AI & Causal Modeling: Modeling Relationships Among Quality of Life at Work and Economic Performances

Michèle Sebag

TAU

CNRS — INRIA — LRI — Université Paris-Sud

Joint work: Olivier Goudet, Diviyan Kalainathan, Philippe Caillou, Isabelle Guyon, Paola Tubaro

AI & Causal Modeling: Modeling Relationships Among Quality of Life at Work and Economic Performances

Michèle Sebag

Tackling the Underspecified

CNRS – INRIA – LRI – Université Paris-Sud

Joint work: Olivier Goudet, Diviyan Kalainathan, Philippe Caillou, Isabelle Guyon, Paola Tubaro

Artificial Intelligence / Machine Learning / Data Science

A Case of Irrational Scientific Exuberance

- Underspecified goals — Big Data cures everything
- Underspecified limitations — Big Data can do anything (if big enough)
- Underspecified caveats — Big Data and Big Brother

Wanted: An AI with common decency

- Fair — no biases
- Accountable — model can be explained
- Transparent — decisions can be explained
- Robust

ML & AI, 2

In practice

- Data are ridden with biases
- Learned models are biased (prejudices are transmissible to AI agents)
- Issues with robustness malicious users; evolution of environments
- Stakeholders with multi-faceted expectations

  Knowledge → Prediction → Control
- Models are used out of their scope

A. Comte	*Savoir pour prévoir, afin de pouvoir*
Zeynep Tufekci	We're building a dystopia just to make people click on ads
C. O'Neill	Weapons of Math Destruction
S.U. Noble	Algorithms of Oppression

Machine Learning: discriminative or generative modelling

Given a training set iid samples $\sim P(X, Y)$

$$\mathcal{E} = \{(\mathbf{x}_i, y_i), \mathbf{x}_i \in \mathbb{R}^d, i \in [[1, n]]\}$$

Find

- Supervised learning: $\hat{h} : X \mapsto Y$ or $\widehat{P}(Y|X)$
- Generative model $\widehat{P}(X, Y)$

Predictive modelling might be based on correlations

If umbrellas in the street, Then it rains

The big data promise:

ML models will expectedly support *interventions*:

- health and nutrition
- education
- economics/management
- climate

Intervention

Pearl 2009

Intervention $do(X = x)$ forces variable X to value x

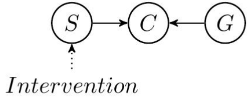
Direct cause $X_i \rightarrow X_j$

$$P_{X_j | \mathrm{do}(X_i = x, \mathbf{X}_{\setminus ij} = \mathbf{c})} \neq P_{X_j | \mathrm{do}(X_i = x', \mathbf{X}_{\setminus ij} = \mathbf{c})}$$

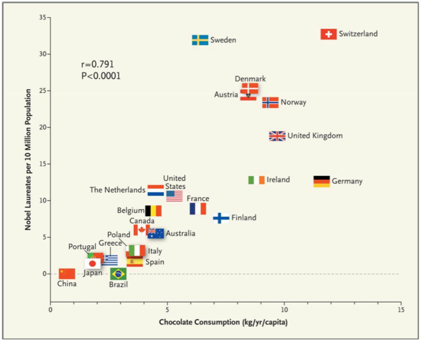
Example

C: Cancer, S : Smoking, G : Genetic factors

$P(C | do\{S = 0, G = 0\}) \neq P(C | do\{S = 1, G = 0\})$



Correlations do not support interventions



F. H. Messerli: *Chocolate Consumption, Cognitive Function, and Nobel Laureates*, N Engl J Med 2012

Causal models are needed to support interventions

An AI with common decency

Desired properties

- Fair — no biases
- Accountable — model can be explained
- Transparent — decisions can be explained
- Robust

Relevance of Causal Modeling

- Decreased sensitivity wrt data distribution
- Support interventions — clamping variable value
- Hopes of explanations

Motivation

Formal Background
The two-variable setting
The general setting

Causal Generative Neural Nets

Application to Human Resources
General Causal Relations
Category 1 (low-tech industry)
Category 2 (medium-low-tech industry)
Category 3 (medium-high-tech industry)
Category 4 (high-tech industry)

