



University of
Zurich^{UZH}

<http://r-bayesian-networks.org/>

GILLES KRATZER, PHD

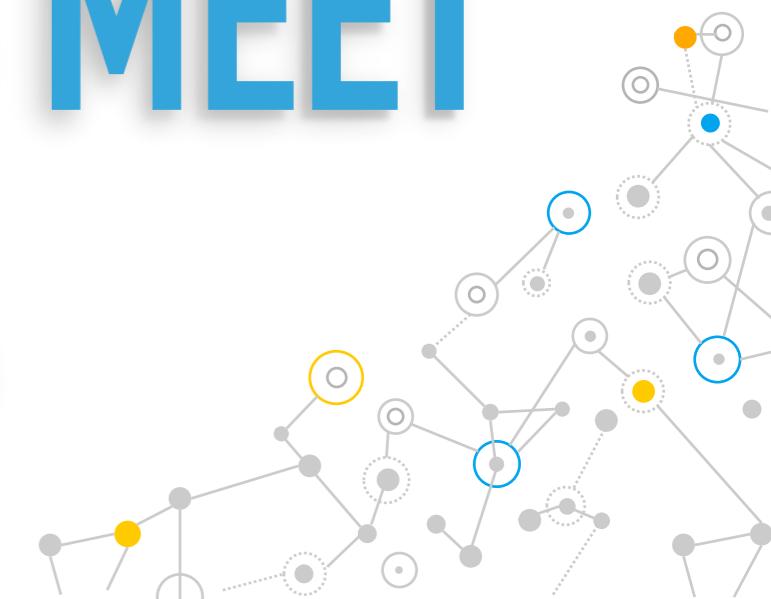
PROF. REINHARD FURRER

NICOLAS HUBER

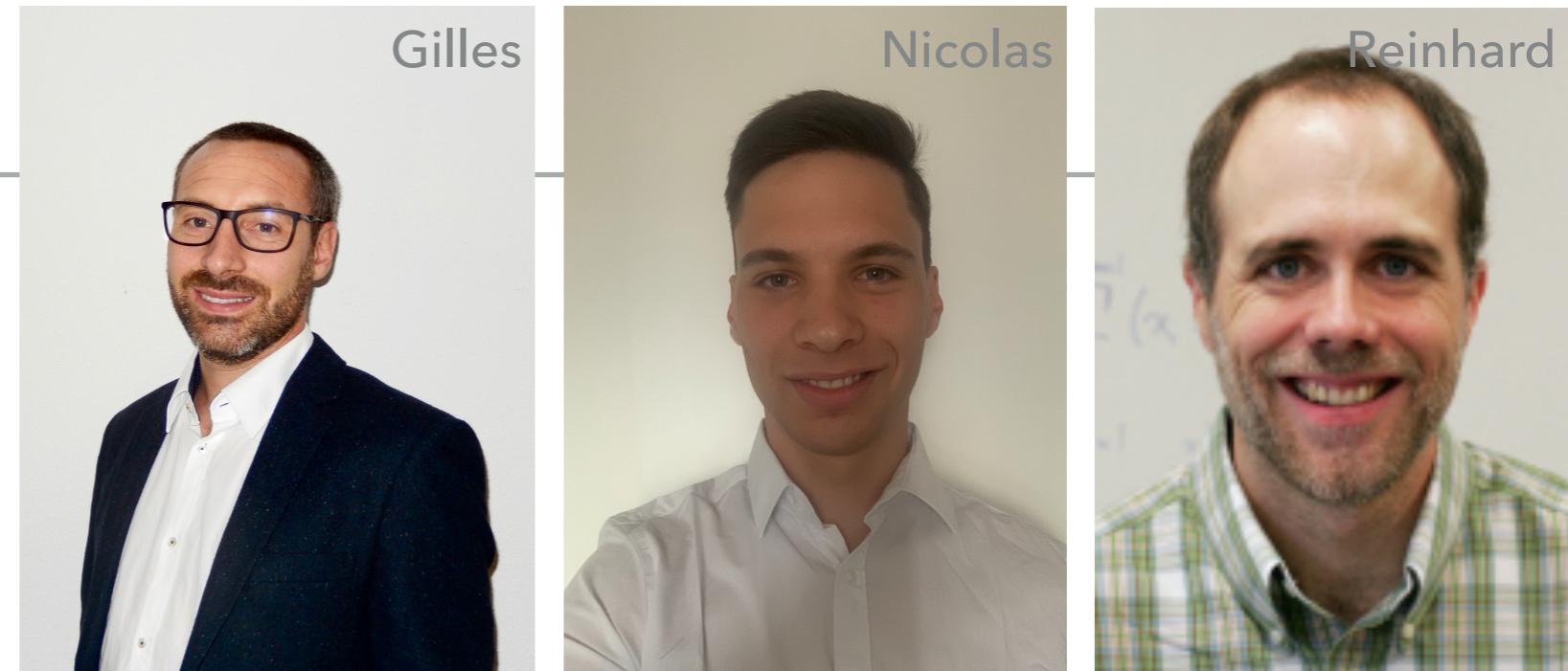
USER! CONFERENCE, ONLINE 07.07.2021

gilles.kratzer@gmail.com
reinhard.furrer@math.uzh.ch
nicolas.huber@uzh.ch

BAYESIAN NETWORKS MEET OBSERVATIONAL DATA



SCHEDULE



11:15 

| Brief introduction on Additive Bayesian modelling

12:00 

| Hands-on exercise: first analysis

13:00 

| More advanced features of Additive Bayesian modelling

13:20 

| Hands-on exercise: advanced features

13:40

| Wrap-up and discussion

13:45

MATERIAL

Material for the workshop

<https://gilleskratzer.github.io/ABN-UseR-2021/>

The screenshot shows the ABN-UseR-2021 workshop website. At the top, there's a blue header with the title "ABN-UseR-2021" and "ABN workshop - UseR! Conference". Below the header is a yellow navigation bar with a "View On GitHub" button and a GitHub icon. The main content area has a light beige background. On the left, there's a sidebar with sections for "General information", "Learning outcomes", and "Schedule (Zurich time)". The "Schedule (Zurich time)" section contains a table with five rows, each representing a time slot from 11:15 to 13:45, a topic, and a material link.

Time	Topic	Material
11:15 – 12:00	Brief theoretical introduction on Additive Bayesian modelling	Presentation
12:00 – 13:00	Hands-on exercises	Presentation / Hands-on exercises / Zip folder
13:00 – 13:20	Advanced ABN modelling	Presentation
13:20 – 13:40	Hands-on exercises on advanced features	Hands-on exercises / Zip folder
13:40 – 13:45	Wrap-up and discussions	

More ressources about ABN

<http://r-bayesian-networks.org/>

The screenshot shows the r-bayesian-networks.org website. At the top, there's a blue header with the title "Additive Bayesian Network Modelling in R" and a subtitle "Bayesian network analysis is a form of probabilistic graphical models which derives from empirical data a directed acyclic graph (DAG)". Below the header is a yellow navigation bar with a "View On GitHub" button and a GitHub icon. The main content area has a light beige background. On the left, there's a sidebar with sections for "Introduction", "Installation", "Quickstart", "Case studies", "Literature", and "Further resources". The "Introduction" section contains a large image of a complex Bayesian network graph with many nodes and arrows. Below the graph, there's a download statistics box ("downloads 693/month CRAN 2.5-0 – 2021-04-23") and a detailed description of what Bayesian network modelling is. The footer includes credits for "Project maintained by gilleskratzer" and "Hosted on GitHub Pages – Theme by mattgraham".

HANDS-ON

<https://gilleskratzer.github.io/ABN-UseR-2021/>

The screenshot shows the ABN-UseR-2021 workshop website. On the left sidebar, there are links for General information, Learning outcomes, and Schedule (Zurich time). The main content area has a yellow header bar with a 'View On GitHub' button and a GitHub logo. Below the header, the title 'ABN-UseR-2021' and subtitle 'ABN workshop - UseR! Conference' are displayed. The 'Schedule (Zurich time)' section lists five sessions with their times, topics, and associated materials:

Time	Topic	Material
11:15 - 12:00	Brief theoretical introduction on Additive Bayesian modelling	Presentation
12:00 - 13:00	Hands-on exercises	Presentation / Hands-on exercises / Zip folder
13:00 - 13:20	Advanced ABN modelling	Presentation
13:20 - 13:40	Hands-on exercises on advanced features	Hands-on exercises / Zip folder
13:40 - 13:45	Wrap-up and discussions	

<https://app.slack.com/>

The screenshot shows the Slack interface for the '#tut_abn' channel. The channel has 15 members. A message from 'useR! 2021 - The L...' at 12:04 AM on July 4th says: 'This is the very beginning of the #tut_abn channel @useR! 2021 global created this channel on July 4th. Additive Bayesian Networks Modeling, Gilles Kratzer, Reinhard Furrer'. Below this, another message from 'useR! 2021 global' at 12:10 AM on July 4th says: 'joined #tut_abn along with 4 others.' The bottom of the screen shows a red 'ROOM 7' button and various Slack interface icons.

