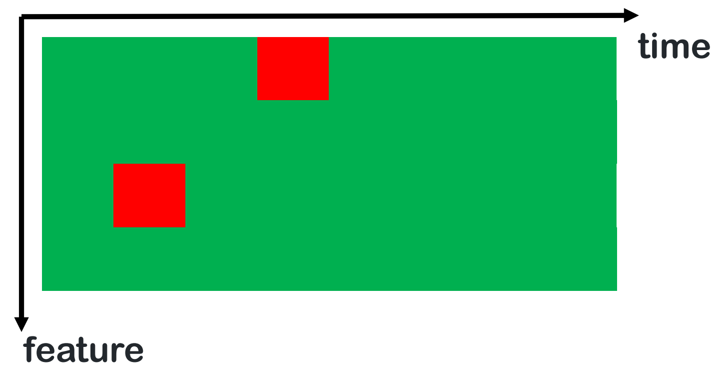
We need “parsimonious” explanations

What do we mean by parsimonious?

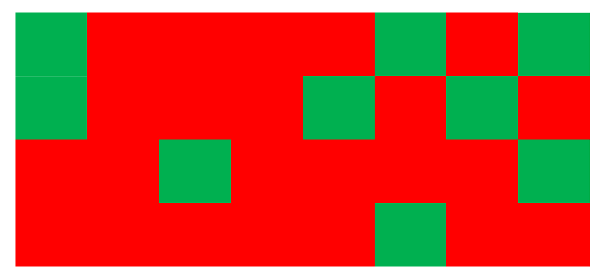
Masks should **not** highlight more features than necessary

$\Longrightarrow$ **Feature selection**

Useless:



Useful:



How to enable parsimony?

User selects desired fraction a of most important features

Dynamask adds a regularization to enforce sparsity:

$$\mathcal{L}_a(\mathbf{M}) = \|\text{vecsort}(\mathbf{M}) - \mathbf{r}_a\|^2$$

We need “congruous” explanations

What do we mean by congruous?

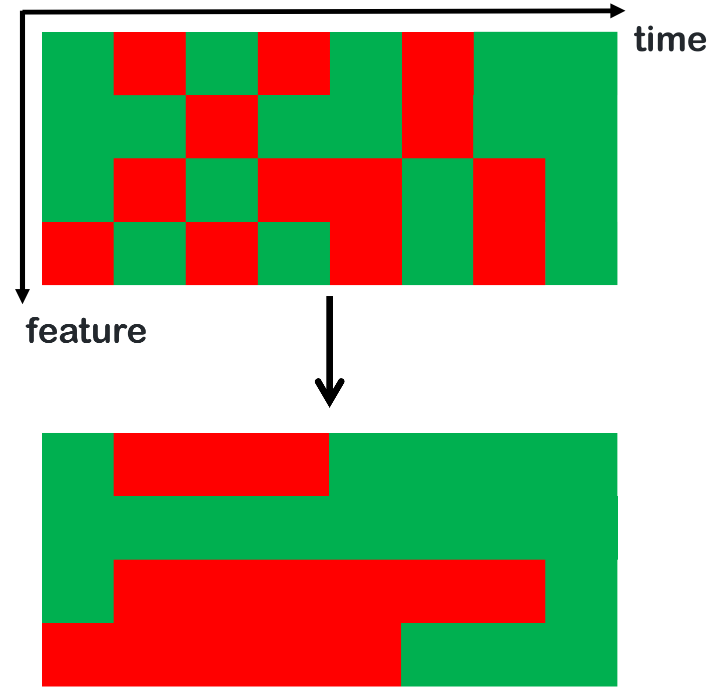
Masks should avoid quick time variations of the saliency

(Robustness)

How to enable congruity?

Dynamask adds a regularization to penalize saliency jumps over time:

$$\mathcal{L}_c(\mathbf{M}) = \sum_{t=1}^{T-1} \sum_{i=1}^{d_X} \left| m_{t+1,i} - m_{t,i} \right|$$

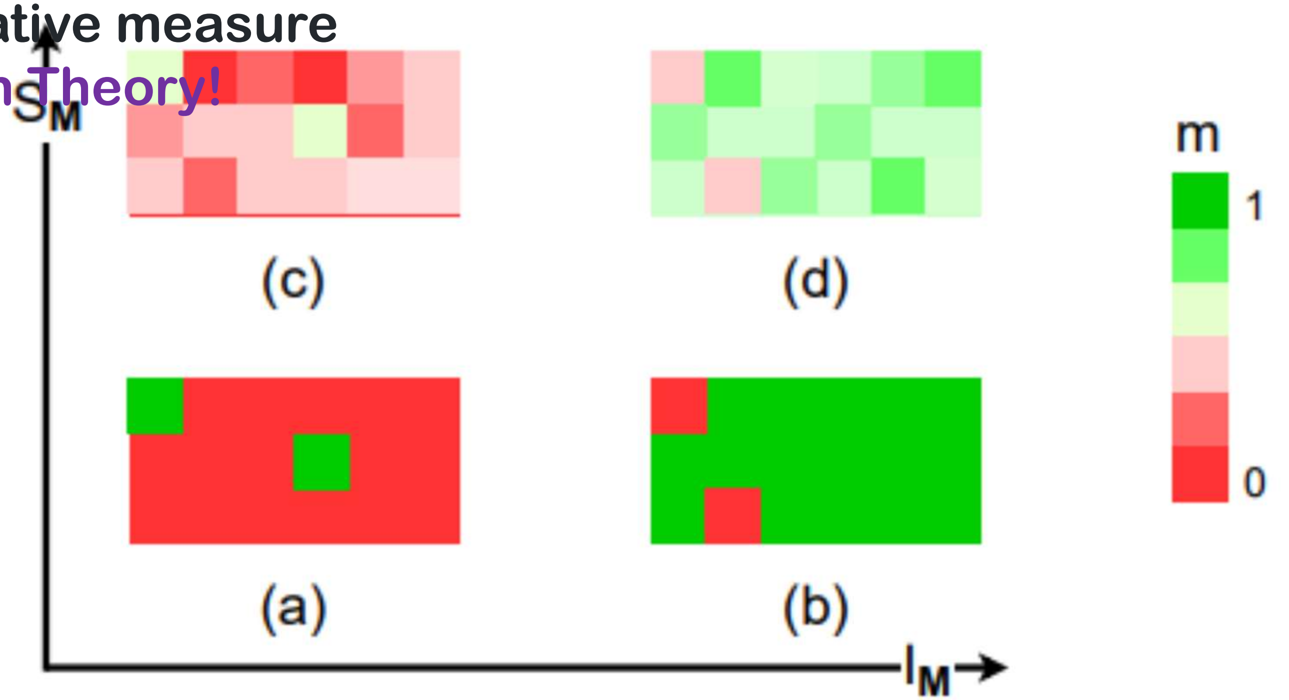


Dynamask enables the saliency map to be “legible”

[Crabbé, vdS, ICML 2021]

How to we know if the “legibility” is achieved by an interpretability method?
We need a quantitative measure
We use Information Theory!

Introduced
Mask information
&
Mask entropy

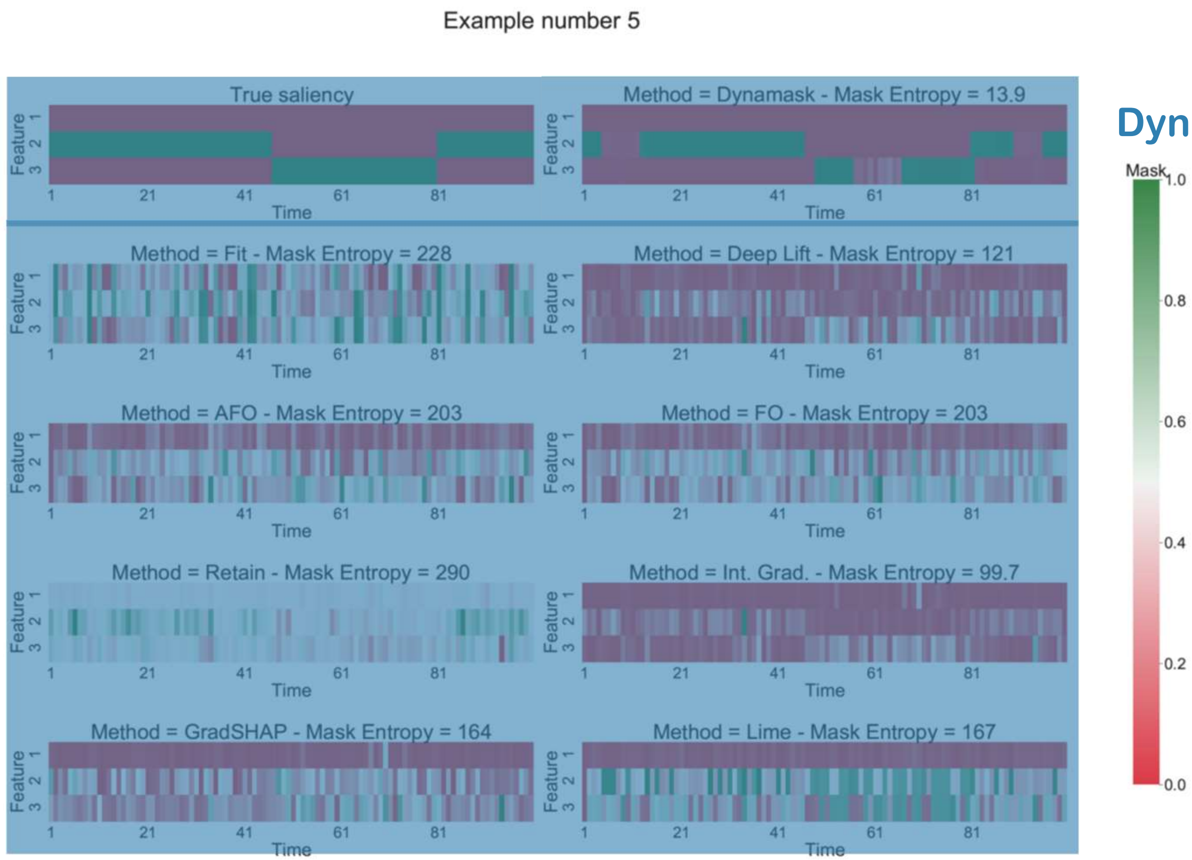


Dynamask - Example

[Crabbé, vdS, ICML 2021]

True saliency

Baseline saliency



Dynamask saliency

Explaining *Unsupervised* Models

[Crabbé, vdS, ICML 2022]

- **Unsupervised learning:** e.g. clustering/phenotyping
- **Self-sup** … urIPS 2020]

