The effect of Tilburg gold on your health

Misja Mikkers & Gertjan Verhoeven

Introduction

Causality

DAGS

Machine learning and causality

Conclusion

Blog and code

The effect of Tilburg gold on your health

Pacmed meetup: Combining causal inference and machine learning in practice

Misja Mikkers & Gertjan Verhoeven

Dutch Healthcare Authority (NZa) & Tilburg University

December 11, 2019

Nederlandse Zorgautoriteit

TILBURG UNIVERSITY

Understanding Society

Introduction

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Causal inference as missing data

The effect of Tilburg gold on your health

Misja Mikkers & Gertjan Verhoeven

Introduction

Causality

DAGS

Machine learning and causality

Conclusion

Blog and code

X_1	X_2	X_3	X_i	I	Y0	Y1
Male	9	14	1	0	67	NA
Female	60	36	0	1	NA	113
Female	7	2	1	1	NA	54
...	...	...	...	...	...	...

Prediction versus understanding

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

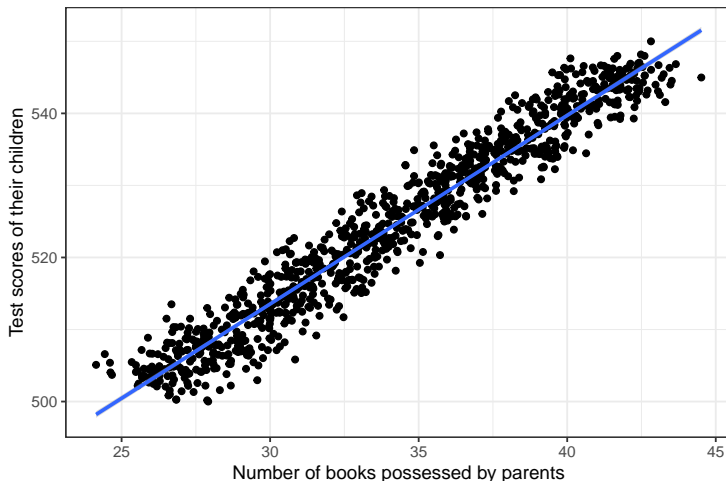**Causality**

DAGS

Machine learning
and causality

Conclusion

Blog and code

Prediction versus understanding

The effect of
Tilburg gold on
your health

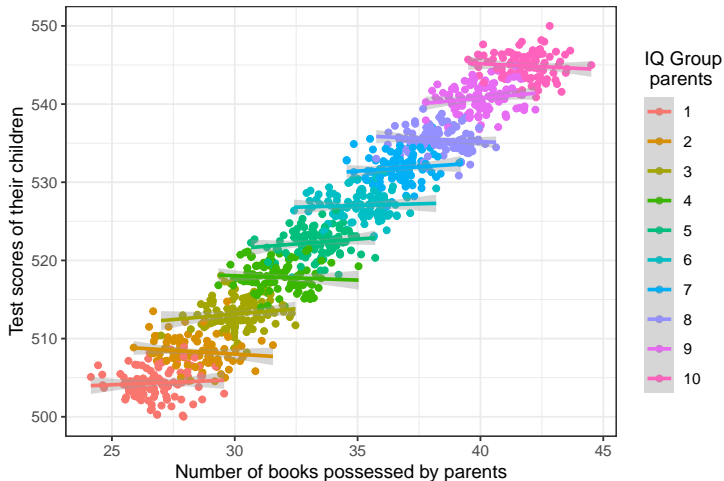**Misja Mikkers &
Gertjan
Verhoeven**

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

Causal graph (DAG) building blocks

The effect of
Tilburg gold on
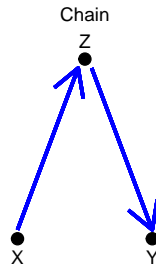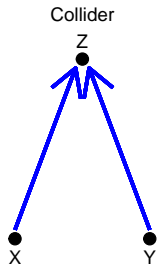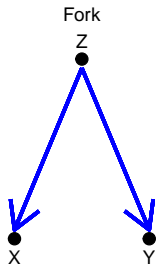your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