MOTIVATIONAL EXAMPLE: CREDIT CARD FRAUD DETECTION PREDICTION

Credit Card Fraud Detection Using Bayesian and Neural Networks

Sam Maes

Karl Tuyls

Bram Vanschoenwinkel

Bernard Manderick

Vrije Universiteit Brussel - Department of Computer Science

Computational Modeling Lab (COMO)

Pleinlaan 2

B-1050 Brussel, Belgium

{sammaes@vub.ac.be,ktuyls@vub.ac.be,bvschoen@vub.ac.be,bernard@arti.vub.ac.be}

Abstract

This paper discusses automated credit card fraud detection by means of machine learning. In an era of digitalization, credit card fraud detection is of great importance to financial institutions. We apply two machine learning techniques suited for reasoning under uncertainty: artificial neural networks and

do the fraud detection. After a process of learning, the program is supposed to be able to correctly classify a transaction it has never seen before as fraudulent or not fraudulent, given some features of that transaction.

The structure of this paper is as follows: first we introduce the reader to the domain of credit card fraud detection. In Sections 3 and 4 we briefly ex-

Credit Card Fraud Detection Using Bayesian and Neural Networks

Sam Maes

Karl Tuyls

Bram Vanschoenwinkel

experiment	$\pm 10\%$ false pos	$\pm 15\%$ false pos
ANN-fig 2(a)	60% true pos	70% true pos
ANN-fig 2(a)	47% true pos	58% true pos
ANN-fig 2(c)	60% true pos	70% true pos
BBN-fig 2(e)	68% true pos	74% true pos
BBN-fig 2(g)	68% true pos	74% true pos

Abstract

This paper discusses credit card fraud detection by means of machine learning. The process of digitalization, creation of databases and the great importance to society of correctly classifying transactions as either normal or fraudulent are discussed.

Table 1: This table compares the results achieved with ANN and BBN, for a false positive rate of respectively 10% and 15%.

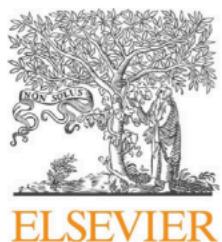
two machine learning techniques suited for reasoning under uncertainty: artificial neural networks and

introduce the reader to the domain of credit card fraud detection. In Sections 3 and 4 we briefly ex-

rocess of learning,
e to correctly clas-
n before as fraud-
ne features of that

s follows: first we

MOTIVATIONAL EXAMPLE: VETERINARY EPIDEMIOLOGY DATA VISUALISATION



Contents lists available at SciVerse ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed



Using Bayesian networks to explore the role of weather as a potential determinant of disease in pigs

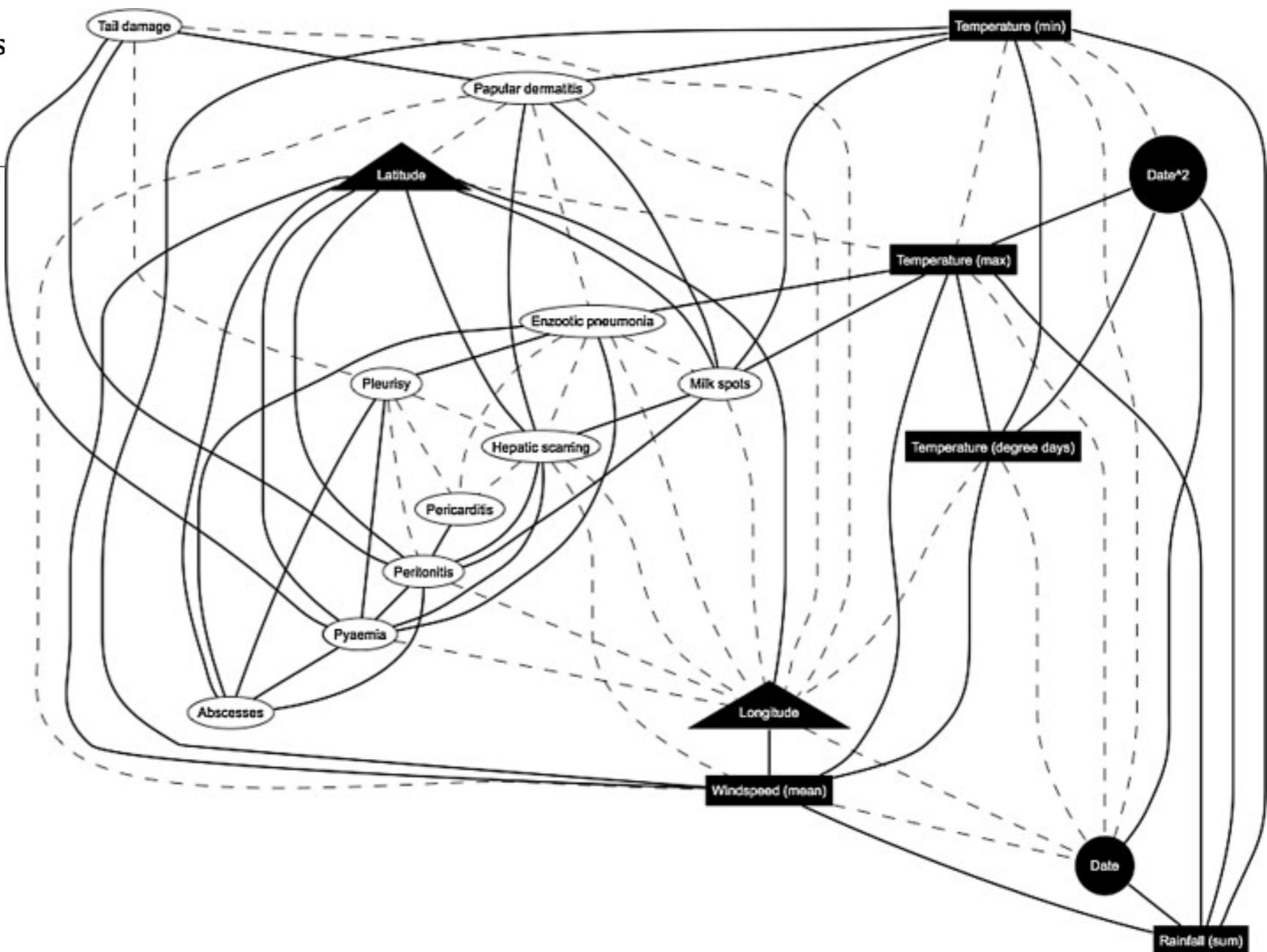


B.J.J. McCormick^a, M.J. Sanchez-Vazquez^b, F.I. Lewis

^a Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

^b OIE Organisation Mondiale de la Santé Animale, 12, rue de Prony, 75017 Paris, France

^c Section of Epidemiology, University of Zurich, Zurich, Switzerland



MOTIVATIONAL EXAMPLE: SOCIAL SCIENCES DATA INTERPRETATION

Discovering complex interrelationships between socioeconomic status and health in Europe: A case study applying Bayesian Networks

Javier Alvarez-Galvez ^{a, b, *}

^a Loyola University Andalusia, Department of International Studies, Campus de Palmas Altas, Faculty of Political Sciences and Law, Seville 41014, Spain

^b Complutense University of Madrid, Department of Sociology IV (Research Methodology and Communication Theory), Campus de Somosaguas, Faculty of Political