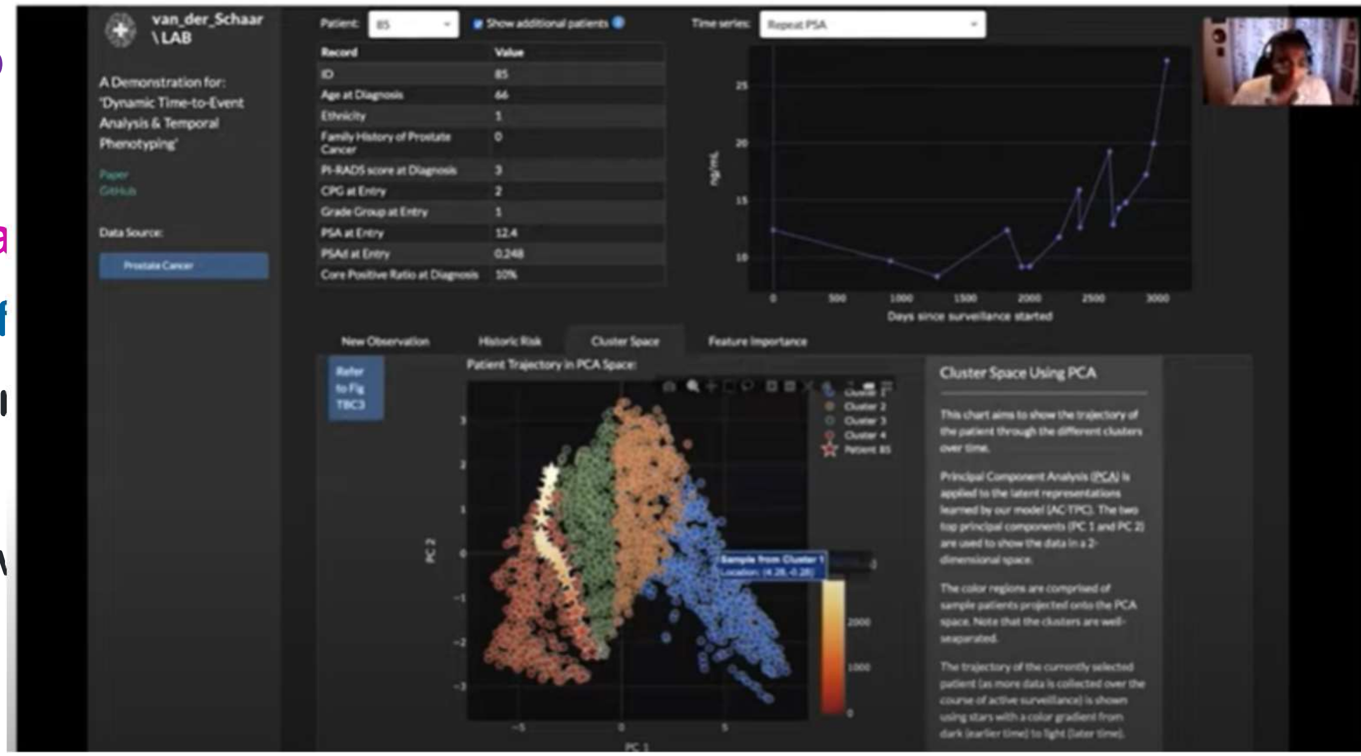
Desiderata

✓ Both f …

✓ Unde … IL encoders

✓ Work … e Functions)

✓ Work … coder, SimCLR)



Medicine – We need to go beyond interpretability of static predictions

Time-series forecasting - Dynamask [ICML 2021]

Unsupervised learning methods – Label-free explainability [ICML 2022]

Heterogeneous effects estimation – ITErpretability [arxiv]

What can be improved?

- **Debug the models**

  Previous methods don't explain why some feature/examples are important

  Previous methods don't clearly indicate when to be sceptical about the model

- **Keep humans in the loop**

  Previous methods don't adapt to the user's knowledge

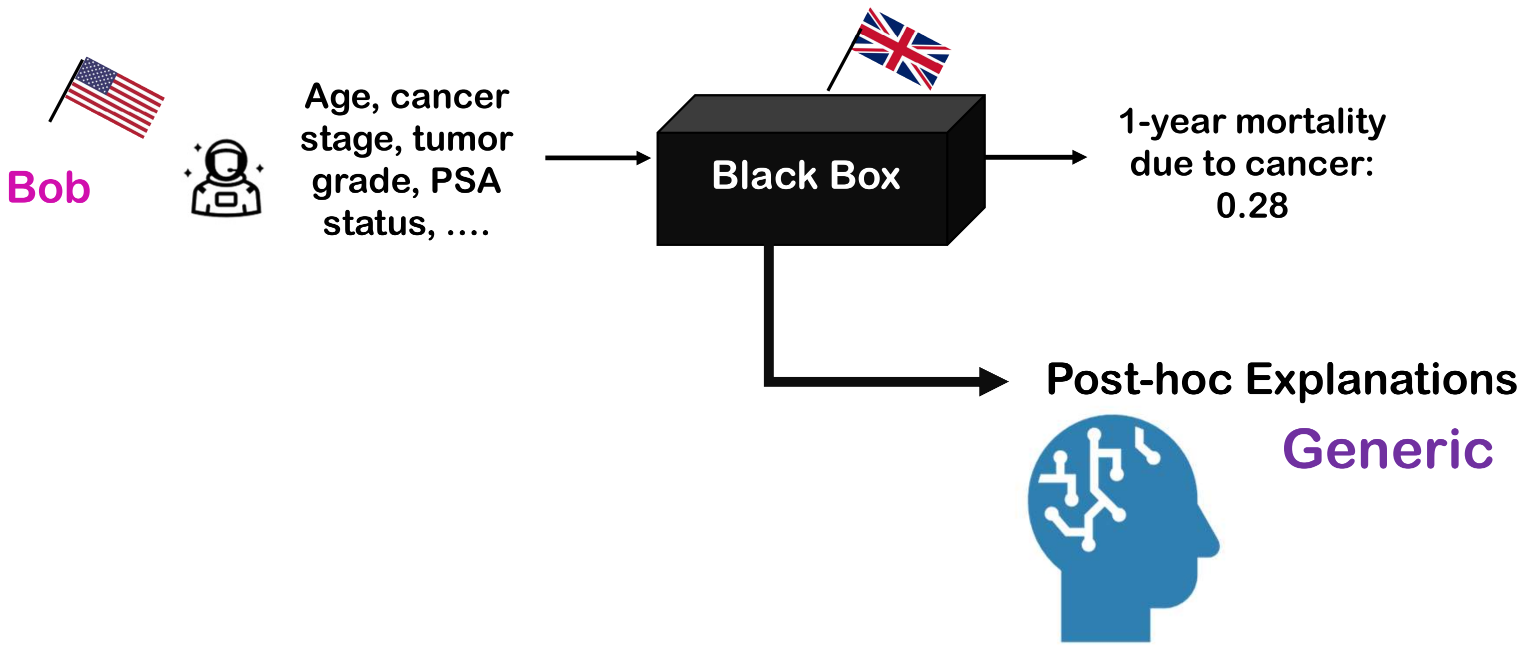
Climbing the "Discovery" Ladder

1. Explanatory features and examples - Beyond interpreting static predictions
   - Interpreting time-series models
   - Unsupervised models
   - Heterogeneous treatment effect models

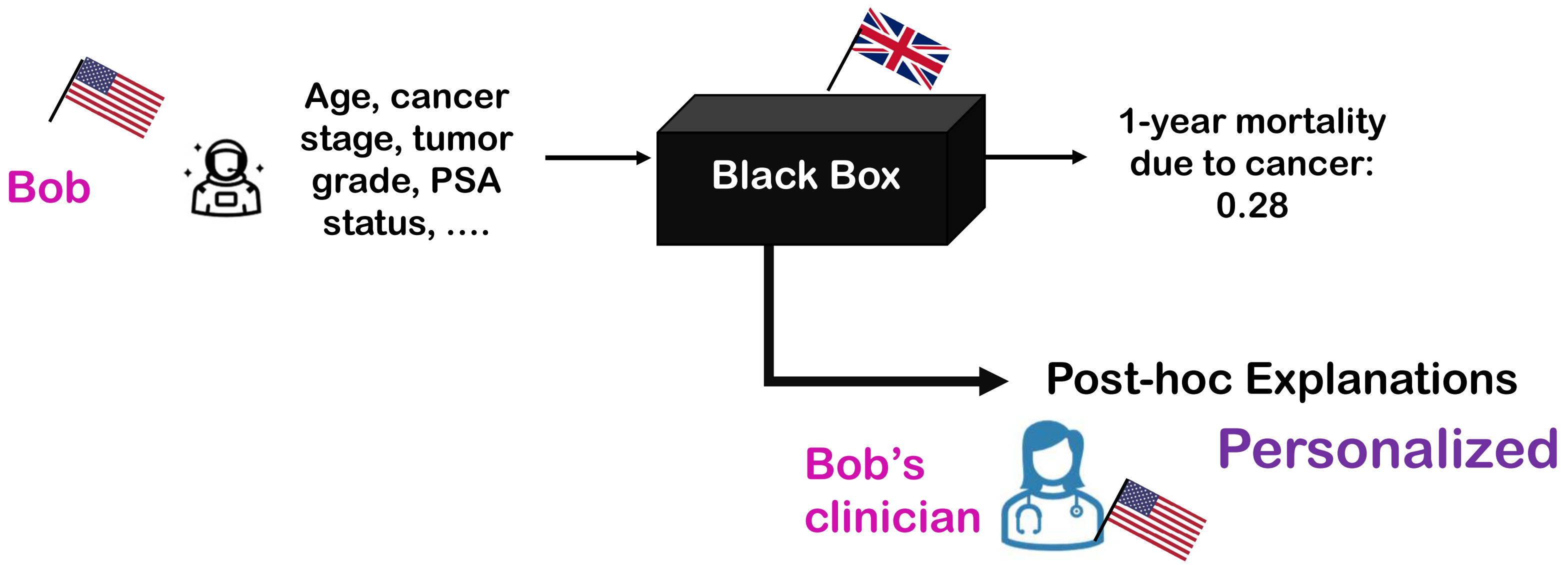
2. Personalized explanations - Beyond "one-size-fits-all" example explanations
   - **Give explanations based on specific patients selected by each clinician**

3. Discover governing equations of medicine - Going beyond association and causal relations

Personalized explanations – Beyond “one-size-fits-all” example explanations



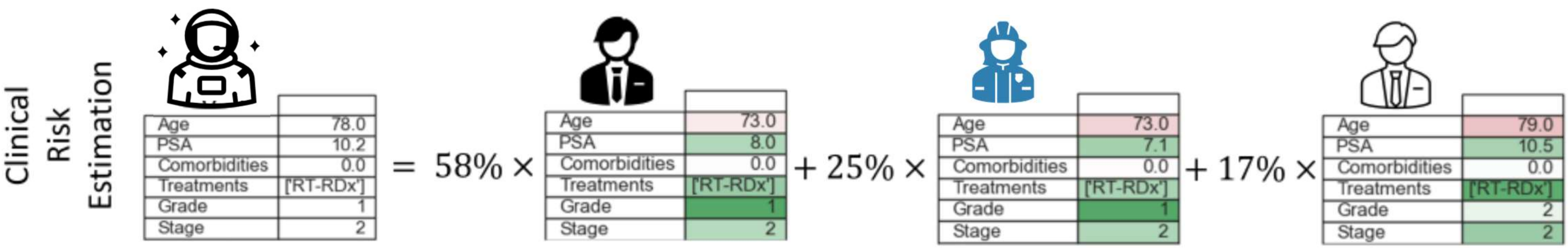
Personalized explanations – Beyond "one-size-fits-all" example explanations



Desiderata

Personalized explanations with reference to a freely selected set of examples, called the **corpus**