Discussion

Causal modelling, how

The royal road: randomized controlled experiments

Duflot Bannerjee 13; Imbens 15; Athey 15

But sometimes these are

- impossible — climate
- unethical — make people smoking
- too expensive — e.g., in economics

Machine Learning alternatives

- Observational data
- Statistical tests
- Learned models
- Prior knowledge / Assumptions / Constraints

Causal modelling, what

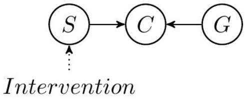
Definition 1, based on interventions Pearl 09, 18

A causes B if setting $A = 0$ yields a B distribution; and setting $A = 1$ ("everything else being equal") yields a different distribution.

$$P(B|\mathrm{do}(A = 1), C \ldots Z) \neq P(B|\mathrm{do}(A = 0), C \ldots Z)$$

Example C: Cancer, S : Smoking, G : Genetic factors

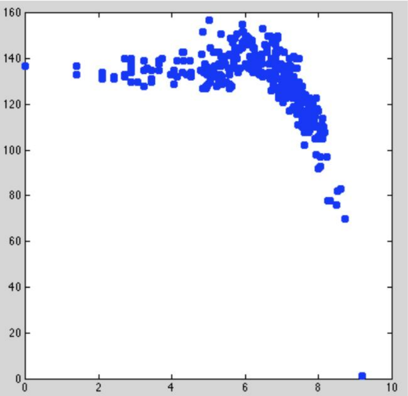
$P(C|do\{S = 0, G = 0\}) \neq P(C|do\{S = 1, G = 1\})$



Causal modeling

Definition 2: Granger causality
A "causes" B if knowing A helps predicting B

Given A, B



A

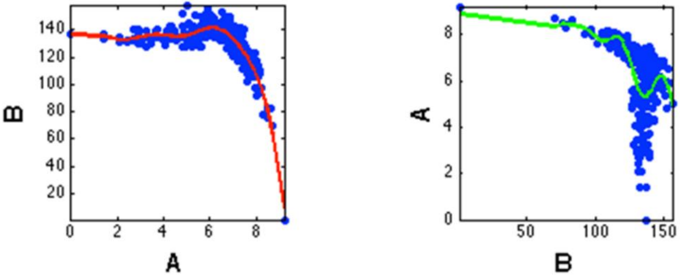
Granger causality

Given A, B, consider models

- $A = f(B)$
- $B = g(A)$

Compare the models

Select the best model: $A \to B$



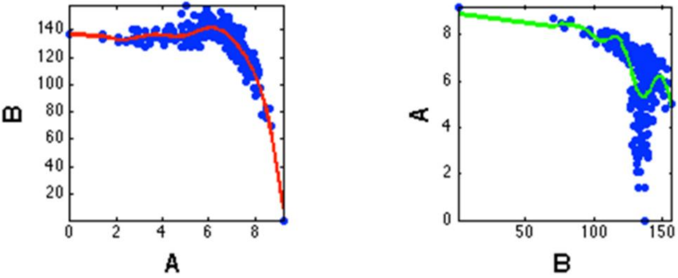
Granger causality

Given A, B, consider models

- $A = f(B)$
- $B = g(A)$

Compare the models

Select the best model: $A \to B$



A: Altitude, B: Temprature
Each point = (altitude, average temperature of a city)

Granger causality, 2

Appropriate to predictive modeling

- Chocolate consumption $\rightarrow$ Nobel prizes (developed countries)
- Umbrellas $\rightarrow$ Rain (effects are used to predict causes)

But inappropriate for prescriptions

- Making it rain by going out with an umbrella ?