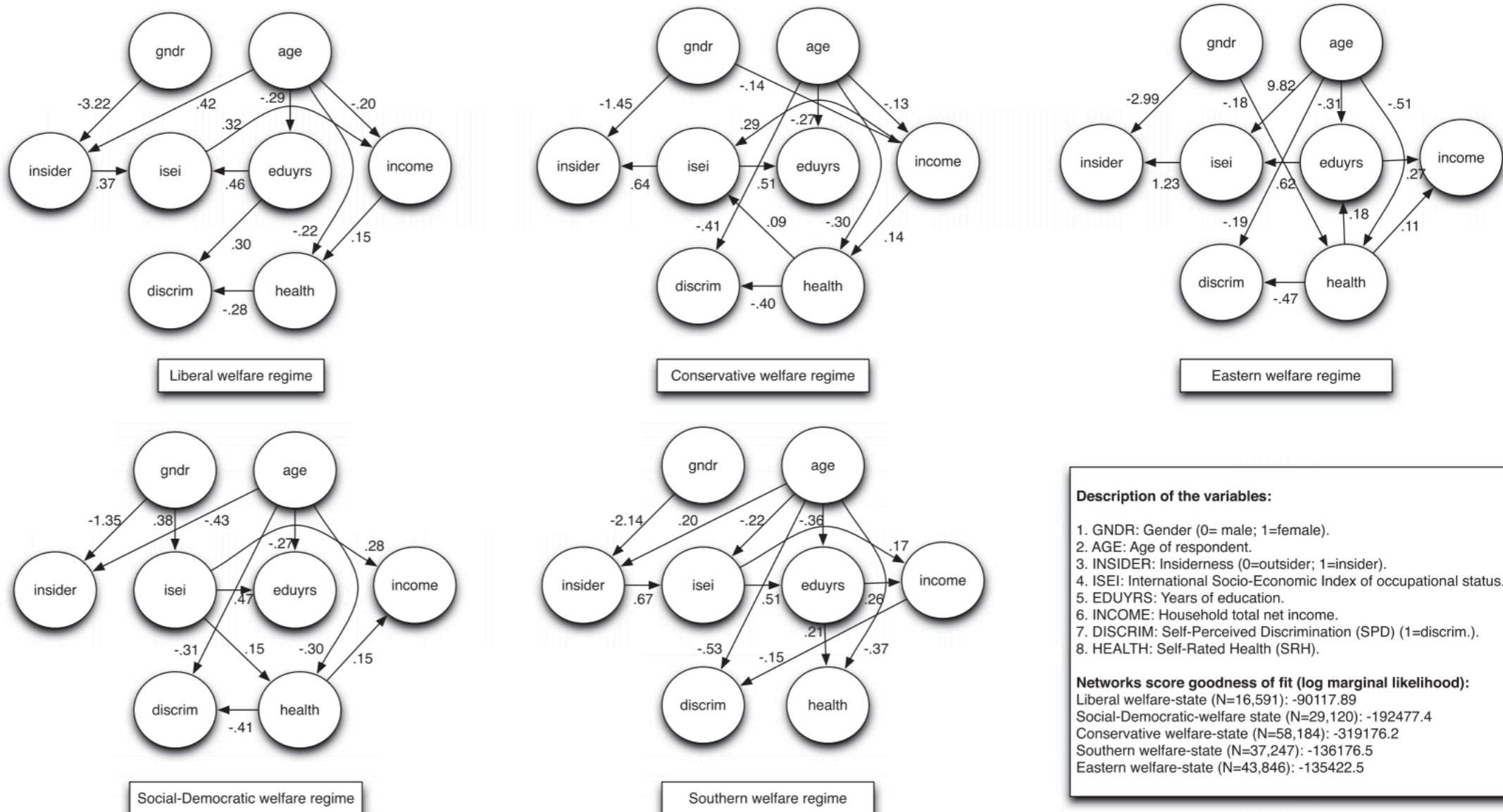
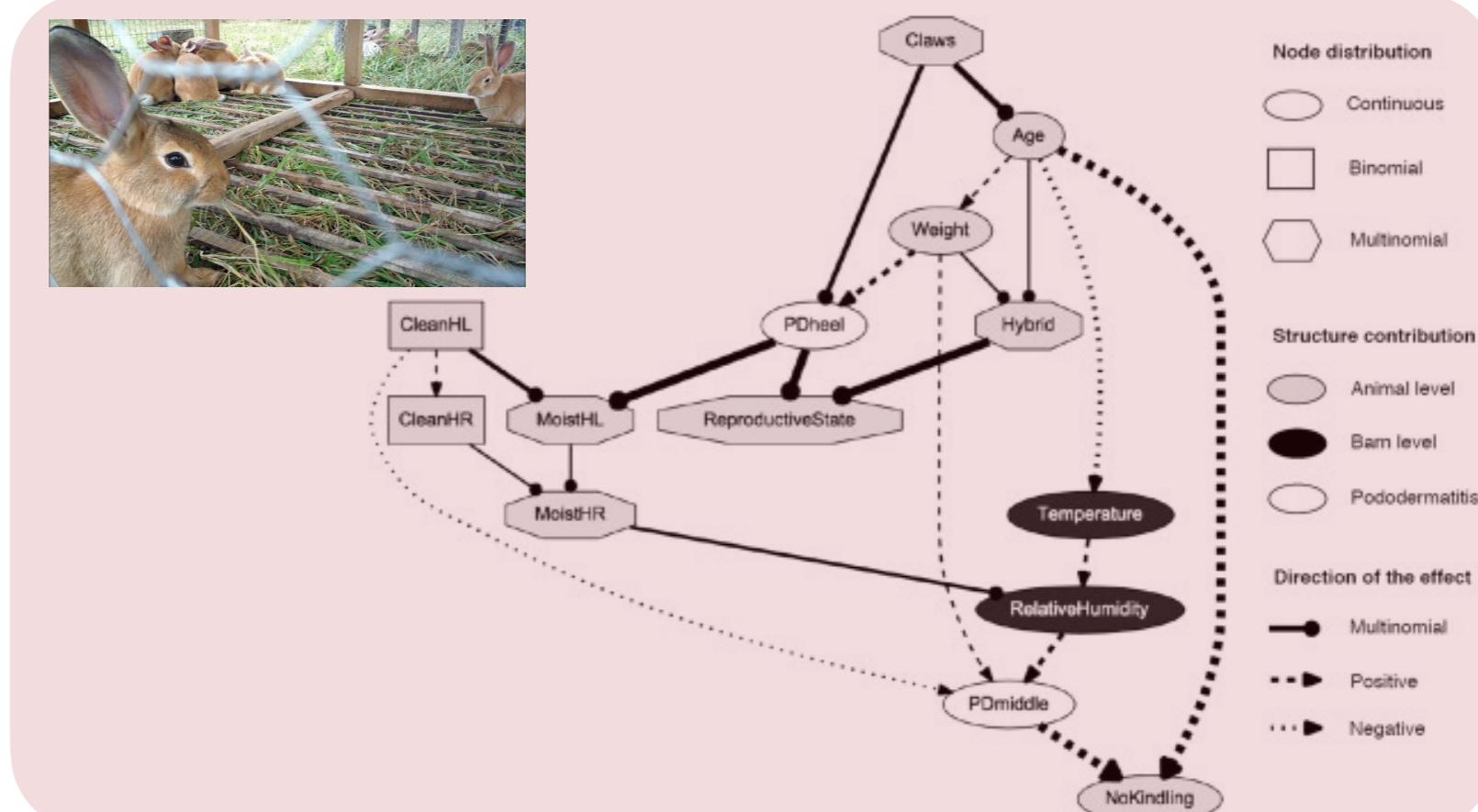