Fork

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

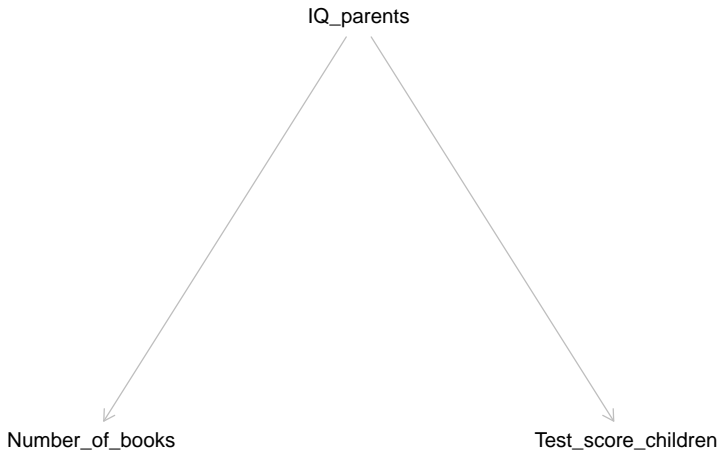Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

IQ_parents

Number_of_books   Test_score_children

Confounding / confounder

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

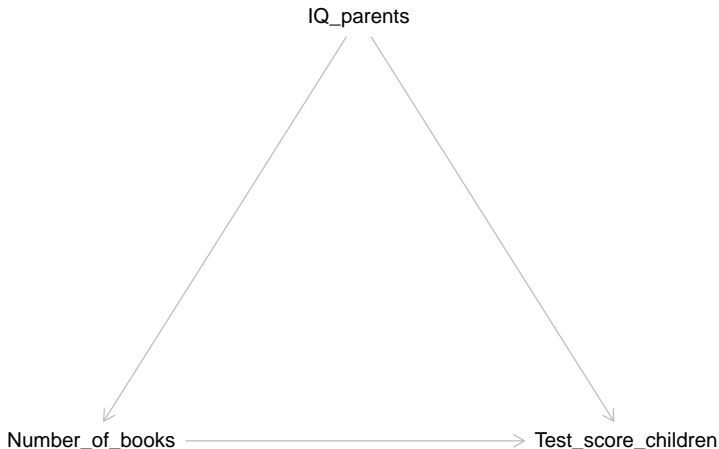**DAGS**

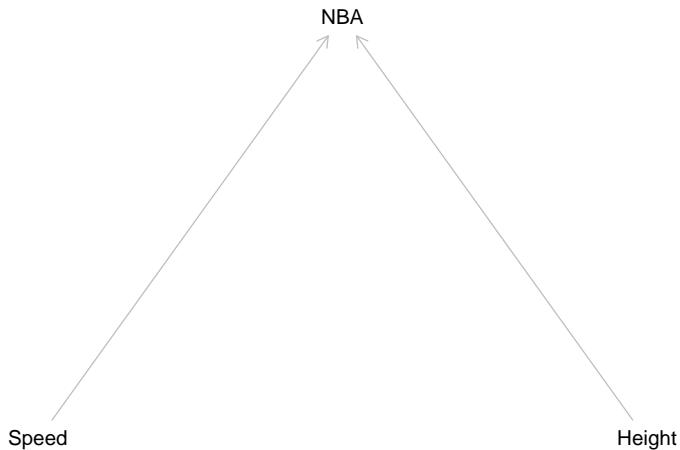Machine learning
and causality

Conclusion

Blog and code

IQ_parents

Number_of_books ⟶ Test_score_children

Collider

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

NBA

Speed

Height

Does being tall makes you fast?

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit

Does being tall makes you fast?