Causality: A machine learning-based approach

Guyon et al, 2014-2015

Pair Cause-Effect Challenges

- Gather data: a sample is a pair of variables (A_i, B_i)
- Its label ℓ_i is the "true" causal relation (e.g., age "causes" salary)

Input

$$\mathcal{E} = \{(A_i, B_i, \ell_i), \ell_i \textit{ in } \{\rightarrow, \leftarrow, \perp\!\!\!\perp\}\}$$

Example A_i, B_i	Label ℓ_i
A_i causes B_i	$\rightarrow$
B_i causes A_i	$\leftarrow$
A_i and B_i are independent	$\perp\!\!\!\perp$

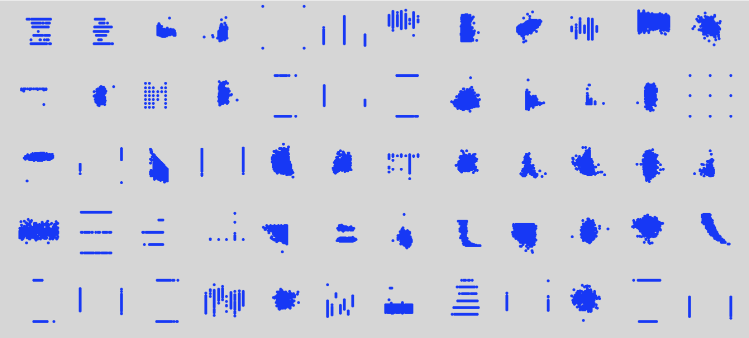
Output

using supervised Machine Learning

$$\text{Hypothesis} : (A, B) \mapsto \text{Label}$$

Causality: A machine learning-based approach, 3

Guyon et al, 2014-2015



The Cause-Effect Pair Challenge, 2

Learn a causality classifier

- Like for a mainstream supervised ML problem

ImageNet 2012

Motivation

Formal Background
- The two-variable setting
- The general setting

Causal Generative Neural Nets

Application to Human Resources
- General Causal Relations
- Category 1 (low-tech industry)
- Category 2 (medium-low-tech industry)
- Category 3 (medium-high-tech industry)
- Category 4 (high-tech industry)

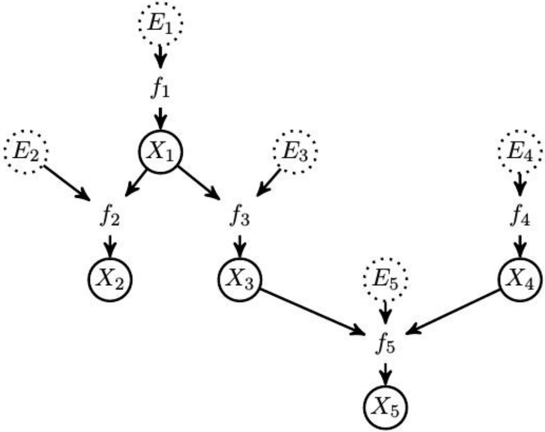
Discussion

Functional Causal Models, a.k.a. Structural Equation Models

Pearl 00-09

$$X_i = f_i(\mathrm{Pa}(X_i), E_i)$$

$\mathrm{Pa}(X_i)$: Direct causes for X_i E_i: noise variables, all unobserved influences



$$\begin{cases} X_1 = f_1(E_1) \\ X_2 = f_2(X_1, E_2) \\ X_3 = f_3(X_1, E_3) \\ X_4 = f_4(E_4) \\ X_5 = f_5(X_3, X_4, E_5) \end{cases}$$

Tasks

- Finding the structure of the graph (no cycles)
- Finding functions (f_i)

Conducting a causal modelling study

Spirtes et al. 01; Tsamardinos et al., 06; Hoyer et al. 09
Daniusis et al., 12; Mooij et al. 16

Milestones

- Testing bivariate independence (statistical tests)
  find edges $X - Y; Y - Z$
- Conditional independence
  prune the edges $X \perp\!\!\!\perp Z | Y$
- Full causal graph modelling
  orient the edges $X \rightarrow Y \rightarrow Z$

Challenges

- Computational complexity **tractable approximation**
- Conditional independence: data hungry tests
- Assuming causal sufficiency **can be relaxed**

$X - Y$ independance

$$P(X,Y) \stackrel{?}{=} P(X).P(Y)$$

Categorical variables

- Entropy $H(X) = -\sum_x p(x) log(p(x))$

  x: value taken by X, $p(x)$ its frequency
- Mutual information $M(X,Y) = H(X) + H(Y) - H(X,Y)$
- Others: χ^2, G-test