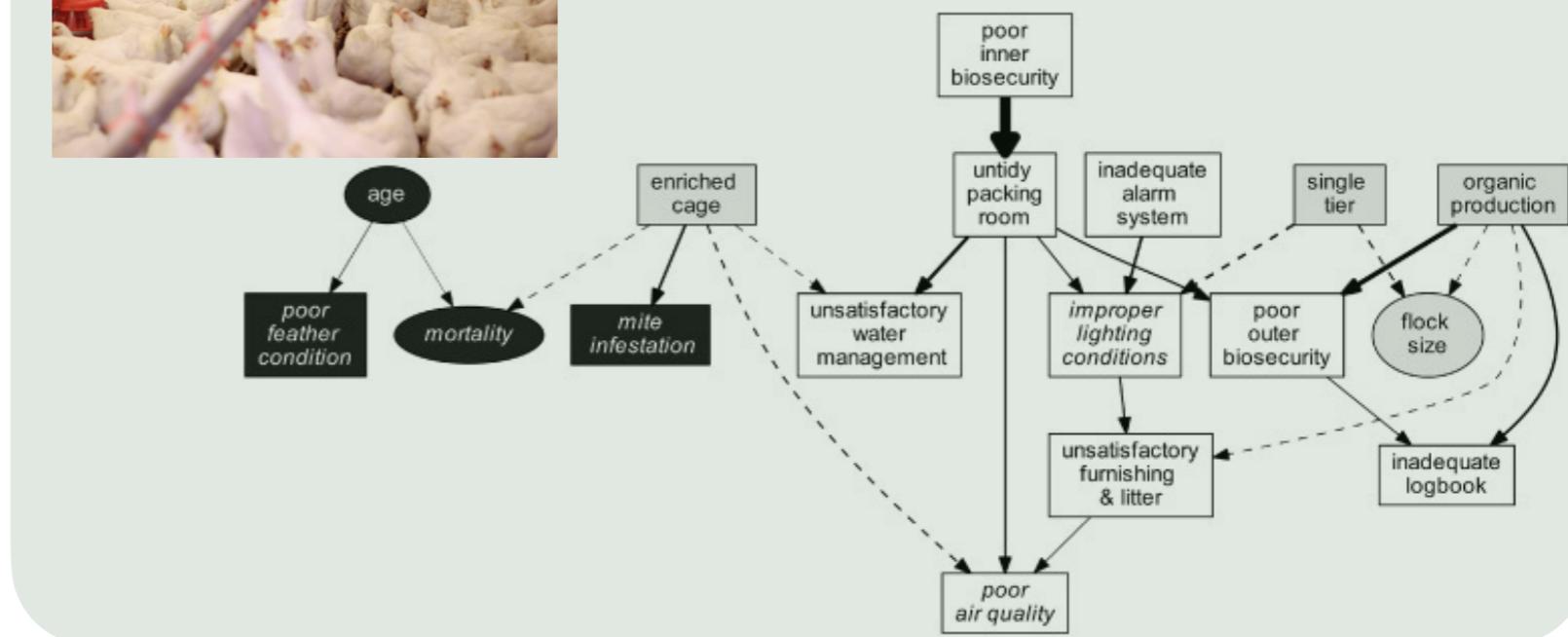
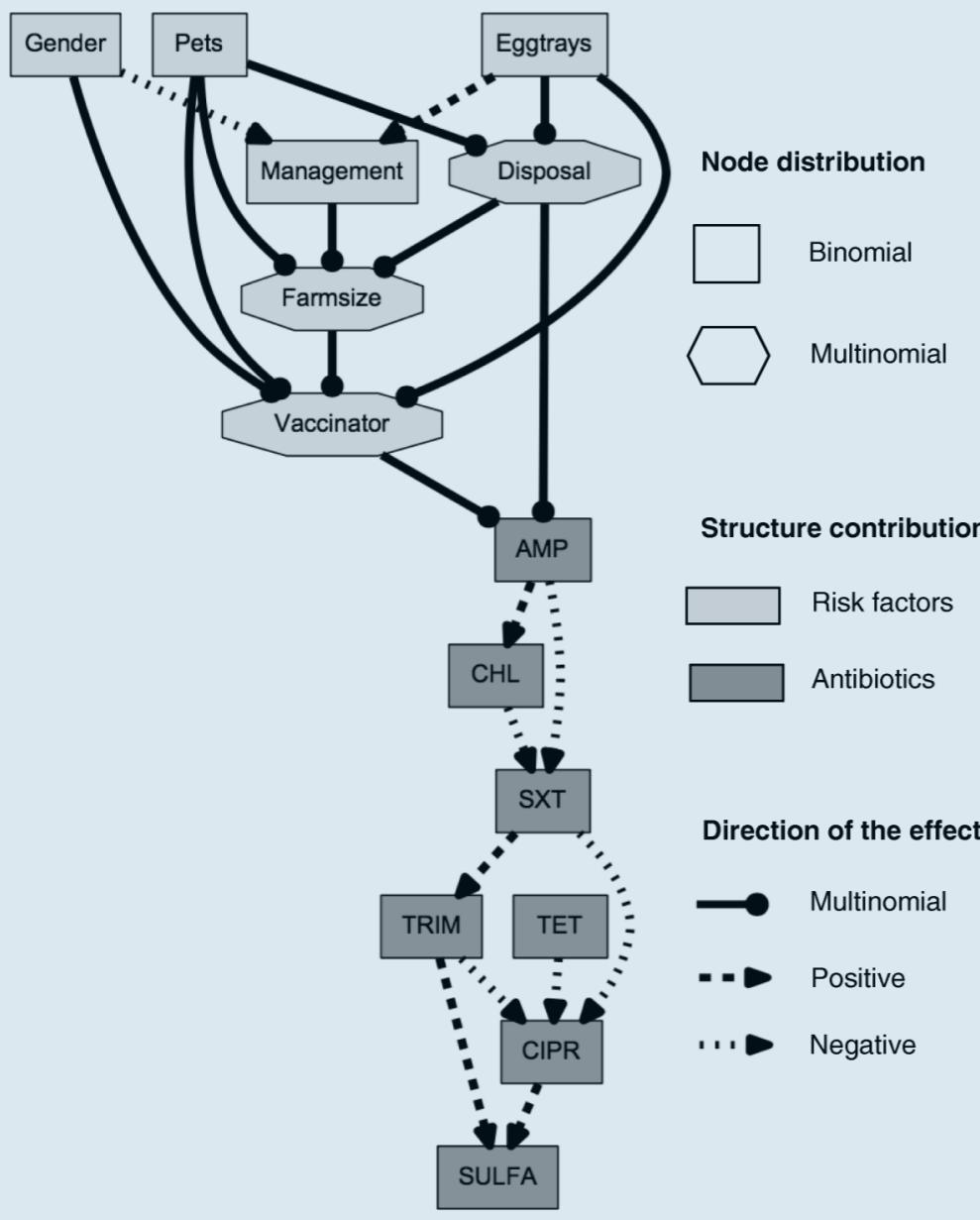