The effect of
Tilburg gold on
your health

**Misja Mikkers &
Gertjan
Verhoeven**

Introduction

Causality

DAGS
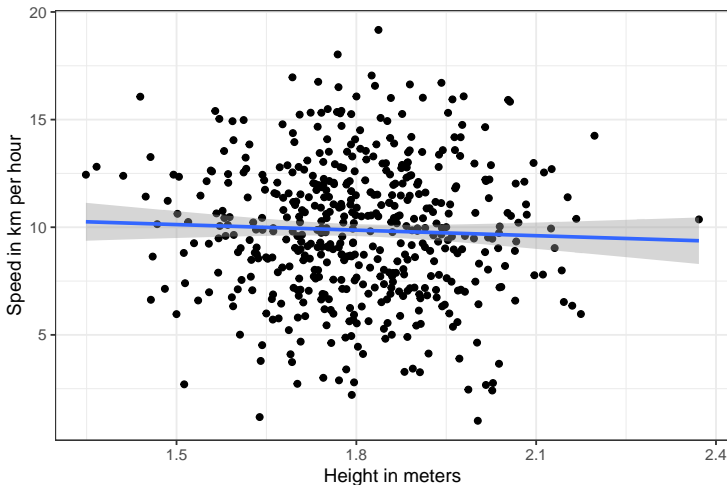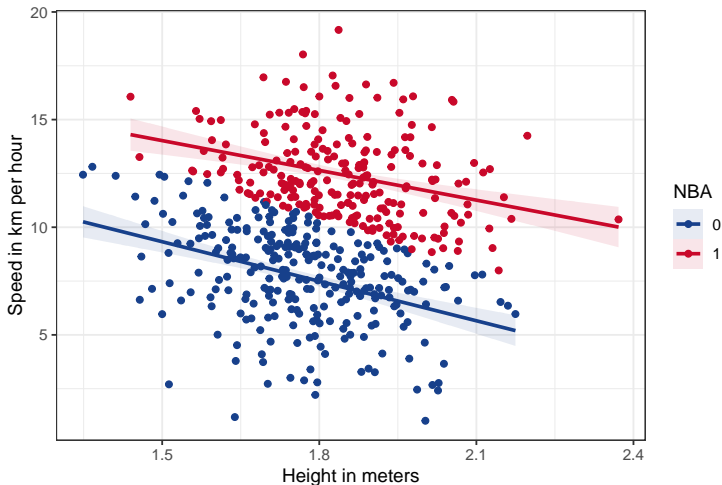
Machine learning
and causality

Conclusion

Blog and code

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit

Chain / Mediator

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit

Budgets and points

The effect of
Tilburg gold on
your health

Misja Mikkers &
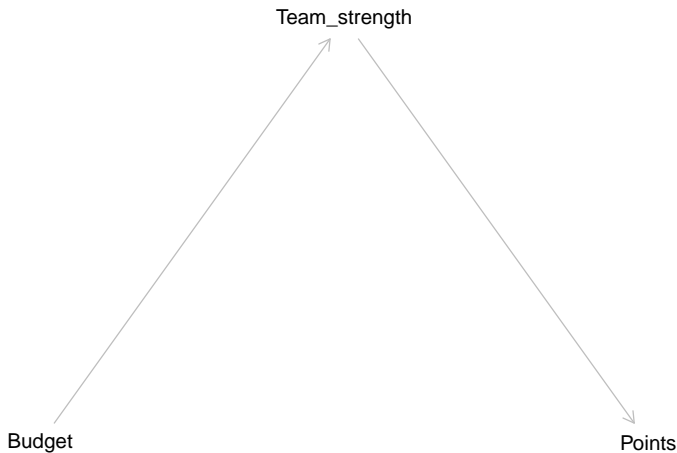Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

Budgets and Points

The effect of
Tilburg gold on
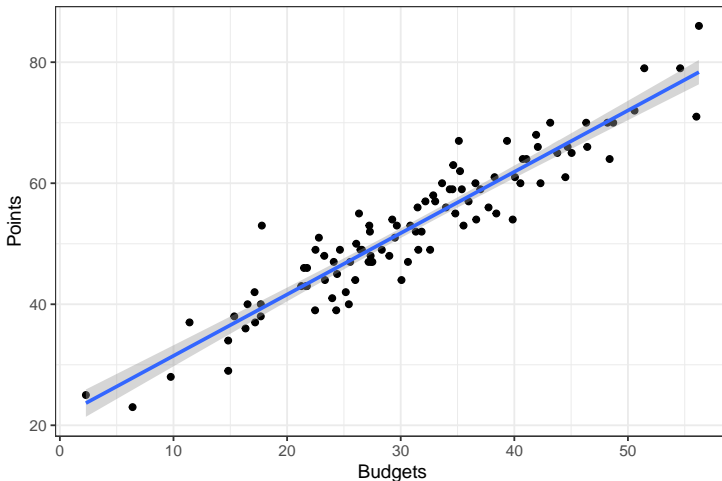your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