Continuous variables

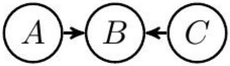
- t-test, z-test
- Hilbert-Schmidt Independence Criterion (HSIC) Gretton et al., 05

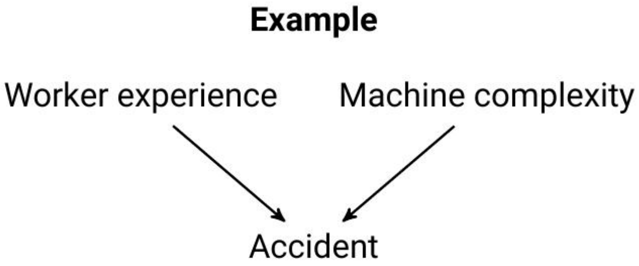
$$Cov(f,g) = \mathbb{E}_{x,y}[f(x)g(y)] - \mathbb{E}_x[f(x)]\mathbb{E}_y[g(y)]$$

- Given $f : X \mapsto \mathbb{R}$ and $g : Y \mapsto \mathbb{R}$
- $Cov(f,g) = 0$ for all f, g iff X and Y are independent

Find V-structure: $A \perp\!\!\!\perp C$ and $A \not\!\perp\!\!\!\perp C|B$

Explaining away causes





Motivation

Formal Background
- The two-variable setting
- The general setting

Causal Generative Neural Nets

Application to Human Resources
- General Causal Relations
- Category 1 (low-tech industry)
- Category 2 (medium-low-tech industry)
- Category 3 (medium-high-tech industry)
- Category 4 (high-tech industry)

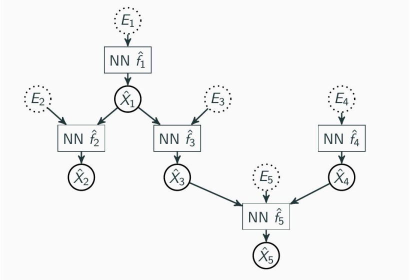
Discussion

Causal Generative Neural Network

Goudet et al. 17

Principle

- Given skeleton — given or extracted
- Given X_i and candidate $Pa(i)$
- Learn $f_i(\mathrm{Pa}(X_i), E_i)$ as a generative neural net
- Train and compare candidates based on scores



NB

- Can handle confounders (X_1 missing $\to$ ($E_2, E_3 \to E_{2,3}$))

Causal Generative Neural Network (2)

Training loss

- Observational data $\mathbf{x} = \{[x_1, \ldots, x_n]\}$ $\qquad x_i$ in $\mathbb{R} * *d$
- $(\mathrm{Graph}, \hat{f})$ $\hat{\mathbf{x}} = \{[\hat{x}_1, \ldots, \hat{x}_{n'}]\}$ $\qquad \hat{x}_i$ in $\mathbb{R} * *d$
- Loss: Maximum Mean Discrepancy $(\mathbf{x}, \hat{\mathbf{x}})$ (+ parsimony term), with k kernel (Gaussian, multi-bandwidth)

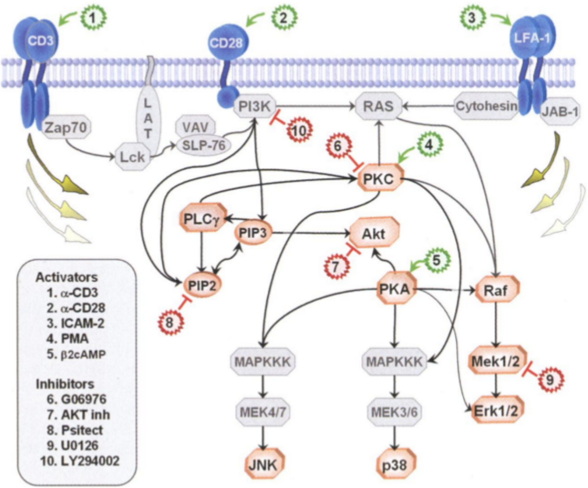
$$\mathrm{MMD}_k(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n^2} \sum_{i,j}^{n} k(x_i, x_j) + \frac{1}{n'^2} \sum_{i,j}^{n'} k(\hat{x}_i, \hat{x}_j) - \frac{2}{n \times n'} \sum_{i=1}^{n} \sum_{j=1}^{n'} k(x_i, \hat{x}_j)$$