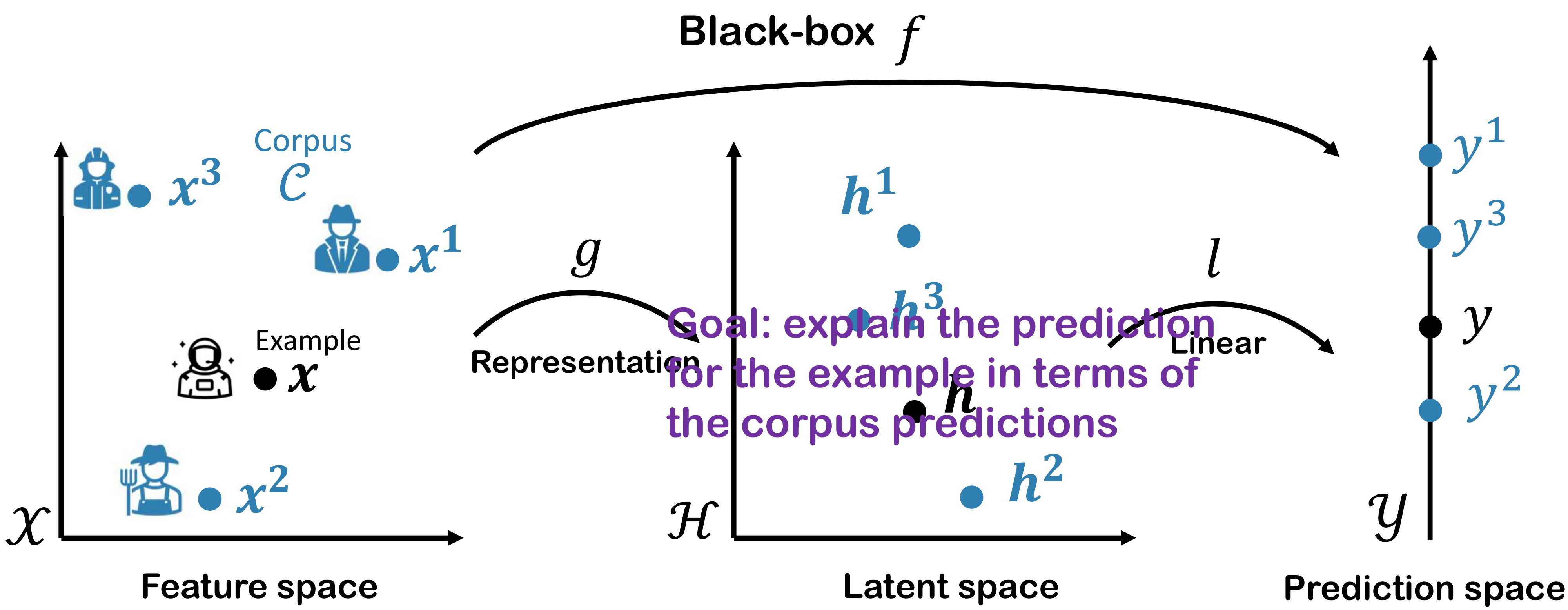
- ✓ **Which corpus examples** explain the prediction issued for a given test example?
- ✓ **What features** of these corpus examples are relevant for the model to relate them to the test example?



Our solution: SimplEx [Crabbe, Qian, Imrie, vdS, NeurIPS 2021]

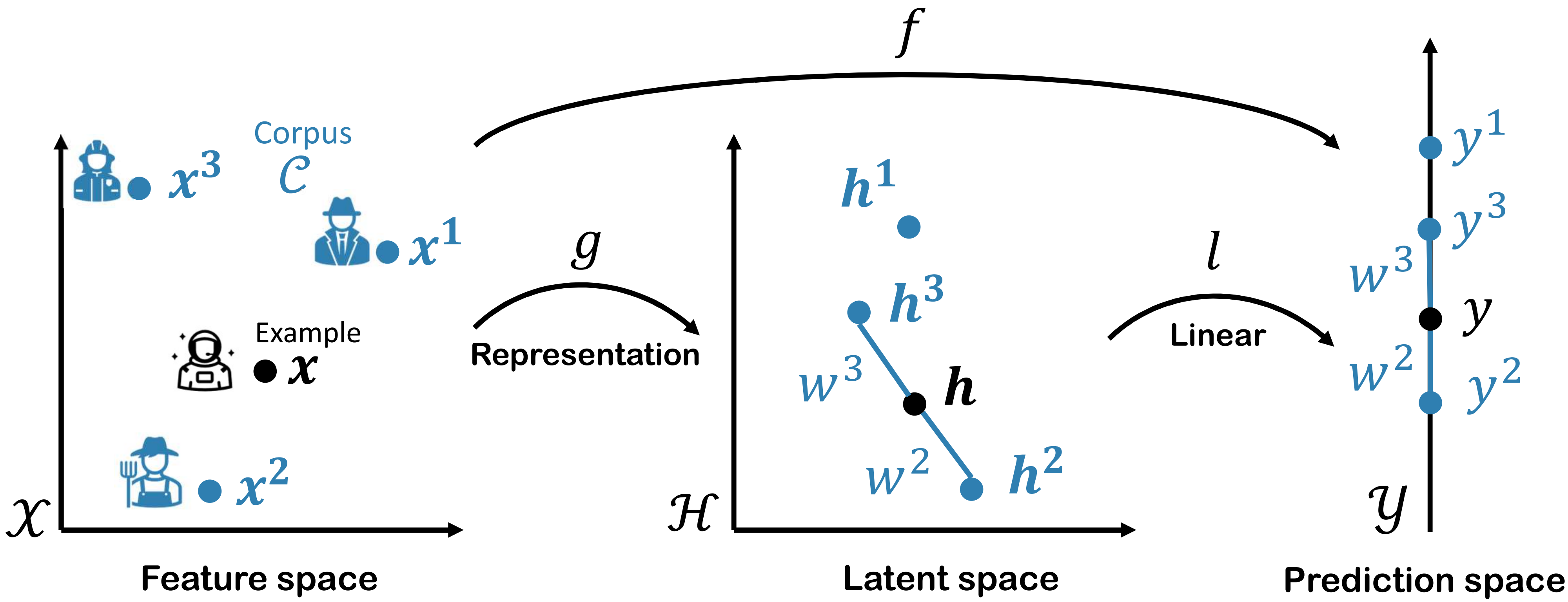
- ✓ SimplEx – able to reconstruct the test latent representation as a mixture of corpus latent representations
- ✓ Novel approach (Integrated Jacobian) allows SimplEx to make explicit the contribution of each corpus feature in the mixture
  - ✓ Bridge between feature importance & example-based explanations
- ✓ SimplEx gives the user freedom to choose the corpus of examples to explain model predictions in a user-centric way
- ✓ SimplEx provides user-centric explanations for any ML methods on diverse data (tabular, imaging, time-series, multi-modal)

SimplEx: Problem set-up



Goal: explain the prediction for the example in terms of the corpus predictions

SimplEx: Key idea



Corpus Decomposition

- **Find the best corpus decomposition of the example**

$$\widehat{\boldsymbol{h}} = \arg\min \left\| \boldsymbol{h} - \widetilde{\boldsymbol{h}} \right\|_{\mathcal{H}} \quad s.t. \quad \widetilde{\boldsymbol{h}} \in \mathcal{CH}(\mathcal{C})$$

How to transfer corpus explanations in the input space?

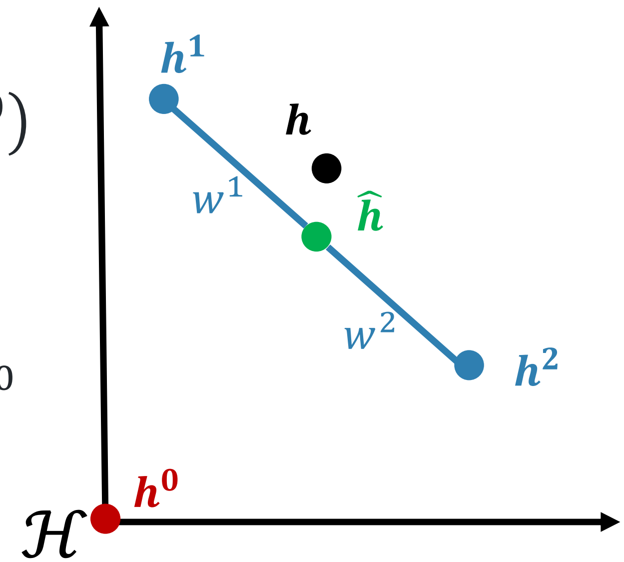
Idea:

fix a baseline input $\boldsymbol{x}^{\mathbf{0}}$ with representation $\boldsymbol{h}^0 = g(\boldsymbol{x}^{\mathbf{0}})$

$$\boldsymbol{h} - \boldsymbol{h}^{\mathbf{0}} \approx \sum_{c=1}^{C} w^c \underline{(\boldsymbol{h}^c - \boldsymbol{h}^{\mathbf{0}})}$$

Compare each corpus member $\boldsymbol{h}^c$ to the baseline $\boldsymbol{h}^0$

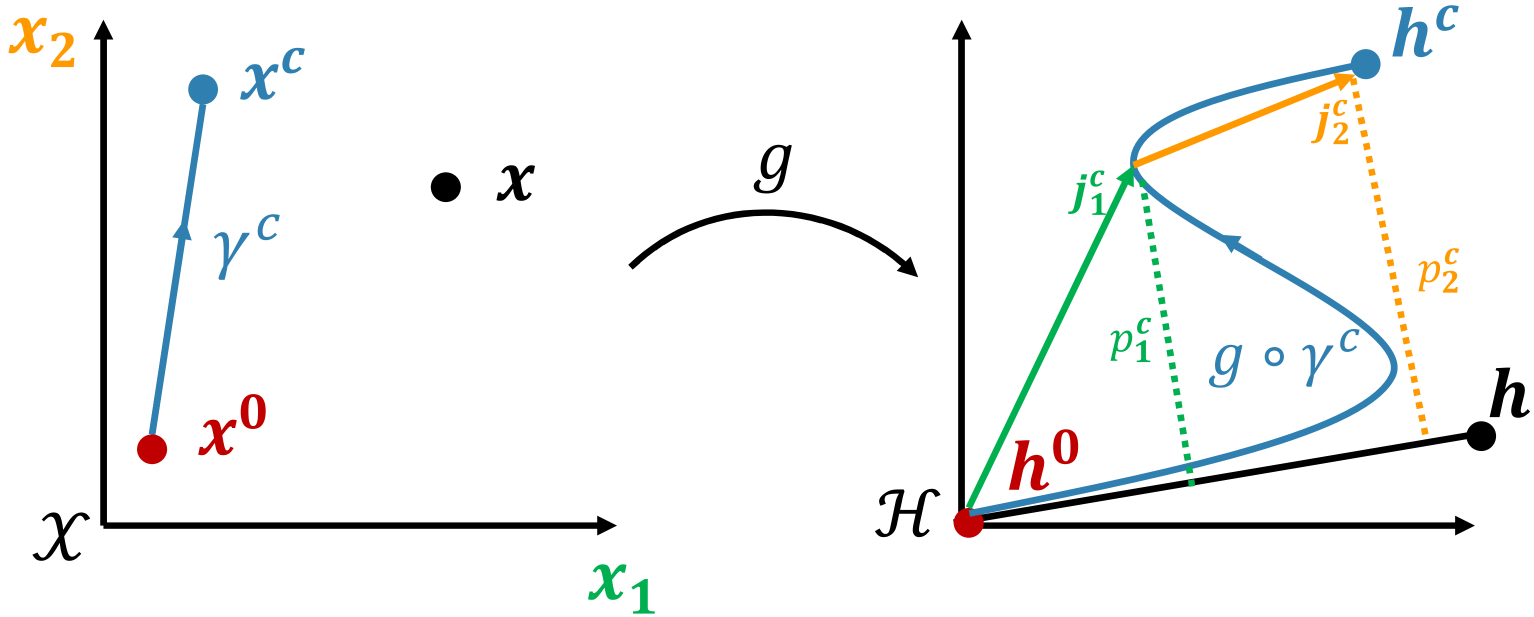
Understand total shift in latent space in terms of individual contributions from each corpus member