Example of a more complex DAG

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality
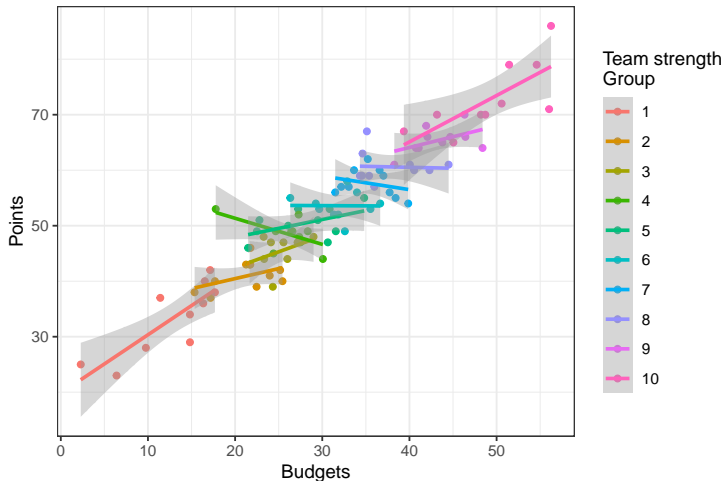
Conclusion

Blog and code

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit
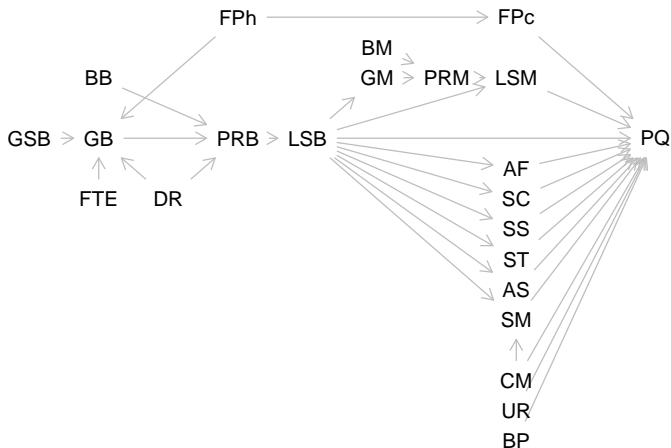
Machine learning and causality

The effect of
Tilburg gold on
your health

**Misja Mikkers &
Gertjan
Verhoeven**

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit

Approach

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction
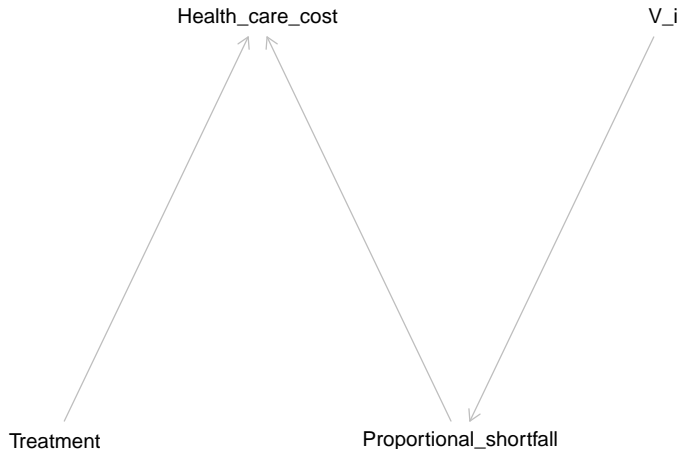
Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

The variable *Proportional_Shortfall* is based on:

$$Proportional\_Shortfall = abs(scale(V1^3 + 2 * V2 +$$
$$3 * V3^2 + 4 * V4 + 5 * V5$$
$$+6 * V6 * V7)) + \epsilon$$

1. Fit Random Forest model on the data 2. Determine the average treatment effect with generalized random forests (grf)

We will fit 2 models for each of these steps:

a. An analysis with all variables ("the wrong model") b. An analysis with all variables, except the collider *Health_care_cost* ("the right model")

Nederlandse
Zorgautoriteit

Summary statistics

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

Statistic	N	Mean	St. Dev.	Min	Max
Proportional Shortfall	1,000	1.080	0.841	0.0003	4.707
Treatment	1,000	0.493	0.500	0	1
Health care cost	1,000	6.887	4.544	−1.642	26.389
V1	1,000	497.768	287.495	1	1,000
V2	1,000	498.749	285.709	2	1,000
V3	1,000	499.119	292.200	4	1,000
V4	1,000	504.484	294.009	1	1,000
V5	1,000	489.628	282.146	1	1,000
V6	1,000	506.039	284.987	2	1,000
V7	1,000	496.600	283.415	1	998
V8	1,000	498.522	287.912	1	999

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit

Predictions Random Forest

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
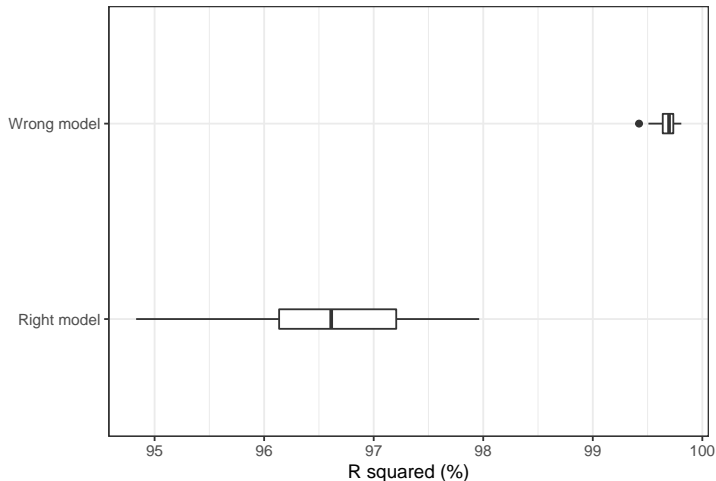Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

Average treatment effect

The effect of
Tilburg gold on
your health

Misja Mikkers &
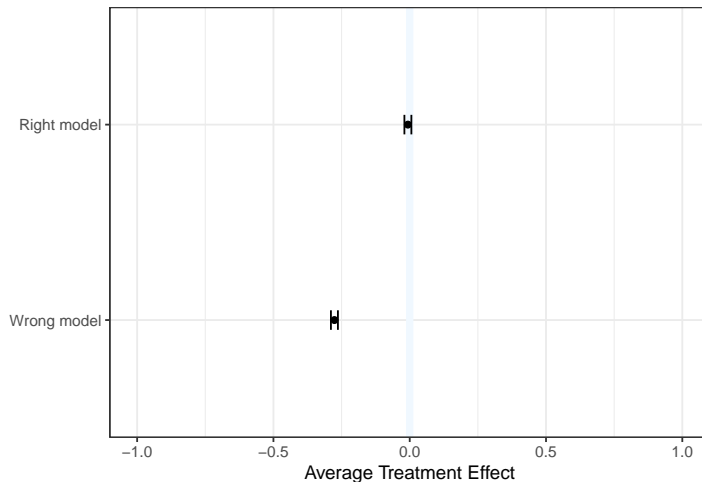Gertjan
Verhoeven

Introduction

Causality

DAGS

Machine learning
and causality

Conclusion

Blog and code

TILBURG
UNIVERSITY

Understanding
Society

Conclusion

The effect of
Tilburg gold on
your health

Misja Mikkers &
Gertjan
Verhoeven

- ▶ Causal models are necessary for inference
- ▶ It is tempting to use all variables in a machine learning model
- ▶ However, this could lead to misleading conclusions

TILBURG
UNIVERSITY

Understanding
Society

Nederlandse
Zorgautoriteit

Blog and code (in Dutch)

The effect of Tilburg gold on your health

Misja Mikkers & Gertjan Verhoeven

Introduction

Causality

DAGS

Machine learning and causality

Conclusion

Blog and code

https://misjamikkers.github.io/post/causaliteit-en-machine-learning/

https://github.com/misjamikkers/Meetup_Informatieberaad

Nederlandse
Zorgautoriteit

TILBURG
UNIVERSITY

Understanding
Society