- For $n, n' \to \infty$ $\qquad$ Gretton 07

$$\mathrm{MMD}_k(\mathbf{x}, \hat{\mathbf{x}}) = 0 \Rightarrow \mathcal{D}(\mathbf{x}) = \mathcal{D}(\hat{\mathbf{x}})$$

Results on real data: causal protein network

Sachs et al. 05



Edge orientation task

All algorithms start from the skeleton of the graph

method	AUPR	SHD	SID
Constraints			
PC-Gauss	0.19 (0.07)	16.4 (1.3)	91.9 (12.3)
PC-HSIC	0.18 (0.01)	17.1 (1.1)	90.8 (2.6)
Pairwise			
ANM	0.34 (0.05)	8.6 (1.3)	85.9 (10.1)
Jarfo	0.33 (0.02)	10.2 (0.8)	92.2 (5.2)
Score-based			
GES	0.26 (0.01)	12.1 (0.3)	92.3 (5.4)
LiNGAM	0.29 (0.03)	10.5 (0.8)	83.1 (4.8)
CAM	0.37 (0.10)	8.5 (2.2)	78.1 (10.3)
CGNN ($\widehat{\text{MMD}}_k$)	0.74* (0.09)	4.3* (1.6)	46.6* (12.4)

AUPR: Area under the Precision Recall Curve
SHD: Structural Hamming Distance
SID: Structural intervention distance

CGNN

Limitations

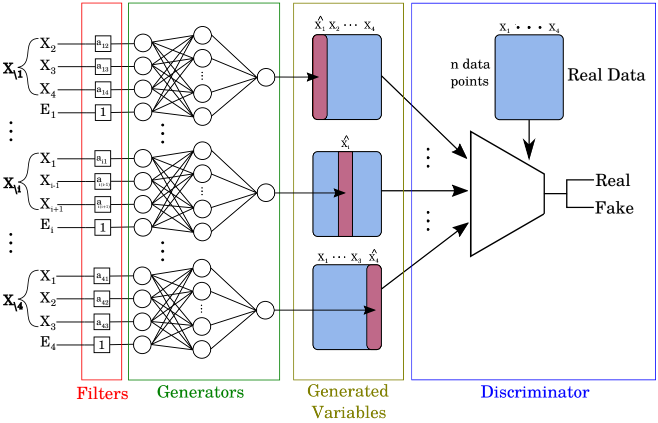
- Combinatorial search in the structure space
- Retraining fully the NN for each candidate graph
- MMD Loss is $O(n^2)$
- Limited to DAG

Structure Agnostic Modeling

Kalainathan et al. 18

- A dynamic process without self-loops:

$$X_i = f_i(X_{\setminus i}, E_i)$$



Structure Agnostic Modeling, 2

Ingredients for $f_i(X_{\setminus i}, E_i)$

- Scaling factors $a_{i,j}$ impact of X_j on X_i
- Dense layer with non-linear activation function
- Linear readout

$$\hat{X}_i = m_i^\top \tanh\left(\bar{W}_i^\top (a_i \odot X) + n_i E_i + b_i\right) + \beta_i$$

Structure Agnostic Modeling, 3

Loss function

- Adversarial learning

$$L_i = \mathbb{E}_{x_i, x_{\backslash i}}[\log D(x_i, x_{\backslash i})] + \mathbb{E}_{e_i, x_{\backslash i}}[\log(1 - D(\hat{f}_i(e_i, x_{\backslash i}), x_{\backslash i}))]$$

- $+$ Regularization   enforcing graph sparsity

$$L_\lambda = \sum_{i=1}^{d} L_i + \lambda \sum_{i=1}^{d} \|a_i\|_1 \quad , \lambda \geq 0$$

A competition between d sparse causal mechanisms $\hat{f}_i$ and a shared discriminator D.