Fig. 1. Bayesian networks describing interrelationships between SES and health in five European welfare states.

EXAMPLE OF SYSTEMS EPIDEMIOLOGY DATA ANALYSED WITH ABN



EXAMPLE OF SYSTEMS EPIDEMIOLOGY DATA ANALYSED WITH ABN

Anti-microbial resistance



- ▶ Multi-drug resistant *Salmonella* isolates (7 antibiotics)
- ▶ 43 poultry farms in Uganda
- ▶ Risk factors: Management practice, farm size, etc ...

MULTIPLE OUTCOMES

Hartnack and al. (2019) in BMC

Animal welfare



- ▶ Welfare control programme after ban of battery cage
- ▶ 193 different poultry farms in Sweden
- ▶ Welfare status depends on many inter-related variables
- ▶ Risk factors: Management practice, weather, etc ...

MULTIDIMENSIONAL

Comin and al. (2019) in PVM

Technopathy in rabbit

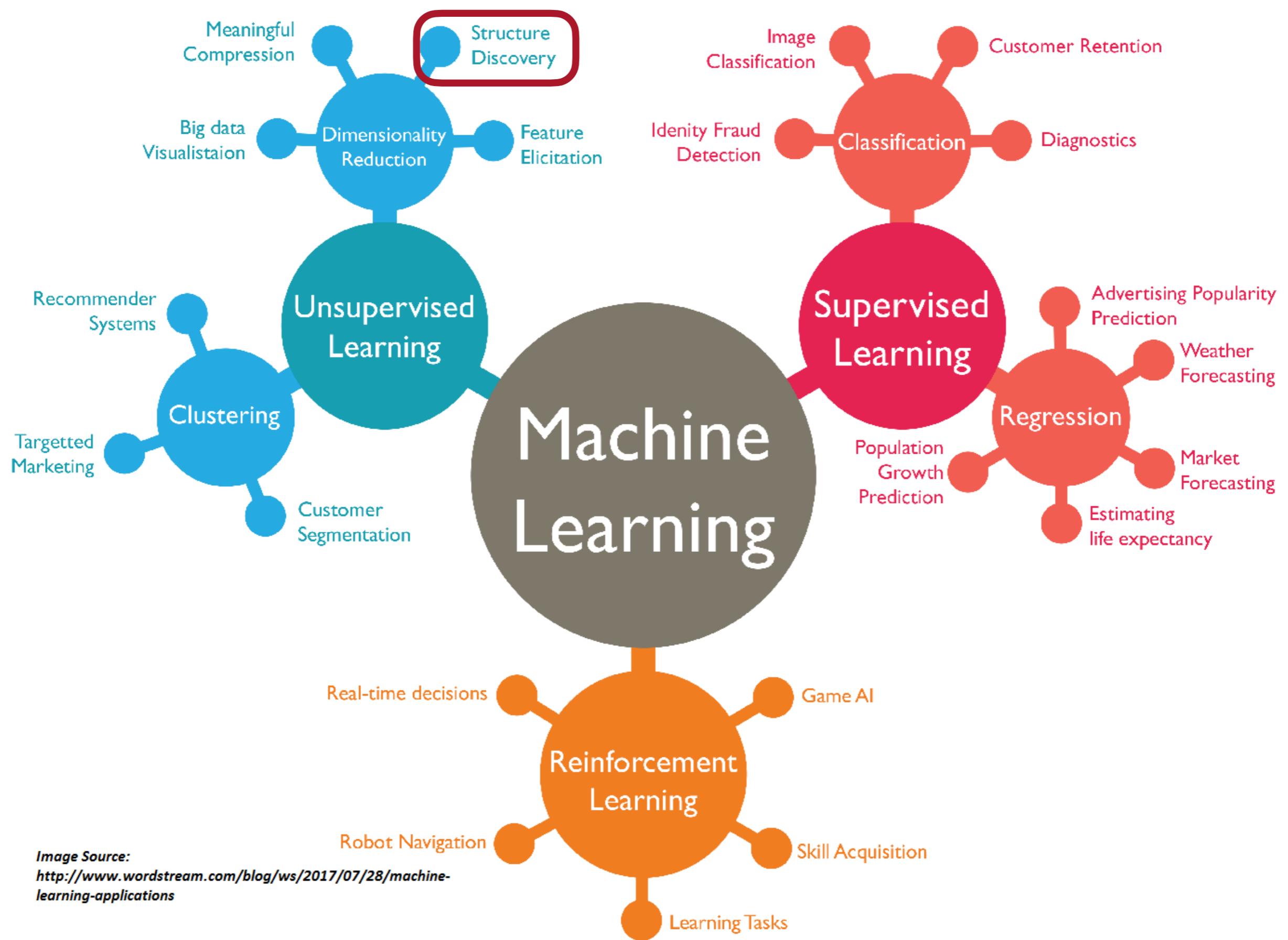


- ▶ Longitudinal study on Pododermatitis in rabbits
- ▶ 3 commercial farms in Switzerland
- ▶ Group housing on litter and plastic slats
- ▶ Main interest: Healing process

HYPOTHESIS GENERATION

Ruchti and al. (2019) in PVM

BAYESIAN NETWORKS IN THE MACHINE LEARNING WORLD



OUTLINE OF THE TALK

Objective of the workshop:

How to learn Bayesian networks from observational data?

OUTLINE OF THE TALK

Objectif of the workshop:

select

How to ~~learn~~ Bayesian networks from observational data?

Bayesian Networks are defined by two elements:

Network structure:

Directed Acyclic Graph (**DAG**): $G = (V, A)$

in which each node $v_i \in V$ corresponds to a random variable X_i

Probability distribution:

Probability distribution X with parameters Θ , which can be factorised into smaller local probability distributions according to the arcs $a_{ij} \in A$ present in the graph.

A BN encodes the factorisation of the joint distribution

$$P(\mathbf{X}) = \prod_{j=1}^n P(X_j | \mathbf{Pa}_j, \Theta_j), \text{ where } \mathbf{Pa}_j \text{ is the set of parents of } X_j$$

ABN WORKFLOW

1. From observational dataset deduce probabilistic model
Epidemiological constrain: mixture of distributions
2. From probabilistic model deduce structure

EXPONENTIAL FAMILY

Observational dataset



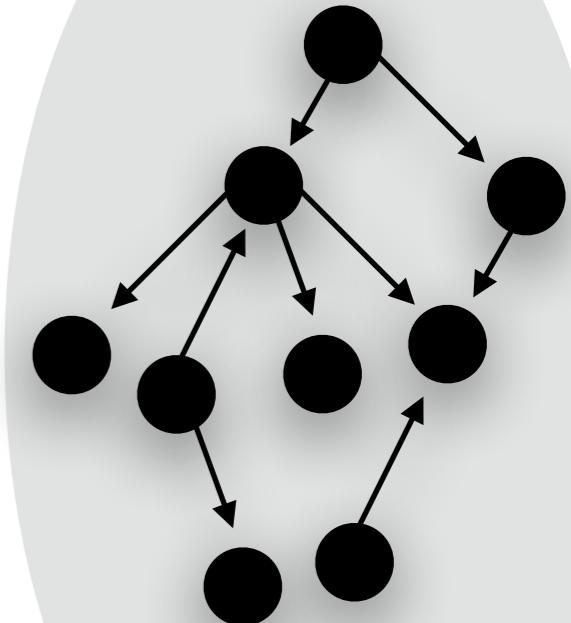
Probabilistic model

$$P(\mathbf{X}) = \prod_{j=1}^n P(X_j | \text{Pa}_j, \Theta_j)$$



Computing directly

Network structure



COMBINATORIAL WALL

# Nodes	# DAGs	Inference
1 - 15 Nodes	$< 10^{41}$ DAGs	Exact inference
16 - 25 Nodes	$< 10^{100}$ DAGs	Exact inference possible
26 - 50 Nodes	$< 10^{400}$ DAGs	Approximate inference
51 - 100 Nodes	$< 10^{1700}$ DAGs	Approximate inference
101 - 1000 Nodes	$< 10^{100000}$ DAGs	(very) approximative inference

Approximations:

- ▶ limiting number of parents per node
- ▶ Decomposable scores/efficient algorithm
- ▶ Score equivalence

SOME ELEMENTS OF PROBABILITY THEORY

- 1** The **conditional probability** of A given B is: $P(A | B) = \frac{P(A, B)}{P(B)}$

- 2** Bayes theorem: $P(A | B) = \frac{P(B | A)P(A)}{P(B)}$

SOME ELEMENTS OF PROBABILITY THEORY

1 The **conditional probability** of A given B is: $P(A | B) = \frac{P(A, B)}{P(B)}$

2 Bayes theorem: $P(A | B) = \frac{P(B | A)P(A)}{P(B)}$

Let A, B and C non intersecting subsets of nodes in a DAG G

1 + 2

A is **conditionally independent** of B given C if: $P(A, B | C) = P(A | C)P(B | C)$

SOME ELEMENTS OF PROBABILITY THEORY

- 1 The **conditional probability** of A given B is: $P(A | B) = \frac{P(A, B)}{P(B)}$
- 2 Bayes theorem: $P(A | B) = \frac{P(B | A)P(A)}{P(B)}$

Let A, B and C non intersecting subsets of nodes in a DAG G

1 + 2

A is **conditionally independent** of B given C if: $P(A, B | C) = P(A | C)P(B | C)$



Theorem (Verma & Pearl, 1988): A is d-separated from B by C if, and only if, the joint distribution over all variables in the graph satisfies:

$$A \perp\!\!\!\perp B | C$$

Link between statistical statement (**conditionally independence**) and a graph propriety (**d-separation**)

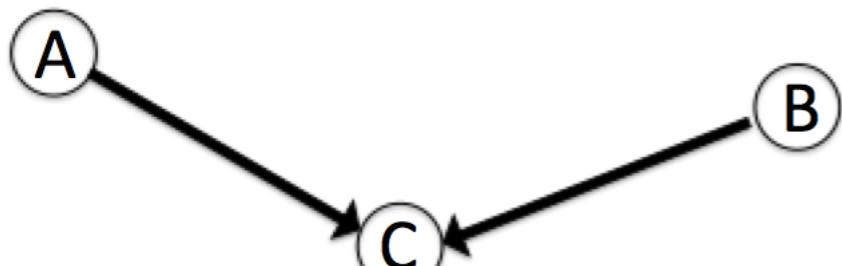
ELEMENT OF GRAPH THEORY

Let A, B and C non intersecting subsets of nodes in a DAG G

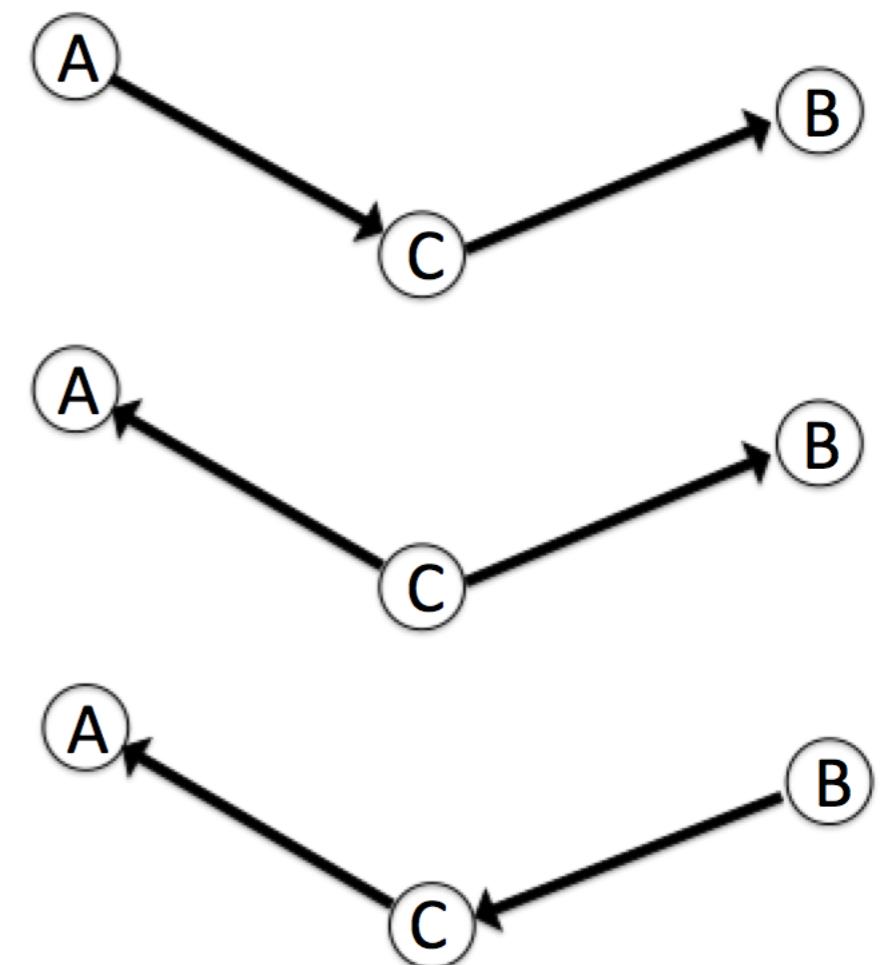
A is **conditionally independent** of B given C if: $A \perp\!\!\!\perp_B | C$