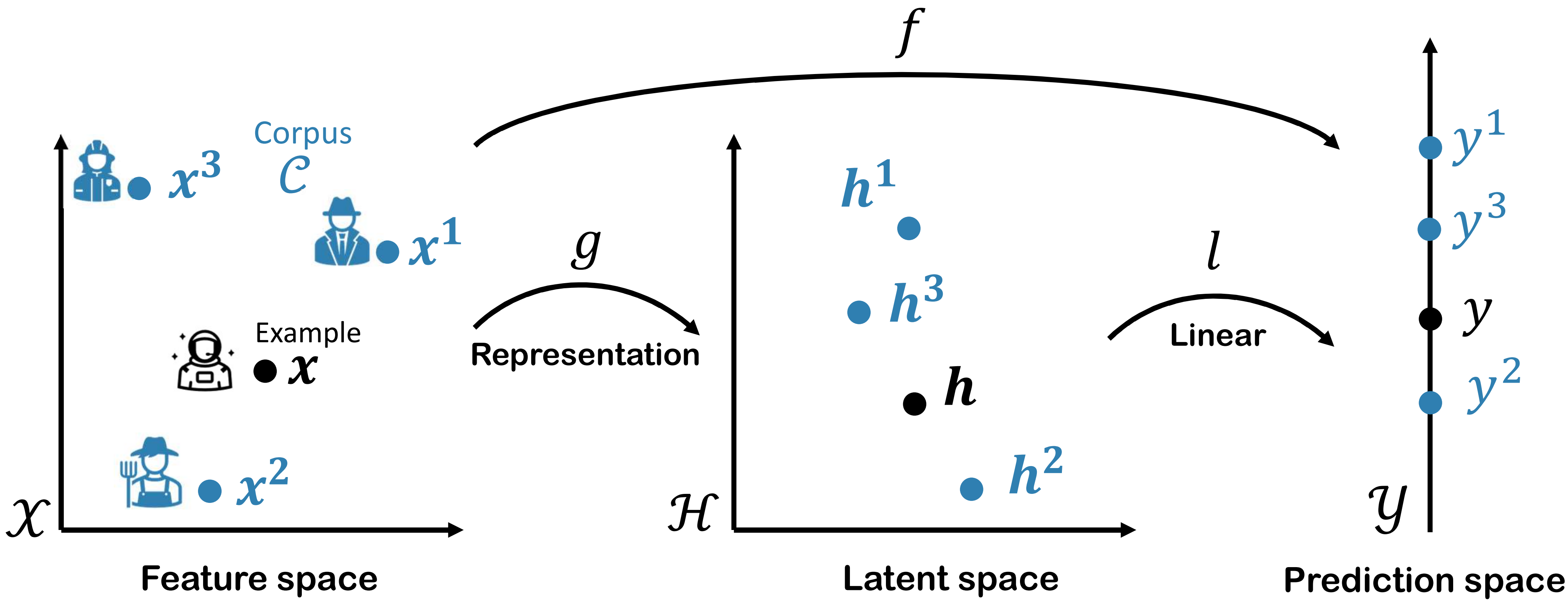
Integrated Jacobian & Projection

$$\boldsymbol{j}_i^c = \int_0^1 \frac{\partial g \circ \gamma^c}{\partial x_i}(t)\, dt$$

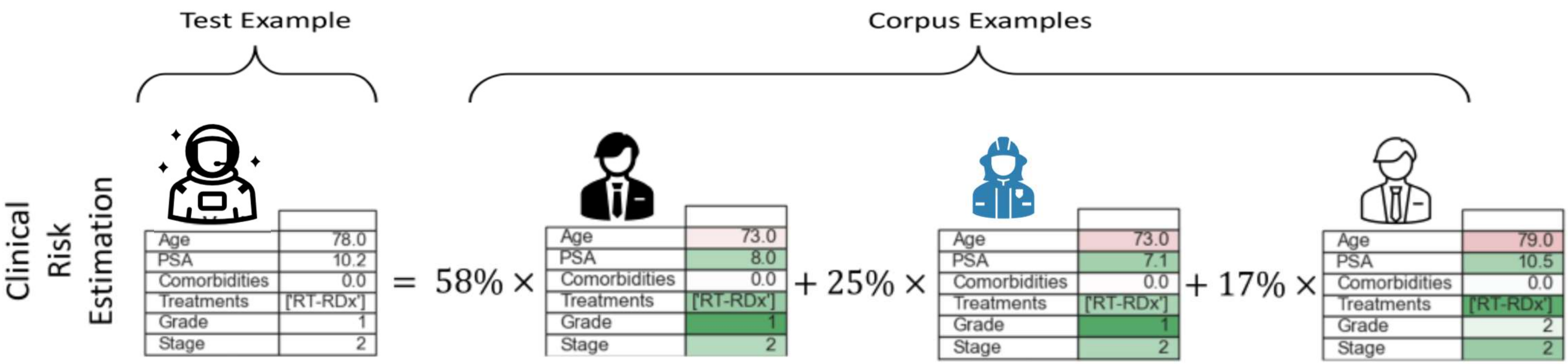
$$p_i^c = \frac{\langle \boldsymbol{h} - \boldsymbol{h}^0, \boldsymbol{j}_i^c \rangle}{\langle \boldsymbol{h} - \boldsymbol{h}^0, \boldsymbol{h} - \boldsymbol{h}^0 \rangle}$$



SimplEx: Feature sensitivity analysis



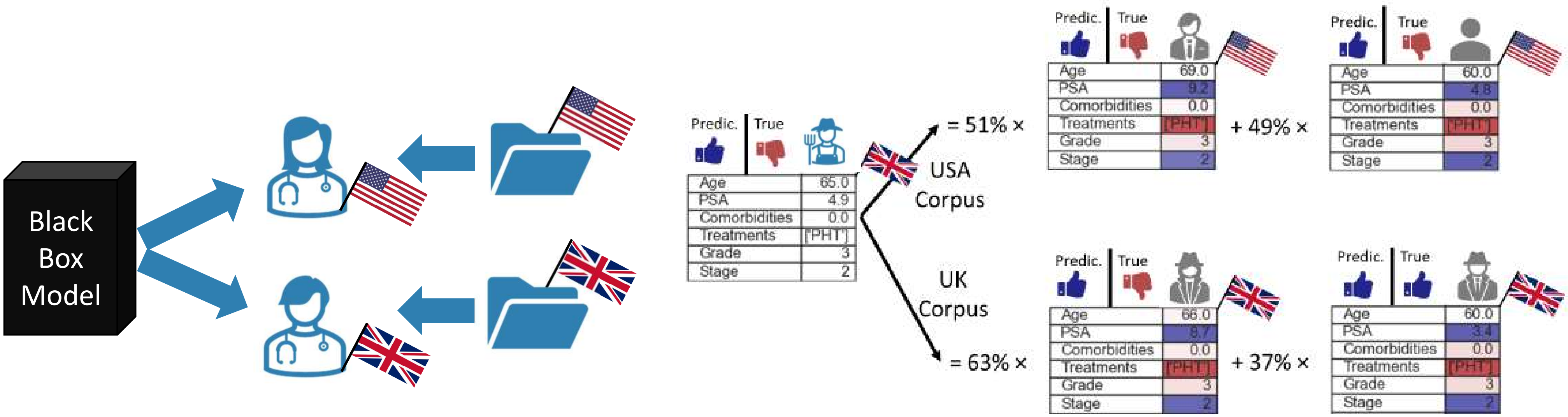
SimplEx Explanations: Going beyond current interpretability



Expanding the picture: SimplEx **unifies** example and feature-based explanations

Enhancing the picture: SimplEx **captures insights** from the model's **latent space**

SimplEx Explanations: Personalized



- **Debug the ML models**

  Present examples for which the model extrapolates

Climbing the "Discovery" Ladder

1. Explanatory features and examples - Beyond interpreting static predictions
   - Interpreting time-series models
   - Unsupervised models
   - Heterogeneous treatment effect models

2. Personalized explanations - Beyond "one-size-fits-all" example explanations
   - Give explanations based on specific patients selected by each clinician

3. Discover governing equations of medicine - Going beyond association and causal relations

Discovery of risk equations using ML

	Explicit function	Implicit function	Ordinary differential equation
Typical form	$y = f(x)$	$f(x, y) = c$	$\frac{dx}{dt} = f(x, t)$

Risk equations

NHS Predict Breast Cancer equations

- if ER+

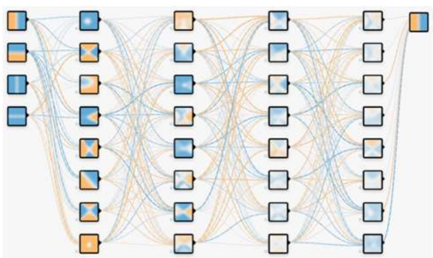
$$H_c(t) = \exp[0.7424402 - 7.527762/\sqrt{t} - 1.812513 * \log(t)/\sqrt{t}]$$

- if ER-

$$H_c(t) = \exp[-1.156036 + 0.4707332/t^2 - 3.51355/t].$$

Turning black boxes into white boxes using symbolic metamodels [Alaa & vdS, NeurIPS 2019] [Crabbe, Zhang, vdS, NeurIPS 2020]

Black-box ML model



$f(\mathbf{x})$

Symbolic Metamodeling

$$g(\mathbf{x}) = G(\mathbf{x}; \theta^*)$$

$$\theta^* = \arg\min_{\theta \in \Theta} \ell(f(\mathbf{x}), G(\mathbf{x}; \theta))$$

Explicit function

$$\alpha_1 X_1 + \alpha_2 X_2^2 + \alpha_3 X_1 X_2$$
$$\alpha_4 X_3^3 + \alpha_5 \log(X_4)$$

$g(\mathbf{x})$

Metamodels

Operates on a trained machine learning model and outputs a symbolic formula describing the model's prediction surface

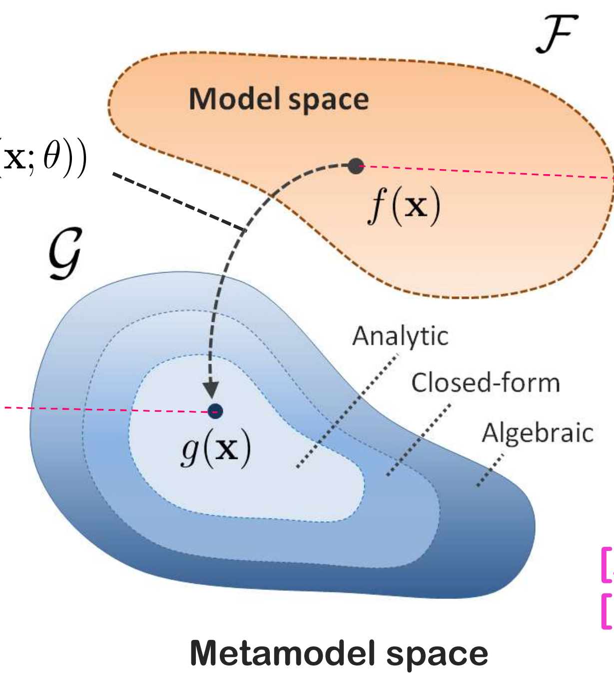
Building transparent risk equations of black-box ML

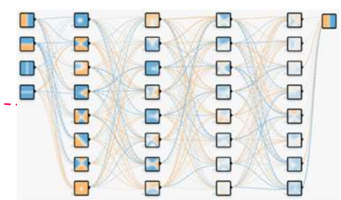
Metamodel representation $g(\mathbf{x}) = G(\mathbf{x}; \theta^*)$

Black-box ML model

Model space (uninterpretable)

$\theta^* = \arg\min_{\theta \in \Theta} \ell(f(\mathbf{x}), G(\mathbf{x}; \theta))$