Discussion

- No combinatorial search   scalability
- Cycles are possible: either genuine; or indicate non-identifiability

Quantitative benchmark - artificial DAG

Directed **acyclic** artificial graphs (DAG) of 20 variables

	PC Gauss	PC HSIC	GES	MMHC	DAGL1	LINGAM	CAM	SAM
Linear	0.36	0.29	0.40	0.36	0.30	0.31	0.29	**0.49**
Sigmoid AM	0.28	0.33	0.18	0.31	0.19	0.19	0.72	**0.73**
Sigmoid Mix	0.22	0.25	0.21	0.22	0.16	0.12	0.15	**0.52**
GP AM	0.21	0.35	0.19	0.21	0.15	0.17	**0.96**	0.74
GP Mix	0.22	0.34	0.18	0.22	0.19	0.14	0.61	**0.66**
Polynomial	0.27	0.31	0.20	0.11	0.26	0.32	0.47	**0.65**
NN	0.40	0.38	0.42	0.11	0.43	0.36	0.22	**0.60**
Execution time	1s	10h	$<$1s	$<$1s	2s	2s	2.5h	1.2h

Quantitative benchmark - artificial DG (with cycles)

Directed **cyclic** artificial graphs of 20 variables

	CCD	PC Gauss	GES	MMHC	DAGL1	LINGAM	CAM	SAM
Linear	0.44	0.44	0.20	0.34	0.26	0.19	0.23	**0.51**
Sigmoid AM	0.31	0.31	0.16	0.32	0.17	0.24	0.37	**0.47**
Sigmoid Mix	0.31	0.35	0.18	0.34	0.19	0.17	0.22	**0.49**
GP AM	0.30	0.32	0.17	0.30	0.15	0.23	0.50	**0.56**
GP Mix	0.24	0.25	0.15	0.24	0.16	0.18	0.26	**0.49**
Polynomial	0.25	0.33	0.20	0.25	0.17	0.22	0.33	**0.42**
NN	0.25	0.18	0.18	0.24	0.18	0.16	0.22	**0.40**
Execution time	1s	1s	$<$1s	$<$1s	2s	2s	2.5h	1.2h

Motivation

Formal Background
The two-variable setting
The general setting

Causal Generative Neural Nets

Application to Human Resources
General Causal Relations
Category 1 (low-tech industry)
Category 2 (medium-low-tech industry)
Category 3 (medium-high-tech industry)
Category 4 (high-tech industry)

Discussion

Causal Modeling and Human Resources

Position of the problem

A Quality of life at work — employee's perspective

B Economic performance — firm's perspective

- ... are correlated

Question: Are there causal relationships ?
$A \rightarrow B$; or $B \rightarrow A$; or $\exists C$ / $C \rightarrow A$ and $C \rightarrow B$

- Answering the question is key to evolve management strategies.

Data

- Gathered by Group Alpha Secafi (trade union advisor)
- Tax files + social audits for 408 firms

Firms

Category 1 255

- Chocolatier, Lesieur, Dassault, Compagnie des fromages, ...

Category 2 312

- ArcelorMittal, St Gobain, Lafarge, Vallourec, Michelin,...

Category 3 197

- Air Liquide, Thalès, Mersen, Filtrauto, Fenwick,...

Category 4 105

- Hispano-Suiza, TurboMéca, Sanofi, Snecma,...

Variables

Economics

- Total number of employees
- Capitalistic intensity, Total payroll, Gini index
- Average salary (of workers, technicians, managers)
- Productivity, Operating profits, Investment rate

People

- Average age, Average seniority, Physical effort,
- Permanent contract rate, Manager rate, Fixed-term contract rate, Temporary job rate, Shift and night work, Turn-over
- Vocational education effort, duration of stints, Average stint rate (for workers, technicians, managers);

Variables, cont'd

Quality of life at work

- Frequency & Gravity of work injuries, Safety expenses, Safety training expenses
- Absenteism (diseases), Occupational-related diseases
- Resignation rate, Termination rate, Participation rate
- Subsidy to the works council

Men/Women

- Percentage of women (employees, managers)
- Wage gap between women and men (average, for workers, technicians, managers)

General Causal Relations

General Causal Relations

Access to training $\nearrow$

- $\searrow$ Gravity of work injuries
- $\searrow$ Occupational-related diseases

Termination rate $\nearrow$

- $\nearrow$ Absenteism (diseases)

Percentage of managers $\nearrow$

- $\nearrow$ Access to training
- $\searrow$ Shift or night working hours

Age $\nearrow$

- $\searrow$ Fixed-term contract rate
- $\searrow$ Productivity (weak impact)

?