$$P(A, B | C) = P(A | C)P(B | C)$$

$A \not\perp\!\!\!\perp_B | C$



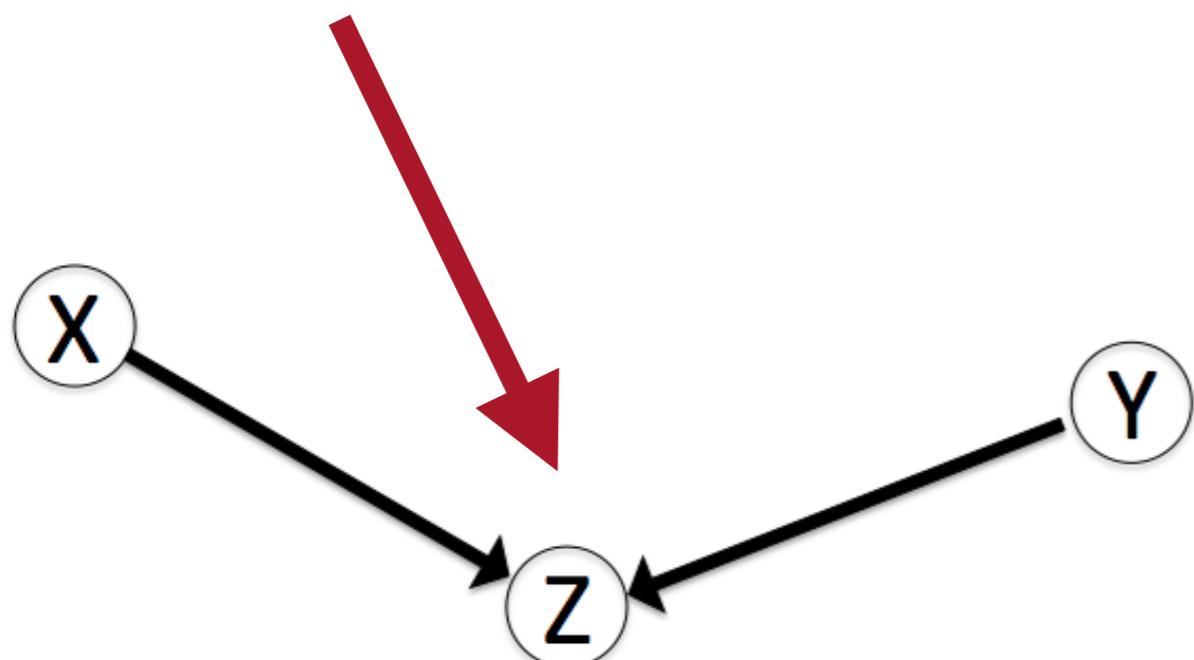
$A \perp\!\!\!\perp_B | C$



Constraint based algorithms

Learning independence relationships

$$P_{X \perp\!\!\!\perp Y|Z} < \alpha$$



Search-and-score algorithms

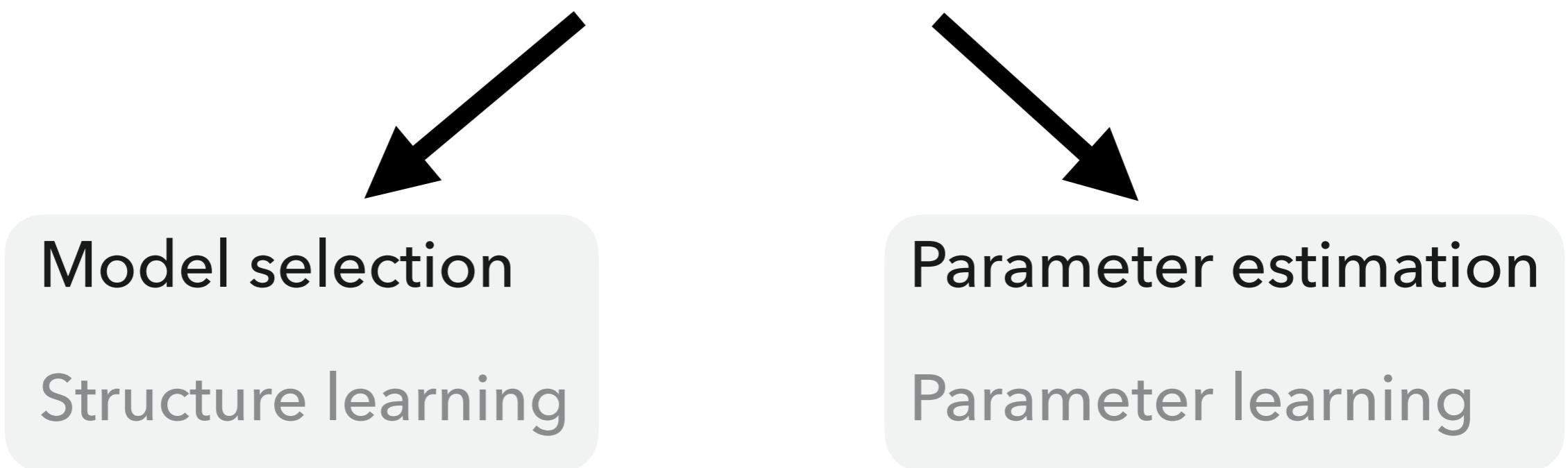
Maximum a posteriori score

Example of scoring functions:

- ▶ Bayesian versus ML scores
 - ▶ log marginal likelihood
 - ▶ Bayesian-Dirichlet (BDeu,BDs,BDe)
 - ▶ Bayesian Information Criterion (BIC)