White-box model

$\alpha_1 X_1 + \alpha_2 X_2^2 + \alpha_3 X_1\, X_2$

$\alpha_4 X_3^3 + \alpha_5 \log(X_4)$

Metamodel

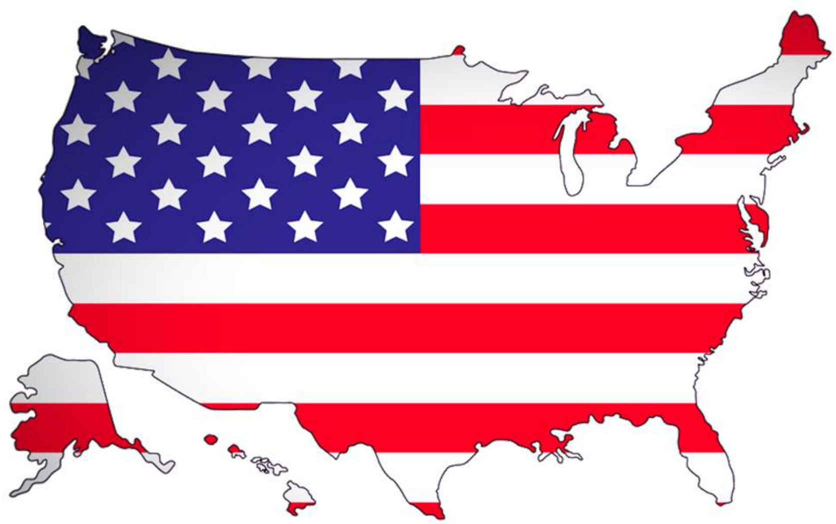
[Alaa & vdS, NeurIPS 2019]
[Crabbe, Zhang, vdS, NeurIPS 2020]

Interpretability using symbolic metamodeling in practice

[Alaa, Gurdasani, Harris, Rashbass & vdS, Nature MI, 2021]

Example: Predicting breast cancer risk survival (5 years)

Breast cancer data → NHS Predict/ AutoPrognosis → Prediction



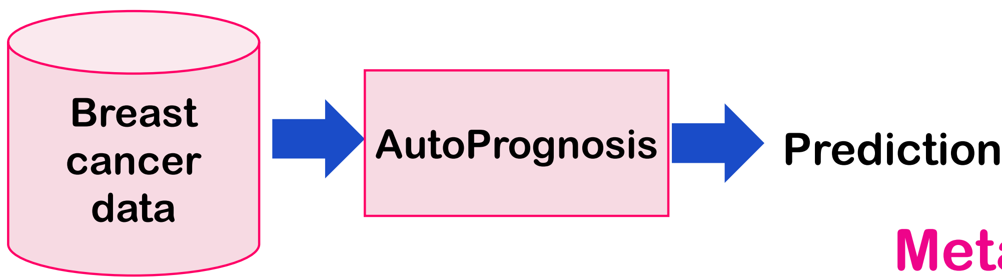
Nearly 1 million patients involved in the analysis.

NCRAS
> 390,000

Interpretability using symbolic metamodeling in practice

[Alaa, Gurdasani, Harris, Rashbass & vdS, Nature MI, 2021]

Example: Predicting breast cancer risk survival (5 years)



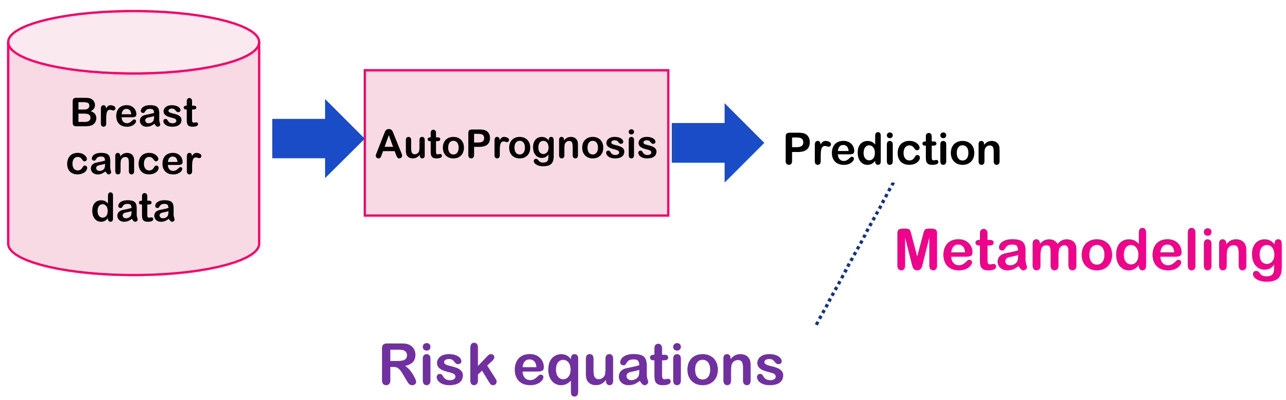
Metamodeling

Method	AUC-ROC
PREDICT	0.75 ± 0.0033
AutoPrognosis	0.84 ± 0.0032

Interpretability using symbolic metamodeling in practice

[Alaa, Gurdasani, Harris, Rashbass & vdS, Nature MI, 2021]

Example: Predicting breast cancer risk survival (5 years)



$$f(Age, ER, HER2, Tumor\ size, Grade, Nodes, Screening)$$

$$\exp\left(\frac{\text{Age}}{5} - \log\left(\frac{\text{Tumor size}}{100}\right) + \frac{1}{10}\log(\text{Nodes})\right) \times \exp\left(\frac{\text{ER}\cdot\text{Nodes}}{20} + \frac{\text{ER}\cdot\text{Tumor size}}{23}\right)$$

Interpretability using symbolic metamodeling in practice

[Alaa, Gurdasani, Harris, Rashbass & vdS, Nature MI, 2021]

Example: Predicting breast cancer risk survival (5 years)

Breast cancer data → AutoPrognosis → Prediction

Metamodeling

Risk equations

Method	AUC-ROC
PREDICT	0.75 ± 0.0033
~~AutoPrognosis~~	~~0.84 ± 0.0032~~
Metamodel	**0.83 ± 0.0020**

Illustration

Nancy

Age
gender
diabetes
BMI

ML Model

Risk prediction

$$g(\mathbf{x}) \quad \begin{aligned} &\alpha_0 \,\text{Age} + \alpha_1 \,\text{BMI}^2 + \alpha_2 \text{Age} \cdot \text{BMI} + \alpha_3 \text{Age} \cdot \text{Gender} \\ &\alpha_4 \text{Gender} \cdot (1 + \alpha_5 \,\text{Diabetes}) + \alpha_6 \log(\text{Age} \cdot \text{Diabetes} + 1) \end{aligned}$$

Explicit risk formulae

Individual-level feature importance

$$\frac{\partial\, g(\boldsymbol{x})}{\partial\, \text{Age}} = \alpha_0 + \alpha_2 \,\text{BMI} + \alpha_3 \,\text{Gender} + \frac{\alpha_6 \,\text{Diabetes}}{\text{Age+1}}$$

Discovery of governing equations using ML

	Explicit function	Implicit function	Ordinary differential equation
Typical form	$y = f(x)$	$f(x, y) = c$	$\frac{dx}{dt} = f(x, t)$

A much harder problem

$$x(t) \xleftrightarrow{\text{ODE}} x'(t)$$

Problem formulation

Dataset

$$\left\{ \begin{array}{c} \mathbf{y}_1(t) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{y}_D(t) \end{array} \right\}$$

$t \in \{t_1, t_2, \dots, T\}$

$\mathbf{y}_i(t) \in \mathbb{R}^J$

Goal: Discover

System of J ODEs

$$\left\{ \begin{array}{c} f_1(\boldsymbol{x}) = \dot{x}_1 \\ \cdot \\ \cdot \\ \cdot \\ f_J(\boldsymbol{x}) = \dot{x}_J \end{array} \right\}$$

$x_j: [0, T] \to \mathbb{R}$

$\boldsymbol{x}(t) = [x_1, \dots, x_J]^T$

$f_j: \mathbb{R}^J \to \mathbb{R}$

Related work

Table 1: Comparison of related works. "Data": the observed variables. "Allowed f^*": the space of discoverable functions. "Est.": the quantities estimated in the intermediate step. "$\dot{x}$ Free": is the method not reliant on $\dot{x}$? "$\boldsymbol{x}(0)$ Free": is the method not reliant on initial condition $\boldsymbol{x}(0)$? "Objective": the objective function. References: [1] Schmidt & Lipson (2009), [2] Brunton et al. (2016), [3] Gaucel et al. (2014) , [4] Chen et al. (2018).

Method	Data	Allowed f^*	Est.	$\dot{x}$ Free	$\boldsymbol{x}(0)$ Free	Objective
Symbolic Reg [1]	$a, \boldsymbol{b}$	Closed-form	None	-	-	$\|\|a - f(\boldsymbol{b})\|\|_2$
2-step Sparse [2]	$\boldsymbol{y}(t)$	$\sum \theta_k h_k(\boldsymbol{x})$	$\widehat{\dot{x}}$	$\times$	$\checkmark$	$\sum_t \|\|\widehat{\dot{x}}(t) - f(\boldsymbol{y}(t))\|\|_2$
2-step Symbolic [3]	$\boldsymbol{y}(t)$	Closed-form	$\widehat{\dot{x}}$	$\times$	$\checkmark$	$\sum_t \|\|\widehat{\dot{x}}(t) - f(\boldsymbol{y}(t))\|\|_2$

Estimate the derivative

Related work

Table 1: Comparison of related works. "Data": the observed variables. "Allowed f^*": the space of discoverable functions. "Est.": the quantities estimated in the intermediate step. "$\dot{x}$ Free": is the method not reliant on $\dot{x}$? "$\boldsymbol{x}(0)$ Free": is the method not reliant on initial condition $\boldsymbol{x}(0)$? "Objective": the objective function. References: [1] Schmidt & Lipson (2009), [2] Brunton et al. (2016), [3] Gaucel et al. (2014) , [4] Chen et al. (2018).

Method	Data	Allowed f^*	Est.	$\dot{x}$ Free	$\boldsymbol{x}(0)$ Free	Objective
Symbolic Reg [1]	$a, \boldsymbol{b}$	Closed-form	None	-	-	$\|\|a - f(\boldsymbol{b})\|\|_2$
2-step Sparse [2]	$\boldsymbol{y}(t)$	$\sum \theta_k h_k(\boldsymbol{x})$	$\widehat{\dot{x}}$	$\times$	$\checkmark$	$\sum_t \|\|\widehat{\dot{x}}(t) - f(\boldsymbol{y}(t))\|\|_2$
2-step Symbolic [3]	$\boldsymbol{y}(t)$	Closed-form	$\widehat{\dot{x}}$	$\times$	$\checkmark$	$\sum_t \|\|\widehat{\dot{x}}(t) - f(\boldsymbol{y}(t))\|\|_2$
ODE Approx [4]	$\boldsymbol{y}(t)$	Neural nets	$\widehat{\boldsymbol{x}}(0)$	$\checkmark$	$\times$	$\sum_t \|\|\boldsymbol{y}(t) - \widehat{\boldsymbol{x}}(t)\|\|_2$

Chen, R. T., Rubanova, Y., Bettencourt, J., & Duvenaud, D. (2018)
Neural ordinary differential equations

Related work

Function approximator approach

Learn true ODE with a function approximator $\hat{f}$ (NN, Gaussian process)

$$
\begin{tikzcd}
\hat{f} \arrow[r] & \hat{\boldsymbol{x}}(0) \arrow[r] & \hat{\boldsymbol{x}} \arrow[r] & \mathcal{L}(\hat{\boldsymbol{x}}, \boldsymbol{y}) \arrow[lll, bend left, "\text{Feedback}"]
\end{tikzcd}
$$

Discovered functions are not closed-form!

More general methods: Neural Laplace [Holt, Qian, vdS, ICML 2022]

Related work

Table 1: Comparison of related works. "Data": the observed variables. "Allowed f^*": the space of discoverable functions. "Est.": the quantities estimated in the intermediate step. "$\dot{x}$ Free": is the method not reliant on $\dot{x}$? "$\boldsymbol{x}(0)$ Free": is the method not reliant on initial condition $\boldsymbol{x}(0)$? "Objective": the objective function. References: [1] Schmidt & Lipson (2009), [2] Brunton et al. (2016), [3] Gaucel et al. (2014) , [4] Chen et al. (2018).