- Productivity $\nearrow \rightarrow$ Participation rate $\nearrow$

Global relations between QLW and performance ?

Failure

- Nothing conclusive

Interpretation

- Exist confounders (controlling QLW and performance) $C \to A$ and $C \to B$
- One such confounder is the activity sector
- In different activity sectors, causal relations are different (hampering their identification)
- $\Rightarrow$ Condition on confounders (independently handle the activity sectors)

Motivation

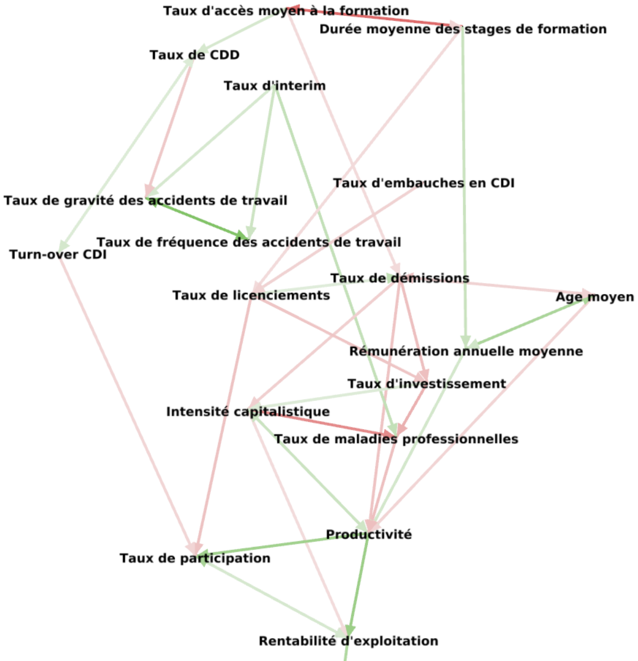
Formal Background
The two-variable setting
The general setting

Causal Generative Neural Nets

Application to Human Resources
General Causal Relations
Category 1 (low-tech industry)
Category 2 (medium-low-tech industry)
Category 3 (medium-high-tech industry)
Category 4 (high-tech industry)

Discussion

Low-tech industry



Category 1 (low-tech industry)

- Resignation rate $\nearrow$, Productivity $\searrow$
- Average salary $\nearrow$, Productivity $\nearrow$ very significant
- Occupational-related diseases $\nearrow$, Productivity $\searrow$
- Temporary job rate $\nearrow$, Gravity of work injuries $\nearrow$
- Permanent contract rate $\nearrow$, Safety training $\searrow$
- Duration training stints $\nearrow$, Termination rate $\searrow$

Motivation

Formal Background
The two-variable setting
The general setting

Causal Generative Neural Nets

Application to Human Resources
General Causal Relations
Category 1 (low-tech industry)
Category 2 (medium-low-tech industry)
Category 3 (medium-high-tech industry)
Category 4 (high-tech industry)

Discussion

Medium-low-tech industry

Category 2 (medium-low-tech industry)

- Permanent contract rate $\nearrow$, Productivity $\nearrow$
- Average salary $\nearrow$, rate of occupational related diseases $\searrow$
- Frequency of work injuries $\nearrow$, Termination rate $\nearrow$
- Percentage of women $\nearrow$, operating profitability $\nearrow$
- Temporary job rate $\nearrow$, Permanent contract rate $\nearrow$

Motivation

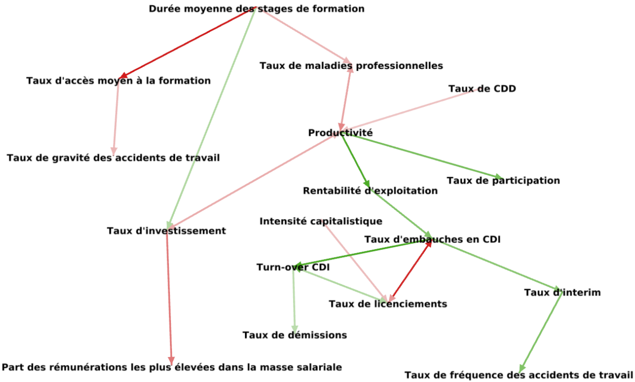
Formal Background
The two-variable setting
The general setting

Causal Generative Neural Nets

Application to Human Resources
General Causal Relations
Category 1 (low-tech industry)
Category 2 (medium-low-tech industry)
Category 3 (medium-high-tech industry)
Category 4 (high-tech industry)

Discussion

Medium-high-tech industry



Category 3 (medium-high-tech industry)

- Occupational-related diseases $\nearrow$, Productivity $\searrow$
- Temporary job rate $\nearrow$,
  - Frequency of work injuries $\nearrow$
  - Average salary $\searrow$
- Average training effort $\nearrow$, Permanent contract rate $\nearrow$
- Average salary workers $\nearrow$, Turn-over Permanent contract rate $\searrow$
- Average age $\nearrow$, Absenteism (diseases) $\nearrow$
- Average seniority $\nearrow$, Occupational-related diseases $\nearrow$
- Training of workers $\nearrow$, Productivity $\nearrow$

Motivation

Formal Background

- The two-variable setting
- The general setting

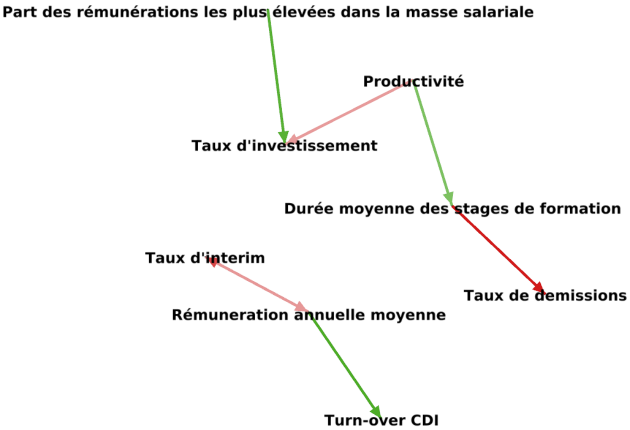
Causal Generative Neural Nets

Application to Human Resources

- General Causal Relations
- Category 1 (low-tech industry)
- Category 2 (medium-low-tech industry)
- Category 3 (medium-high-tech industry)
- Category 4 (high-tech industry)

Discussion

High-tech industry



Category 4 (high-tech industry)

- Manager rate $\nearrow$, Termination rate $\searrow$
- Total number of employees $\nearrow$, Training of managers $\nearrow$
- Occupational-related diseases $\nearrow$, Productivity $\searrow$

Motivation

Formal Background
The two-variable setting
The general setting

Causal Generative Neural Nets

Application to Human Resources
General Causal Relations
Category 1 (low-tech industry)
Category 2 (medium-low-tech industry)
Category 3 (medium-high-tech industry)
Category 4 (high-tech industry)

Discussion

Outcomes & Limitations

Causal modeling and exploratory analysis

- Efficient filtering of plausible relations (several orders of magnitude);
- Complementary w.r.t. visual inspection (experts can be fooled and make sense of correlations & hazards);
- Multi-factorial relations ? yes; but even harder to interpret.

Not a ready-made analysis

- Causal relations must be
  - interpreted
  - confirmed by field experiments; polls; interviews.

Perspectives

Spurious relations due to redundant variables

- E.g., in analytical accounting
- Variable selection
- Feature construction dimensionality reduction

Handling confounders

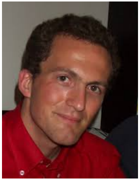
- Based on prior knowledge (e.g. using industry sectors)
- But many confounders are plausible (age and size of firm; company ownership and shareholdings; capitalistic intensity);
- Based on posterior analysis: finding latent dependency structures

Webb et al., 18

Thanks!

Diviyan Kalainathan

Olivier Goudet



Philippe Caillou

Isabelle Guyon

Paola Tubaro