LEARNING BAYESIAN NETWORKS VIA SEARCH-AND-SCORE ALGORITHMS

$$\mathcal{M} = (\mathcal{S}, \Theta_{\mathcal{M}})$$

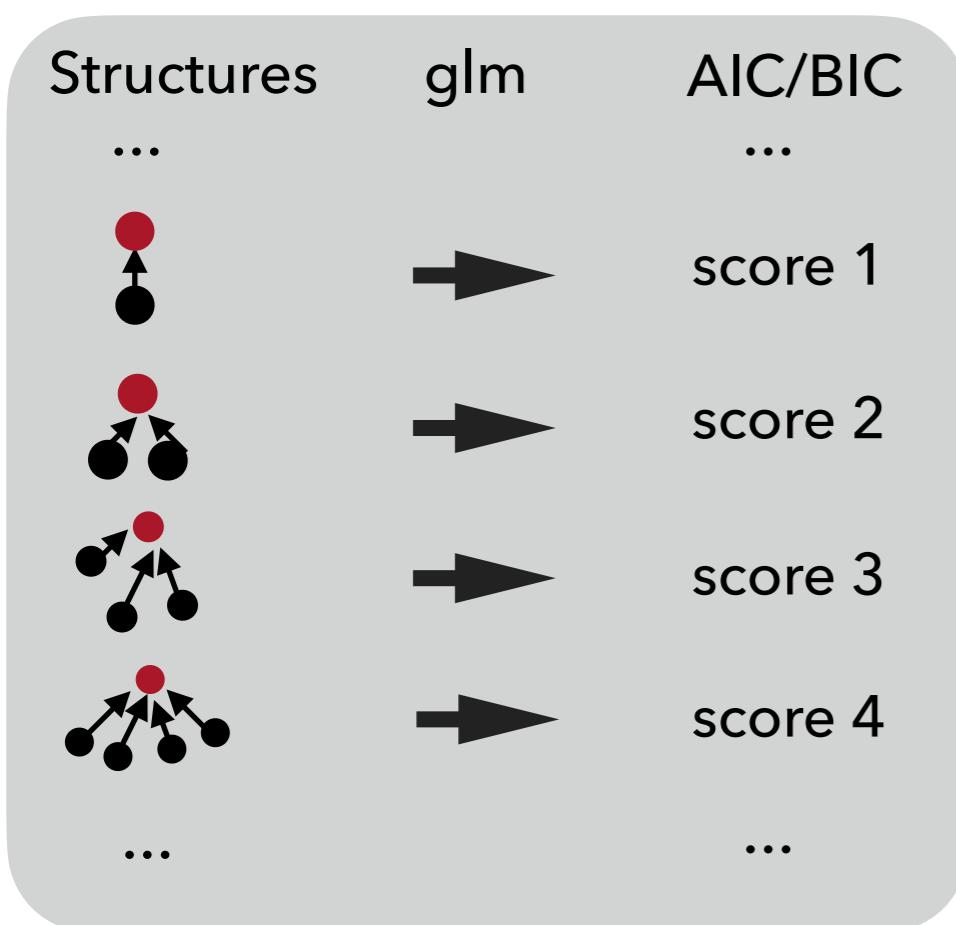


$$P(\mathcal{M}|\mathcal{D}) = \underbrace{P(\Theta_{\mathcal{M}}, \mathcal{S}|\mathcal{D})}_{\text{model learning}} = \underbrace{P(\Theta_{\mathcal{M}}|\mathcal{S}, \mathcal{D})}_{\text{parameter learning}} \cdot \underbrace{P(\mathcal{S}|\mathcal{D})}_{\text{structure learning}}$$

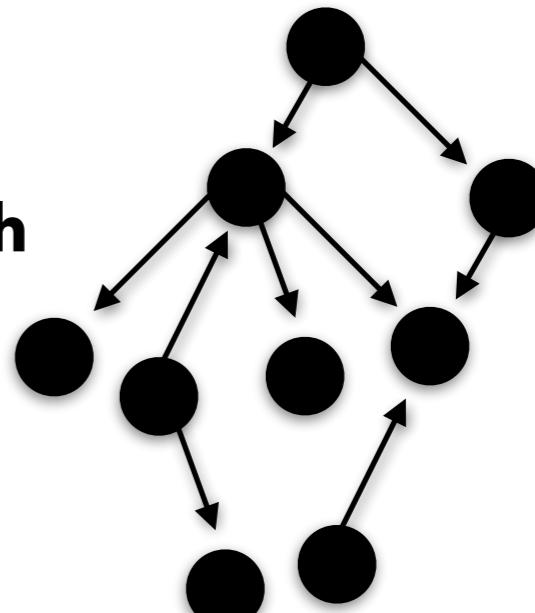
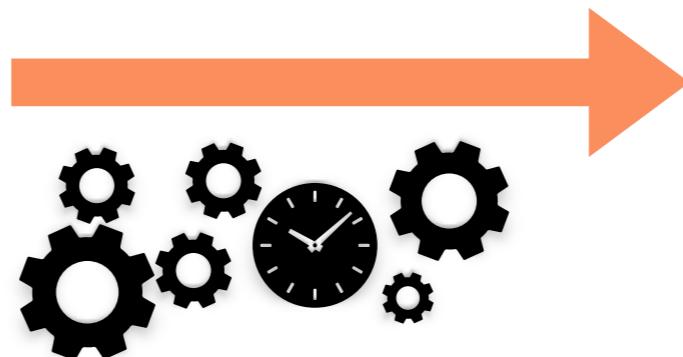
DEEP DIVE INTO ABN