Method	Data	Allowed f^*	Est.	$\dot{x}$ Free	$\boldsymbol{x}(0)$ Free	Objective
Symbolic Reg [1]	$a, \boldsymbol{b}$	Closed-form	None	-	-	$\|\|a - f(\boldsymbol{b})\|\|_2$
2-step Sparse [2]	$\boldsymbol{y}(t)$	$\sum \theta_k h_k(\boldsymbol{x})$	$\hat{\dot{x}}$	$\times$	✓	$\sum_t \|\|\hat{\dot{x}}(t) - f(\boldsymbol{y}(t))\|\|_2$
2-step Symbolic [3]	$\boldsymbol{y}(t)$	Closed-form	$\hat{\dot{x}}$	$\times$	✓	$\sum_t \|\|\hat{\dot{x}}(t) - f(\boldsymbol{y}(t))\|\|_2$
ODE Approx [4]	$\boldsymbol{y}(t)$	Neural nets	$\hat{\boldsymbol{x}}(0)$	✓	$\times$	$\sum_t \|\|\boldsymbol{y}(t) - \hat{\boldsymbol{x}}(t)\|\|_2$
D-CODE	$\boldsymbol{y}(t)$	Closed-form	$\hat{\boldsymbol{x}}$	✓	✓	

D-CODE: Discovering Closed-Form ODEs
[Qian, Kacprzyk, vdS, ICLR 2022]

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\dot{x}_j(t) = f_j(\boldsymbol{x}(t)), \ \forall j = 1, \ldots, J, \ \forall t \in [0, T] \tag{1}$$

Characterize an ODE without referring to the derivative!

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\dot{x}_j(t) = f_j(\boldsymbol{x}(t)),\ \forall j = 1, \ldots, J,\ \forall t \in [0, T] \tag{1}$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set of continuously differentiable functions. We define the functionals

$$C_j\boxed{(f, \boldsymbol{x}, g)} := \int_0^T f\big(\boldsymbol{x}(t)\big) g(t) dt + \int_0^T x_j(t) \dot{g}(t) dt; \quad \forall j \in \{1, 2, \ldots, J\}$$

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\dot{x}_j(t) = f_j(\boldsymbol{x}(t)),\ \forall j = 1, \ldots, J,\ \forall t \in [0, T] \tag{1}$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set of continuously differentiable functions. We define the functionals

$$C_j(f, \boldsymbol{x}, g) := \boxed{\int_0^T f\big(\boldsymbol{x}(t)\big) g(t) dt} + \boxed{\int_0^T x_j(t) \dot{g}(t) dt;} \quad \forall j \in \{1, 2, \ldots, J\}$$

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\dot{x}_j(t) = f_j(\boldsymbol{x}(t)),\ \forall j = 1, \ldots, J,\ \forall t \in [0, T] \tag{1}$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set of continuously differentiable functions. We define the functionals

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f\left(\boxed{\boldsymbol{x}(t)}\right) g(t) dt + \int_0^T \boxed{x_j(t)}\, \dot{g}(t) dt; \quad \forall j \in \{1, 2, \ldots, J\}$$

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\dot{x}_j(t) = f_j(\boldsymbol{x}(t)),\ \forall j = 1, \dots, J,\ \forall t \in [0, T] \tag{1}$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set of continuously differentiable functions. We define the functionals

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f\big(\boldsymbol{x}(t)\big) g(t) dt + \int_0^T x_j(t) \boxed{\dot{g}(t)} dt; \quad \forall j \in \{1, 2, \dots, J\}$$

Hackbusch, W. (2017)
Variational Formulation

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\boxed{\dot{x}_j(t) = f_j(\boldsymbol{x}(t))}, \ \forall j = 1, \dots, J, \ \forall t \in [0, T] \qquad (1)$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set of continuously differentiable functions. We define the functionals

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f\big(\boldsymbol{x}(t)\big) g(t) dt + \int_0^T x_j(t) \dot{g}(t) dt; \quad \forall j \in \{1, 2, \dots, J\}$$

Proposition 1. *(Hackbusch, 2017) Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, a continuously differentiable function $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, and continuous functions $f_j : \mathbb{R}^J \to \mathbb{R}$ for $j = 1, \dots, J$. Then $\boldsymbol{x}$ is the solution to the system of ODEs in Equation 1 if and only if*

$$\boxed{C_j(f_j, \boldsymbol{x}, g) = 0,} \ \forall j \in \{1, \dots, J\}, \ \forall g \in \mathcal{C}^1[0, T], \ g(0) = g(T) = 0$$

Hackbusch, W. (2017)
Variational Formulation

D-CODE: motivation

Variational formulation of ordinary differential equations

$$\boxed{\dot{x}_j(t) = f_j(\boldsymbol{x}(t))}, \ \forall j = 1, \ldots, J, \ \forall t \in [0, T] \qquad (1)$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set of continuously differentiable functions. We define the functionals

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f\big(\boldsymbol{x}(t)\big) g(t) dt + \int_0^T x_j(t) \dot{g}(t) dt; \quad \forall j \in \{1, 2, \ldots, J\}$$

Proposition 1. *(Hackbusch, 2017) Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, a continuously differentiable function $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, and continuous functions $f_j : \mathbb{R}^J \to \mathbb{R}$ for $j = 1, \ldots, J$. Then $\boldsymbol{x}$ is the solution to the system of ODEs in Equation 1 if and only if*

$$\boxed{C_j(f_j, \boldsymbol{x}, g) = 0,} \ \forall j \in \{1, \ldots, J\}, \ \boxed{\forall g \in \mathcal{C}^1[0, T], \ g(0) = g(T) = 0}$$

Hackbusch, W. (2017)
Variational Formulation

D-CODE: motivation

Variational formulation of ordinary differential equations

Proposition 1. *(Hackbusch, 2017) Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, a continuously differentiable function $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, and continuous functions $f_j : \mathbb{R}^J \to \mathbb{R}$ for $j = 1, \ldots, J$. Then $\boldsymbol{x}$ is the solution to the system of ODEs in Equation 1 if and only if*

$$\boxed{C_j(f_j, \boldsymbol{x}, g) = 0}, \ \forall j \in \{1, \ldots, J\}, \ \boxed{\forall g \in \mathcal{C}^1[0, T], \ g(0) = g(T) = 0}$$

Proposition 1. *(Hackbusch, 2017) Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, a continuously differentiable function $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, and continuous functions $f_j : \mathbb{R}^J \to \mathbb{R}$ for $j = 1, \ldots, J$. Then $\boldsymbol{x}$ is the solution to the system of ODEs in Equation 1 if and only if*

$$\boxed{C_j(f_j, \boldsymbol{x}, g) = 0}, \ \forall j \in \{1, \ldots, J\}, \ \boxed{\forall g \in \mathcal{C}^1[0, T], \ g(0) = g(T) = 0}$$

We expect that minimising value of this functional corresponds to finding better approximations of the true ODE.

D-CODE: theory

$$d_{\boldsymbol{x}}(f, f^*) := ||f \circ \boldsymbol{x} - f^* \circ \boldsymbol{x}||_2 = ||(f - f^*) \circ \boldsymbol{x}||_2$$

Theorem 1. *Consider $J \in \mathbb{N}^+$, $j \in \{1, \dots, J\}$, $T \in \mathbb{R}^+$. Let $f^* : \mathbb{R}^J \to \mathbb{R}$ be a continuous function, and let $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$ be a continuously differentiable function satisfying $\dot{x}_j(t) = f^*(\boldsymbol{x}(t))$. Consider a sequence of functions $(\widehat{\boldsymbol{x}}_k)$, where $\widehat{\boldsymbol{x}}_k : [0, T] \to \mathbb{R}^J$ is a continuously differentiable function. If $(\widehat{\boldsymbol{x}}_k)$ converges to $\boldsymbol{x}$ in L^2 norm. Then for any Lipschitz continuous function f*

$$\lim_{S \to \infty} \lim_{k \to \infty} \sum_{s=1}^{S} C_j(f, \widehat{\boldsymbol{x}}_k, g_s)^2 = d_{\boldsymbol{x}}(f, f^*)^2, \tag{7}$$

where $\{g_1, g_2, \dots\}$ is a Hilbert (orthonormal) basis for $L^2[0, T]$ such that $\forall i$, $g_i(0) = g_i(T) = 0$ and $g_i \in \mathcal{C}^1[0, T]$.

D-CODE: theory

$$d_{\boldsymbol{x}}(f, f^*) := ||f \circ \boldsymbol{x} - f^* \circ \boldsymbol{x}||_2 = ||(f - f^*) \circ \boldsymbol{x}||_2$$

Theorem 1. *Consider* $J \in \mathbb{N}^+$, $j \in \{1, \ldots, J\}$, $T \in \mathbb{R}^+$. *Let* $f^* : \mathbb{R}^J \to \mathbb{R}$ *be a continuous function, and let* $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$ *be a continuously differentiable function satisfying* $\dot{x}_j(t) = f^*(\boldsymbol{x}(t))$. *Consider a sequence of functions* $(\widehat{\boldsymbol{x}}_k)$, *where* $\widehat{\boldsymbol{x}}_k : [0, T] \to \mathbb{R}^J$ *is a continuously differentiable function. If* $(\widehat{\boldsymbol{x}}_k)$ *converges to* $\boldsymbol{x}$ *in* L^2 *norm. Then for any Lipschitz continuous function* f

$$\lim_{S \to \infty} \lim_{k \to \infty} \sum_{s=1}^{S} C_j(f, \widehat{\boldsymbol{x}}_k, g_s)^2 = d_{\boldsymbol{x}}(f, f^*)^2, \tag{7}$$

where $\{g_1, g_2, \ldots\}$ *is a Hilbert (orthonormal) basis for* $L^2[0, T]$ *such that* $\forall i$, $g_i(0) = g_i(T) = 0$ *and* $g_i \in \mathcal{C}^1[0, T]$.

D-CODE: theory

$$d_{\boldsymbol{x}}(f, f^*) := ||f \circ \boldsymbol{x} - f^* \circ \boldsymbol{x}||_2 = ||(f - f^*) \circ \boldsymbol{x}||_2$$

Theorem 1. *Consider $J \in \mathbb{N}^+$, $j \in \{1, \ldots, J\}$, $T \in \mathbb{R}^+$. Let $f^* : \mathbb{R}^J \to \mathbb{R}$ be a continuous function, and let $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$ be a continuously differentiable function satisfying $\dot{x}_j(t) = f^*(\boldsymbol{x}(t))$. Consider a sequence of functions $(\widehat{\boldsymbol{x}}_k)$, where $\widehat{\boldsymbol{x}}_k : [0, T] \to \mathbb{R}^J$ is a continuously differentiable function. If $(\widehat{\boldsymbol{x}}_k)$ converges to $\boldsymbol{x}$ in L^2 norm. Then for any Lipschitz continuous function f*

$$\lim_{S \to \infty} \lim_{k \to \infty} \sum_{s=1}^{S} C_j(f, \widehat{\boldsymbol{x}}_k, g_s)^2 = d_{\boldsymbol{x}}(f, f^*)^2, \tag{7}$$

where $\{g_1, g_2, \ldots\}$ is a Hilbert (orthonormal) basis for $L^2[0, T]$ such that $\forall i$, $g_i(0) = g_i(T) = 0$ and $g_i \in \mathcal{C}^1[0, T]$.