From now on ... ABN specific

Search and score algorithm

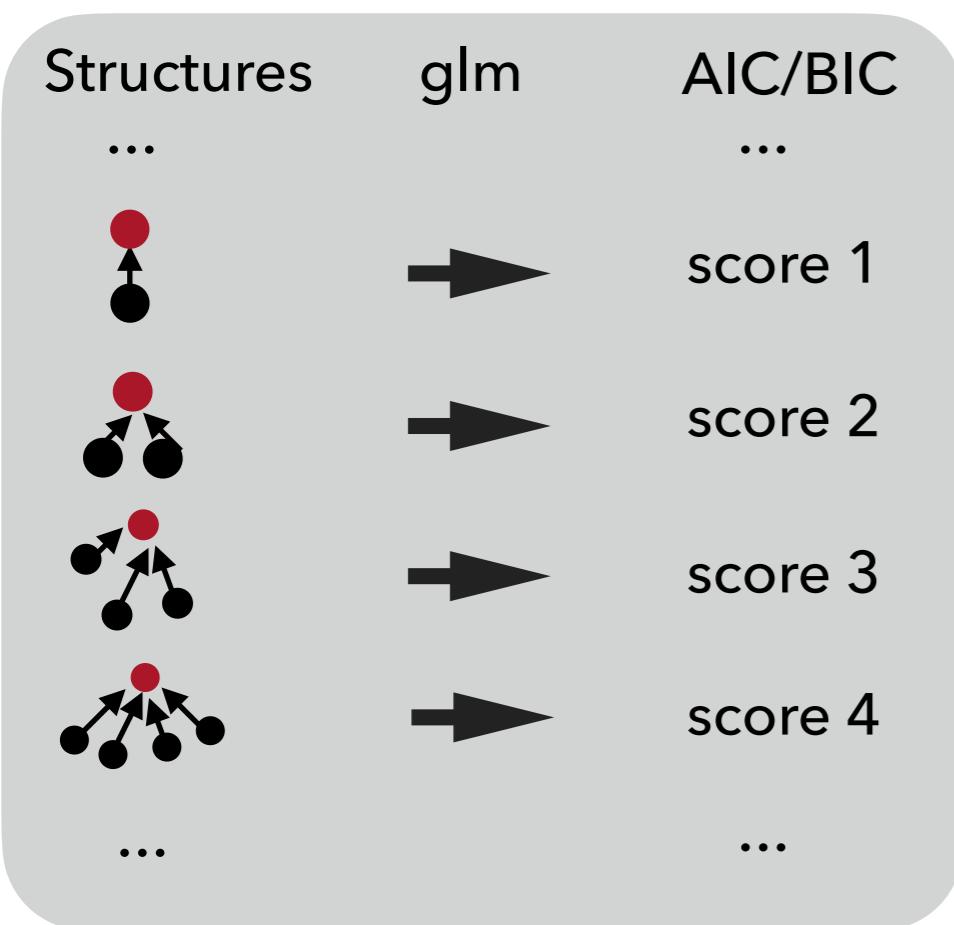


Exact or heuristic search

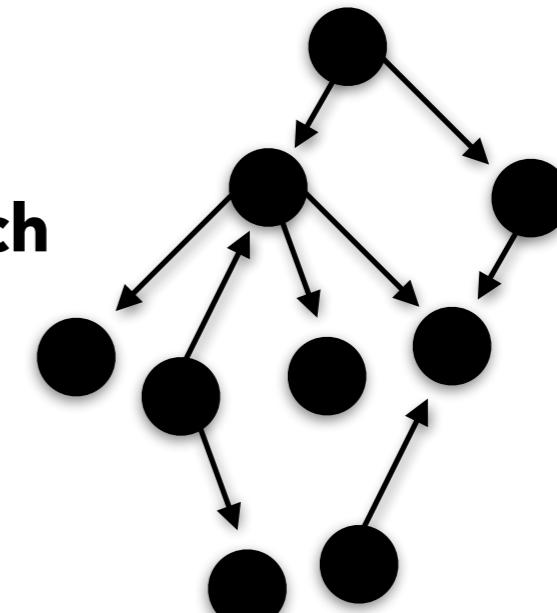


Bayesian network with highest posterior probability

Search and score algorithm



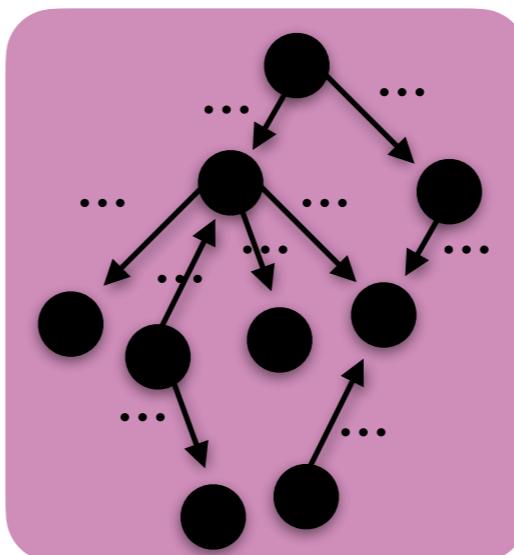
Exact or heuristic search



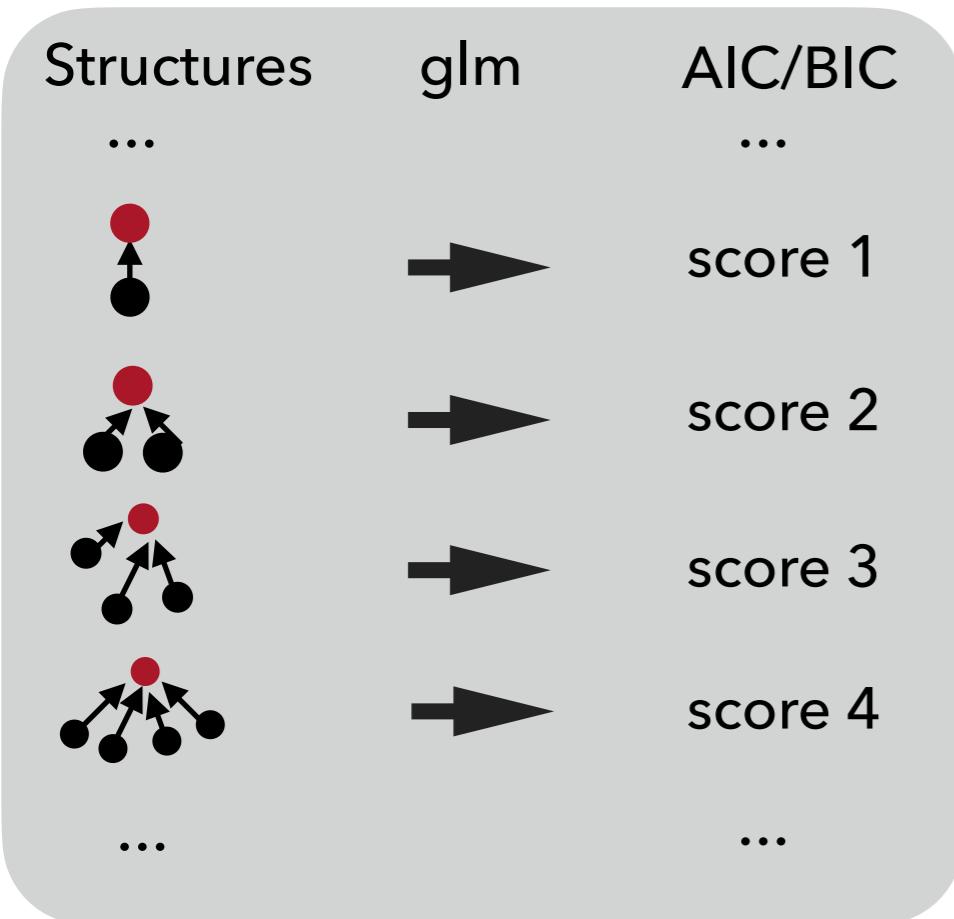
Bayesian network with highest posterior probability

Parameter estimation

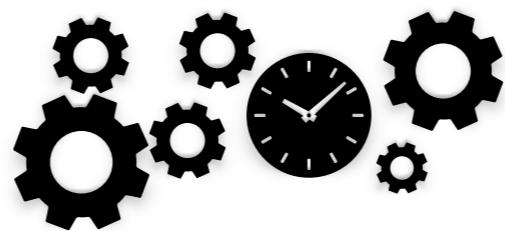
- ▶ compute marginal posterior density
- ▶ regression estimate



Search and score algorithm

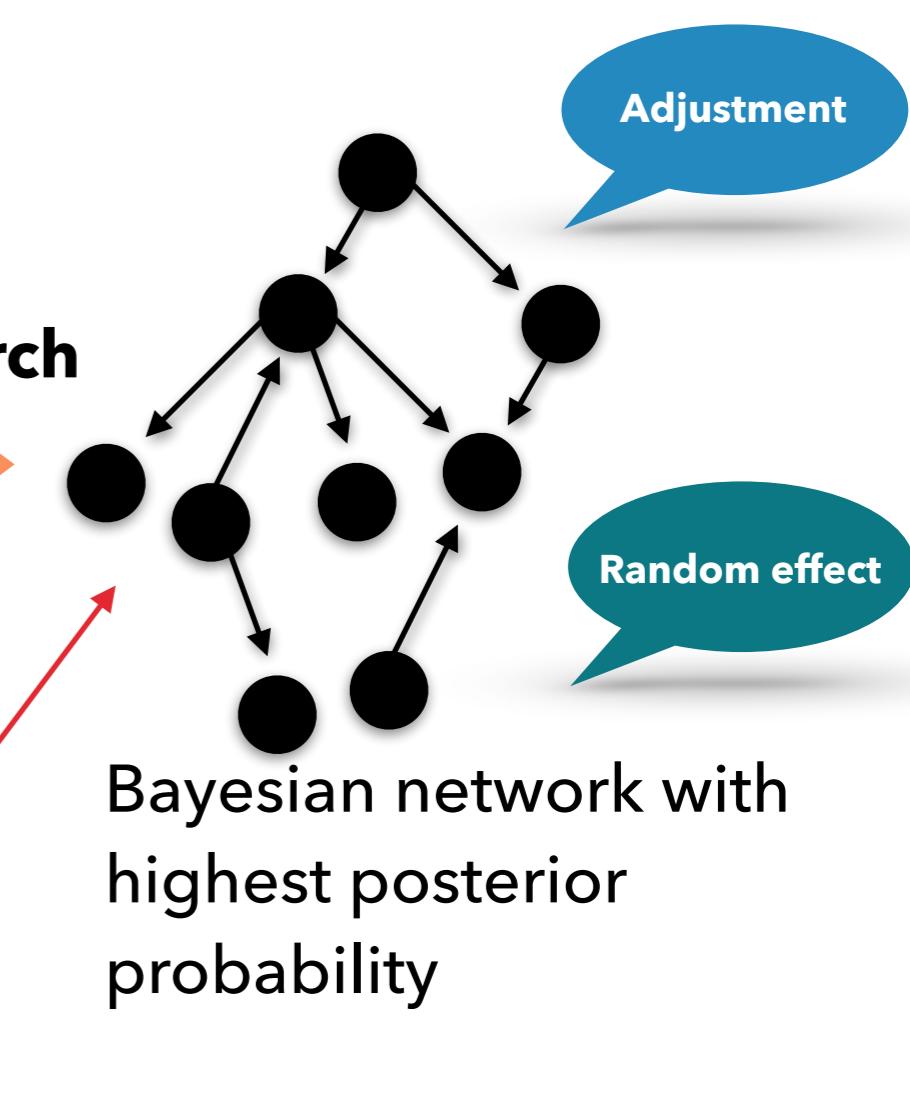


Exact or heuristic search



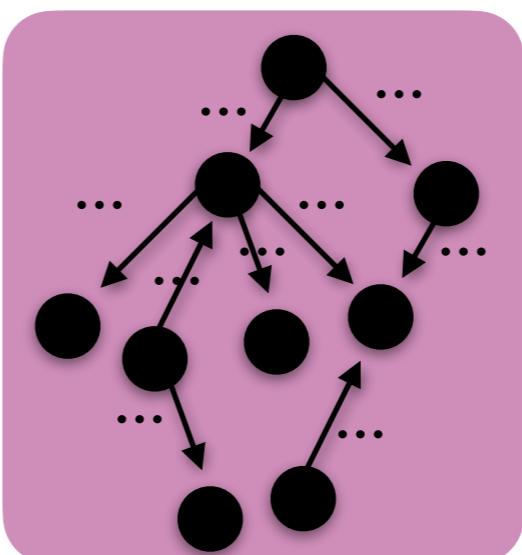
Causality!

Ban/Retain structures



Parameter estimation

- ▶ compute marginal posterior density
- ▶ regression estimate



Using R

```
buildScoreCache()
mostProbable()
fitAbn()
```

Core functions

buildScoreCache

Bayes

MLE

mostProbable

searchHillClimber

searchHeuristic

fitAbn

Bayes

MLE

Ancillary functions for analysis

toGraphviz

plotAbn

essentialGraph

Information theory

discretization

entropyData

miData

Graph inference

mb

linkStrength

Model inference

scoreContribution

Ancillary functions for simulation

simulateDag

simulateAbn

DAG metrics

createDag

compareDag

infoDag

Legend



Bayesian



MLE



Not applicable



Category

SELECTED BIBLIOGRAPHY