Natural choice

$$g_s(t) = \sqrt{2/T} \sin(s\pi t/T)$$

D-CODE: algorithm

Preprocessing

Optimization

D-CODE: algorithm

Preprocessing

$$\left\{ \begin{array}{c} \mathbf{y}_1(t) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{y}_D(t) \end{array} \right\}$$

$t \in \{t_1, t_2, \dots, T\}$

$\mathbf{y}_i(t) \in \mathbb{R}^J$

Denoise & Interpolate

Optimization

D-CODE: algorithm

Preprocessing

$$\begin{Bmatrix} \mathbf{y}_1(t) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{y}_D(t) \end{Bmatrix} \xrightarrow{\text{Estimate}} \begin{Bmatrix} \hat{\boldsymbol{x}}_1(t) \\ \cdot \\ \cdot \\ \cdot \\ \hat{\boldsymbol{x}}_D(t) \end{Bmatrix}$$

$t \in \{t_1, t_2, \dots, T\}$

$\mathbf{y}_i(t) \in \mathbb{R}^J$

$t \in [0, T]$

$\boldsymbol{x}_i(t) \in \mathbb{R}^J$

Optimization

D-CODE: algorithm

Preprocessing

$$\begin{Bmatrix} \mathbf{y}_1(t) \\ \vdots \\ \mathbf{y}_D(t) \end{Bmatrix} \xrightarrow{\text{Estimate}} \begin{Bmatrix} \hat{\boldsymbol{x}}_1(t) \\ \vdots \\ \hat{\boldsymbol{x}}_D(t) \end{Bmatrix}$$

$t \in \{t_1, t_2, \dots, T\}$

$\mathbf{y}_i(t) \in \mathbb{R}^J$

$t \in [0, T]$

$\boldsymbol{x}_i(t) \in \mathbb{R}^J$

Optimization

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f(\boldsymbol{x}(t))g(t)dt + \int_0^T x_j(t)\dot{g}(t)dt$$

$$\widehat{f}_j = \arg\min_f \sum_{i=1}^{N} \sum_{s=1}^{S} C_j(f, \widehat{\boldsymbol{x}}_i, g_s)^2$$

D-CODE: algorithm

Preprocessing

$$\begin{Bmatrix} \mathbf{y}_1(t) \\ \vdots \\ \mathbf{y}_D(t) \end{Bmatrix} \xrightarrow{\text{Estimate}} \begin{Bmatrix} \hat{\boldsymbol{x}}_1(t) \\ \vdots \\ \hat{\boldsymbol{x}}_D(t) \end{Bmatrix}$$

$t \in \{t_1, t_2, \dots, T\}$

$\mathbf{y}_i(t) \in \mathbb{R}^J$

$t \in [0, T]$

$\boldsymbol{x}_i(t) \in \mathbb{R}^J$

Optimization

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f(\boldsymbol{x}(t)) g(t) dt + \int_0^T x_j(t) \dot{g}(t) dt$$

$$\widehat{f}_j = \arg\min_f \sum_{i=1}^{N} \sum_{s=1}^{S} C_j(f, \widehat{\boldsymbol{x}}_i, g_s)^2$$

Prespecified testing functions

$$g_s(t) = \sqrt{2/T} \sin(s\pi t / T)$$

Symbolic regression

D-CODE: experiments

Dynamical systems:
- Gompertz model
- Generalized logistic model
- Glycolytic oscillator
- Lorenz system

Benchmarks:
Two-step symbolic regression with

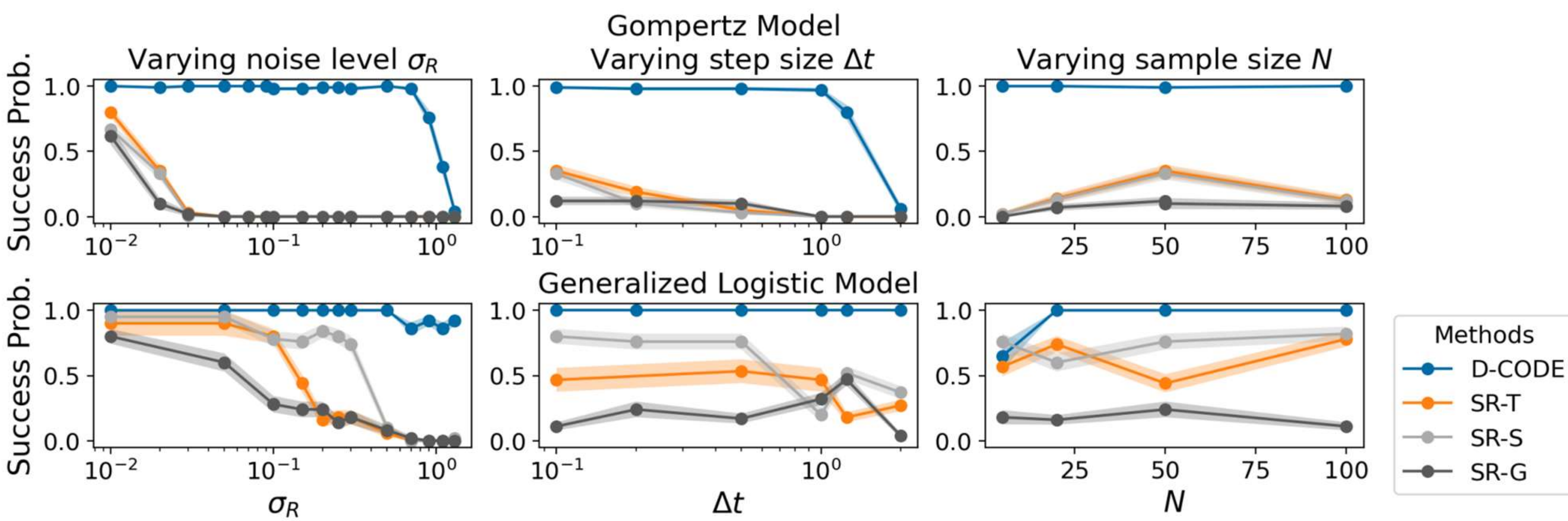
a) total variation regularized differentiation (SR-T)
b) spline-smoothed differentiation (SR-S)
c) Gaussian process smoothed differentiation (SR-G)

D-CODE: Experiments

$\dot{x}(t) = -\theta_1 x(t) \cdot \log\left(\theta_2 x(t)\right)$ Gompertz Model

$\dot{x}(t) = \theta_1 x(t) \cdot \left(1 - x(t)^{\theta_2}\right)$ Generalized Logistic Model

asymmetric growth with saturation



van_der_Schaar \LAB

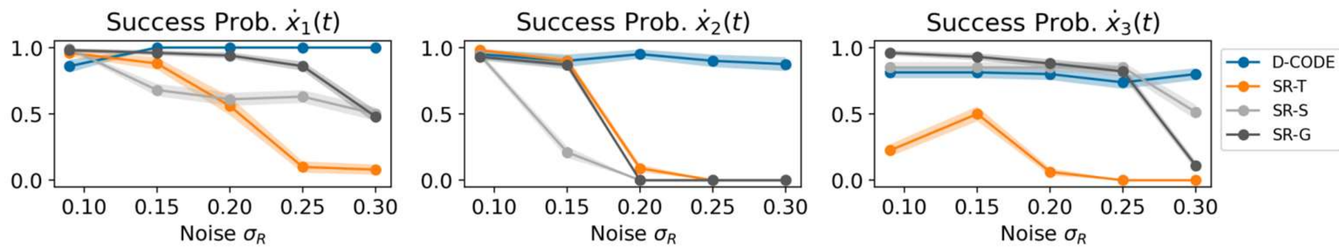
vanderschaar-lab.com

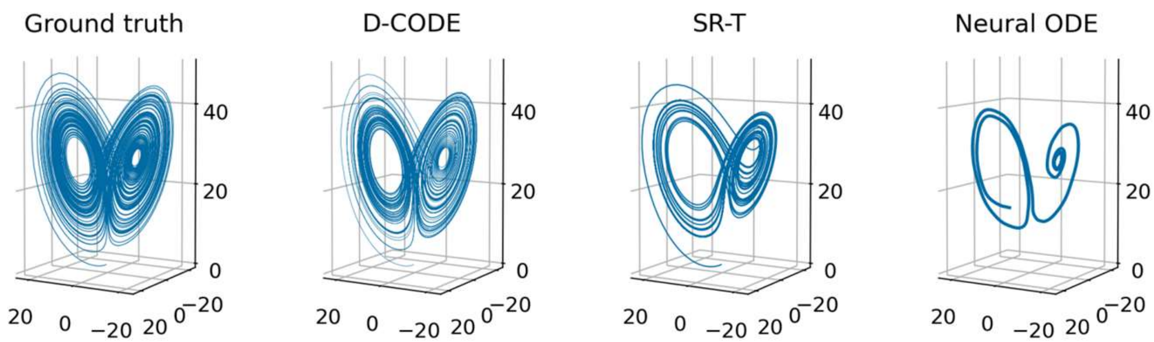
UNIVERSITY OF CAMBRIDGE

D-CODE: Experiments

Chaotic Lorenz system. The Lorenz system is a model system for chaotic dynamics, defined as:

$$\dot{x}_1(t) = \theta_1\big(x_2(t) - x_1(t)\big); \;\; \dot{x}_2(t) = x_1(t)\big(\theta_2 - x_3(t)\big) - x_2(t); \;\; \dot{x}_3(t) = x_1(t)x_2(t) - \theta_3 x_3(t)$$

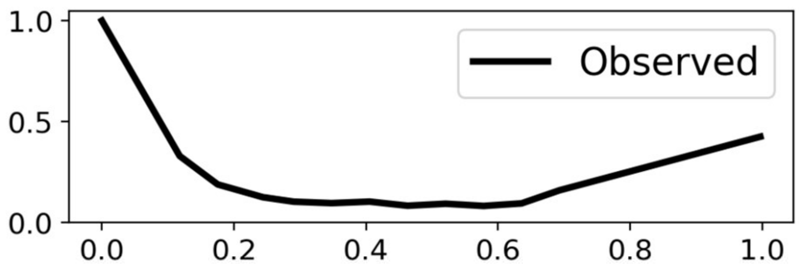




chaotic &
non-periodic systems

D-CODE in action

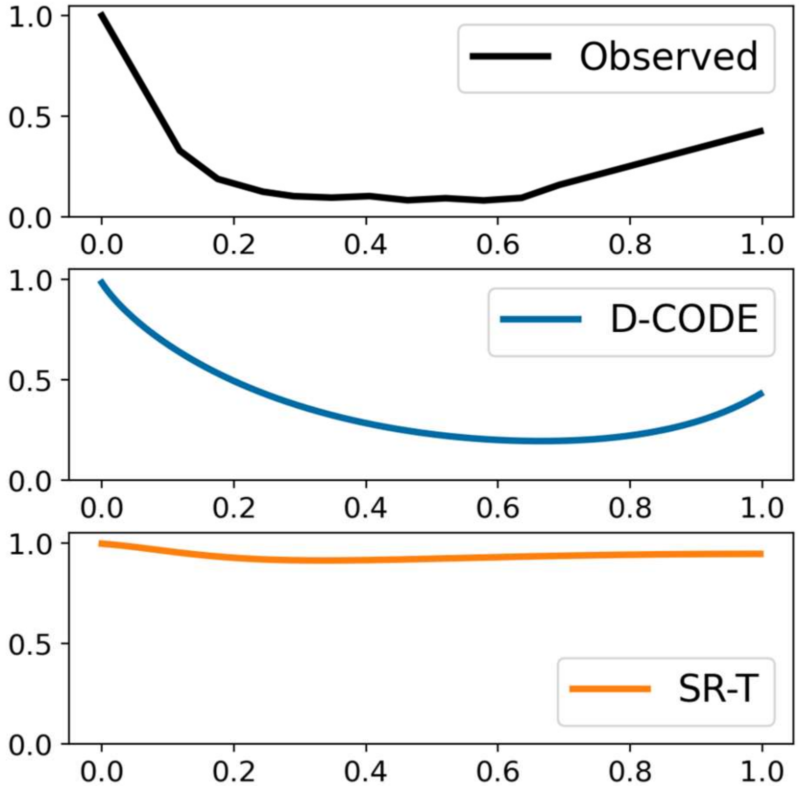
Discover temporal effects of chemotherapy on tumor volume



Dataset: 8 clinical trials on cancer patients

D-CODE in action

Discover temporal effects of chemotherapy on tumor volume



Dataset: 8 clinical trials on cancer patients

The following two ODEs are discovered by D-CODE and SR-T.

$$\dot{x}(t) = 4.48t^2 x(t) + \log(t); \quad \text{D-CODE}$$

$$\dot{x}(t) = 4x(t)\log\big(tx(t)\big)\log\big(tx(t) + 2t\big); \quad \text{SR-T}$$

Discovery of governing equations using ML

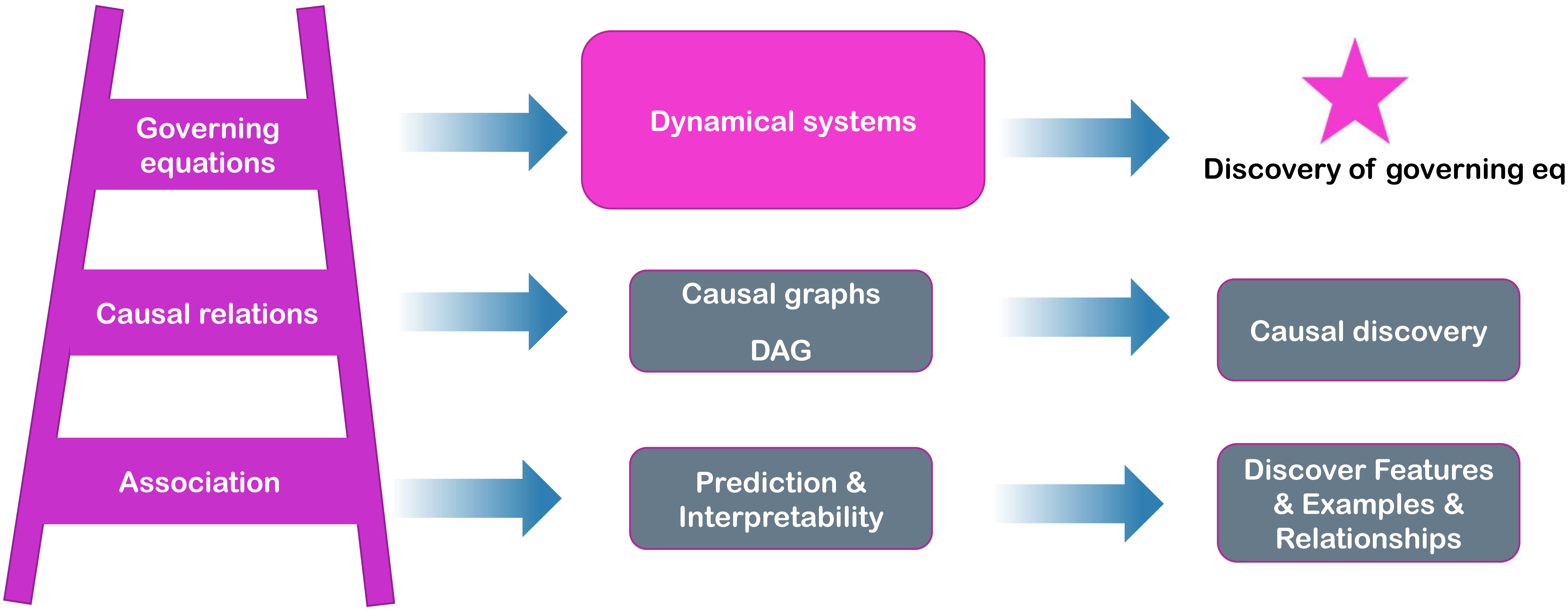
	Explicit function	Implicit function	Ordinary differential equation	Partial differential equation
Typical form	$y = f(x)$	$f(x, y) = c$	$\frac{dx}{dt} = f(x, t)$	$\frac{\partial u}{\partial t} = f(u, x)$

Symbolic Metamodels [NeurIPS '19, '20]

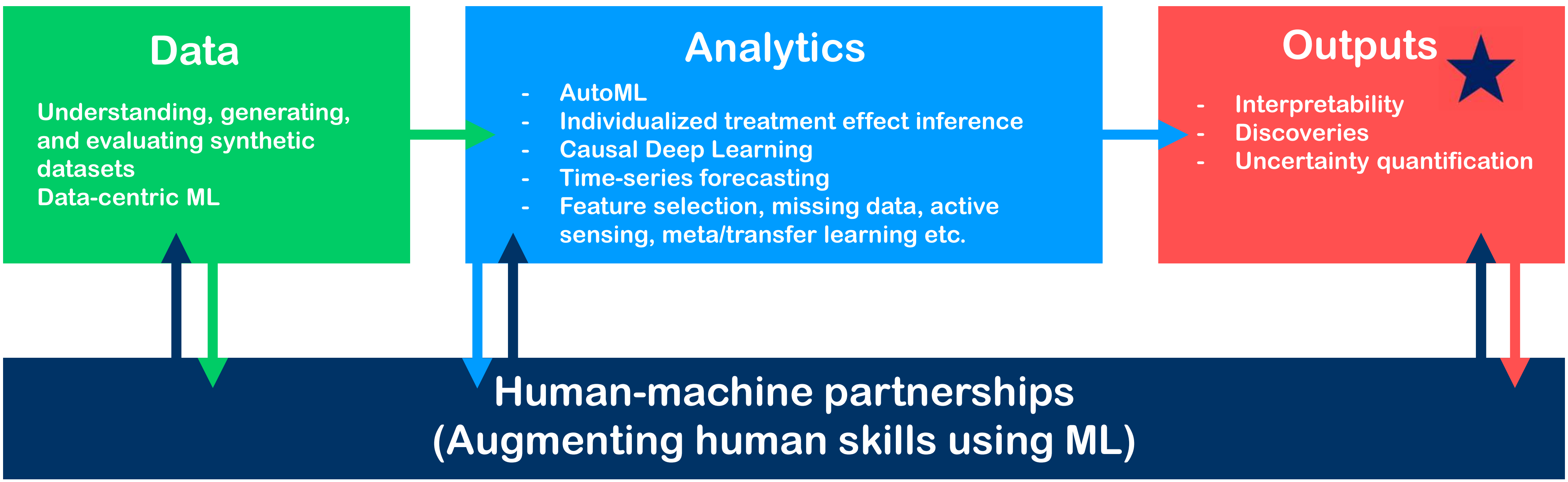
D-Code [ICLR '22]

D-CIPHER [archive]

Let us climb the “Discovery” Ladder!



Cutting-edge ML methods for healthcare



van_der_Schaar \LAB

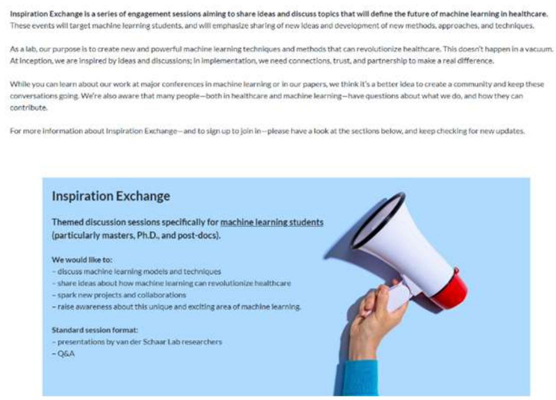
vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

Engagement sessions: Inspiration Exchange

Online engagement sessions for ML researchers in healthcare; themed presentations & Q&A

https://www.vanderschaar-lab.com/
→ Engagement sessions
→ Inspiration Exchange



Inspiration Exchange is a series of engagement sessions aiming to share ideas and discuss topics that will define the future of machine learning in healthcare. These events will target machine learning students, and will emphasize sharing of new ideas and development of new methods, approaches, and techniques.

As a lab, our purpose is to create new and powerful machine learning techniques and methods that can revolutionize healthcare. This doesn't happen in a vacuum. At inception, we are inspired by ideas and discussions; in implementation, we need connections, trust, and partnership to make a real difference.

While you can learn about our work at major conferences in machine learning or in our papers, we think it's a better idea to create a community and keep these conversations going. We're also aware that many people—both in healthcare and machine learning—have questions about what we do, and how they can contribute.

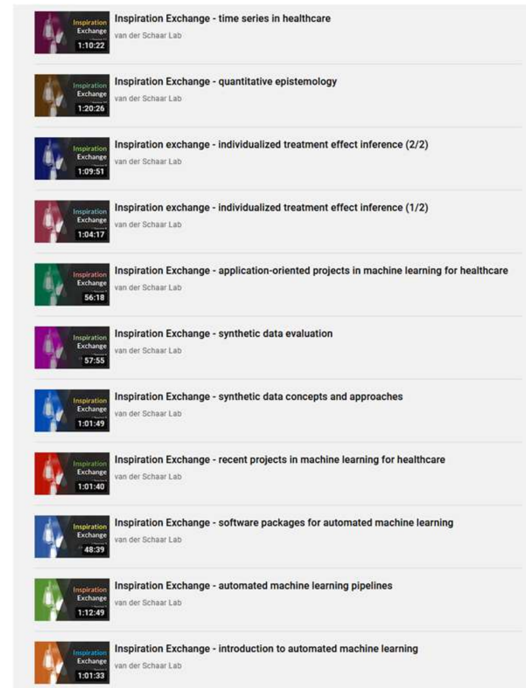
For more information about Inspiration Exchange—and to sign up to join in—please have a look at the sections below, and keep checking for new updates.

Inspiration Exchange

Themed discussion sessions specifically for machine learning students (particularly masters, Ph.D., and post-docs).

We would like to:
- discuss machine learning models and techniques
- share ideas about how machine learning can revolutionize healthcare
- spark new projects and collaborations
- raise awareness about this unique and exciting area of machine learning.

Standard session format:
- presentations by van der Schaar Lab researchers
- Q&A



Inspiration Exchange - time series in healthcare — van der Schaar Lab — 1:10:22
Inspiration Exchange - quantitative epistemology — van der Schaar Lab — 1:20:26
Inspiration exchange - individualized treatment effect inference (2/2) — van der Schaar Lab — 1:09:51
Inspiration exchange - individualized treatment effect inference (1/2) — van der Schaar Lab — 1:04:17
Inspiration Exchange - application-oriented projects in machine learning for healthcare — van der Schaar Lab — 56:18
Inspiration Exchange - synthetic data evaluation — van der Schaar Lab — 57:55
Inspiration Exchange - synthetic data concepts and approaches — van der Schaar Lab — 1:01:49
Inspiration Exchange - recent projects in machine learning for healthcare — van der Schaar Lab — 1:01:40
Inspiration Exchange - software packages for automated machine learning — van der Schaar Lab — 48:39
Inspiration Exchange - automated machine learning pipelines — van der Schaar Lab — 1:12:49
Inspiration Exchange - introduction to automated machine learning — van der Schaar Lab — 1:01:33

van_der_Schaar \LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

PhD & Postdoc positions with the van der Schaar Lab

University of Cambridge // position starting Winter, Spring and Fall 2023

What we can offer you:

1. World-leading research
A small research team that punches well above its weight. 26 papers at ICLR, ICML and NeurIPS in the last year alone.

2. Freedom to think big and explore
Ground-breaking work across an enormous range of ML sub-fields – from new deep learning theory and algorithms to methods for causal inference, graphical models, statistical machine learning, time series data, interpretable ML, and reinforcement learning.

3. Projects with a purpose; work that can change the world
Targeted & practical research projects; opportunities to apply ML to real-world problems in healthcare and help catalyze a revolution in medicine.

4. Unmatched prospects
Alumni around the world are leaders in their fields, with some continuing to full professorships and others joining top private-sector teams - DeepMind, Google Research, Meta

Interested? Let us know.

https://www.vanderschaar-lab.com/join-the-van-der-schaar-lab/

van_der_Schaar
\LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE