PADERBORN UNIVERSITY

Causal Machine Learning Approaches For Modelling Data Center Heat Recovery: A Physical Testbed Study.

David Zapata Gonzalez, Marcel Meyer and Oliver Müller

HotCarbon - Workshop on Sustainable Computer Systems. Cambridge, July 2025

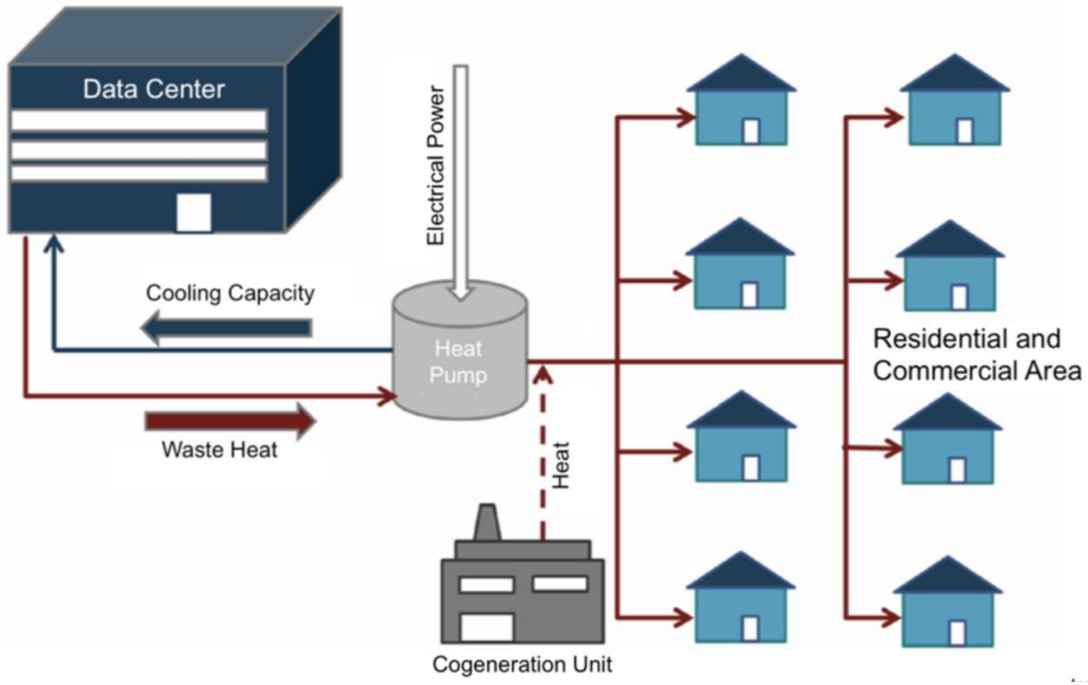
Motivation - Data Centers Must Utilize Waste Heat

EU Energy Efficiency Directive 2023/1791 - Germany: Energy Efficiency Act (EnEfG)

The heat produced by data centers is typically removed through cooling systems; however, it can be instead reused as a sustainable alternative to fossil fuels for residential and industrial heating.

Legal Requirements in Germany: New data centers with a 300 kW non-redundant connection must utilize a minimum share of their waste heat.

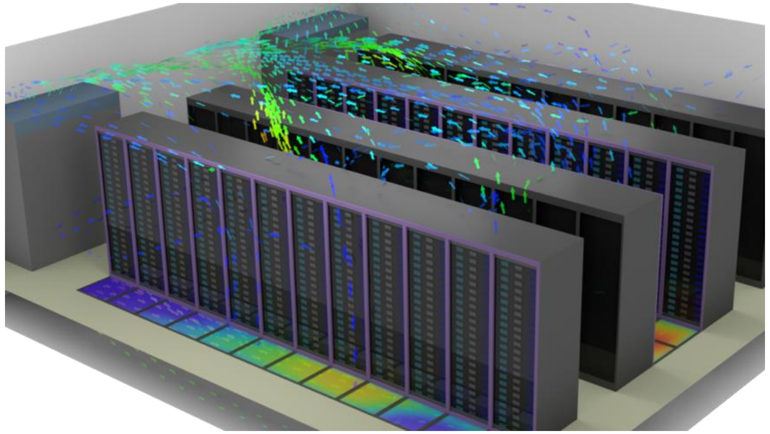
- Commissioned from July 1, 2026: At least 10%.
- Commissioned from July 1, 2027: At least 15%.
- Commissioned from July 1, 2028: At least 20%.



https://www.dotmagazine.online/issues/on-the-edge-building-the-foundations-for-the-future

Motivation

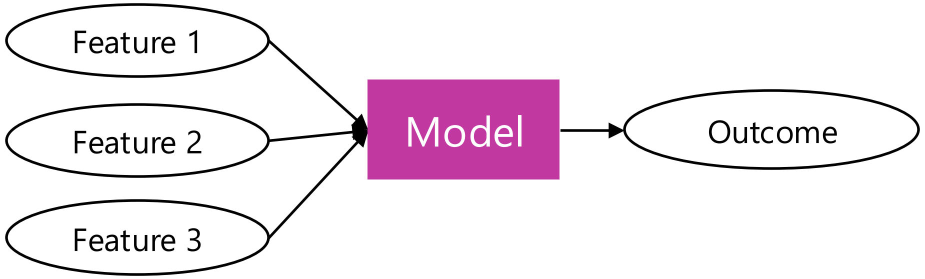
Simulation Models



https://blogs.sw.siemens.com/simcenter/eleven-top-tips-for-energy-efficient-data-center-design-and-operation/

Data-driven Models

Machine Learning



(Athavale et al., 2019)

Causal Machine Learning (CML)

- ✓ Principles from causal inference (Graphical causal models and causal discovery) and flexible ML modelling (Pearl, 2009)
- ✓ Robust data-driven models

PADERBORN UNIVERSITY

Motivation

Evaluation challenge of Causal ML

- ✓ Interventions are rare in real data centers
- ✓ Most approaches are evaluated with simulation or observational data (no interventions)

DC with heat recovery testbed

- ✓ Building a testbed is necessary to assess effects of interventions (Gamella et al., 2025; Wiesner et al., 2023)

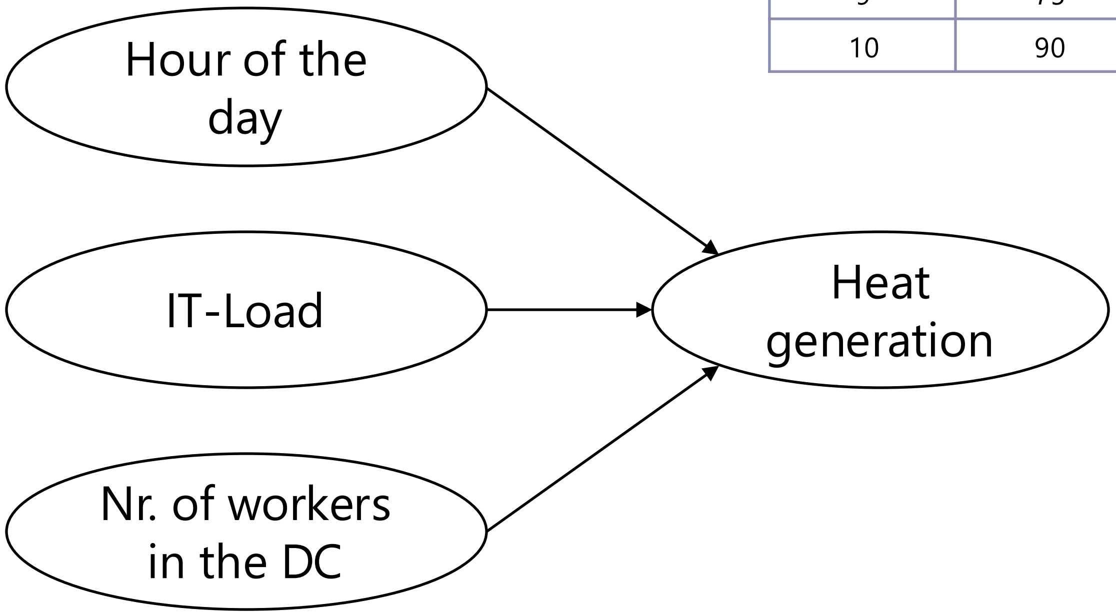
Background – Causal Graphs and Causal Discovery

Hour of the day	IT-Load	Nr. of workers in the DC	Heat generation
7	50	2	200
8	60	5	250
9	75	6	300
10	90	7	450

PADERBORN UNIVERSITY

Background – Causal Graphs and Causal Discovery

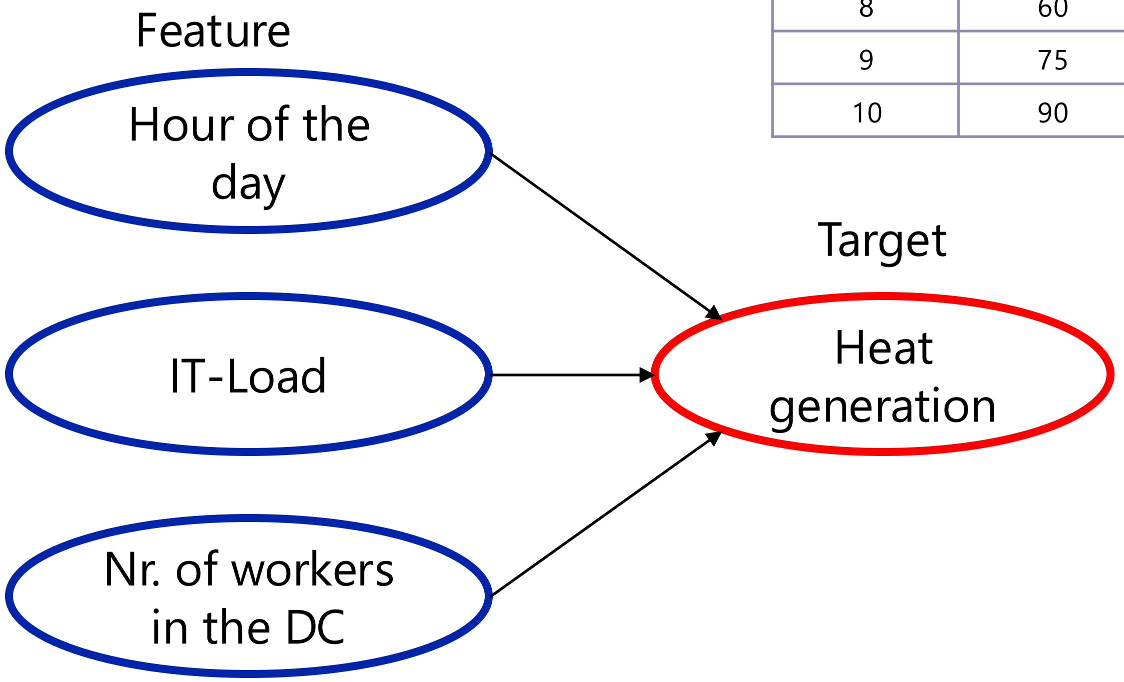
Hour of the day	IT-Load	Nr. of workers in the DC	Heat generation
7	50	2	200
8	60	5	250
9	75	6	300
10	90	7	450



Machine Learning

Background – Causal Graphs and Causal Discovery

Hour of the day	IT-Load	Nr. of workers in the DC	Heat generation
7	50	2	200
8	60	5	250
9	75	6	300
10	90	7	450

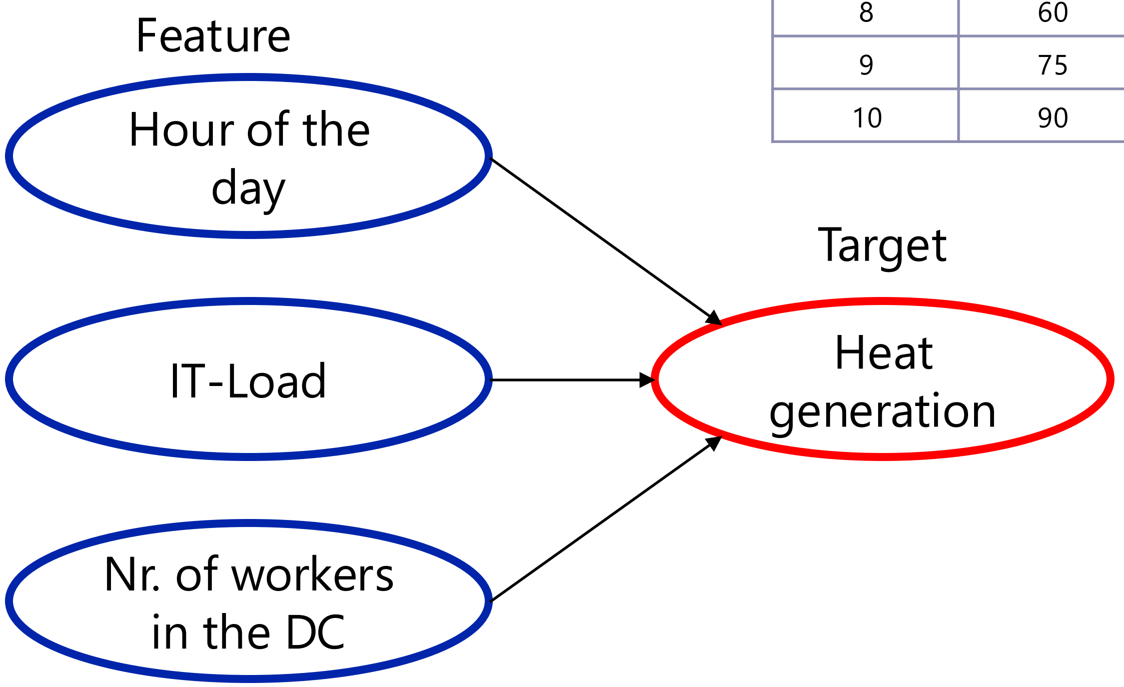


Machine Learning

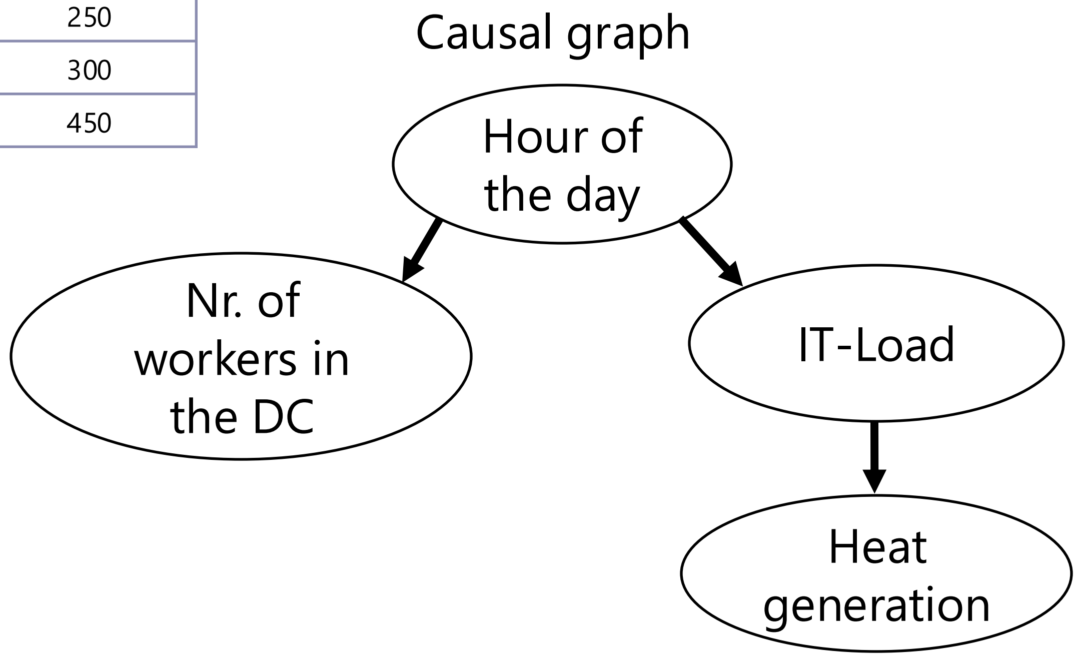
PADERBORN UNIVERSITY

Background – Causal Graphs and Causal Discovery

Hour of the day	IT-Load	Nr. of workers in the DC	Heat generation
7	50	2	200
8	60	5	250
9	75	6	300
10	90	7	450



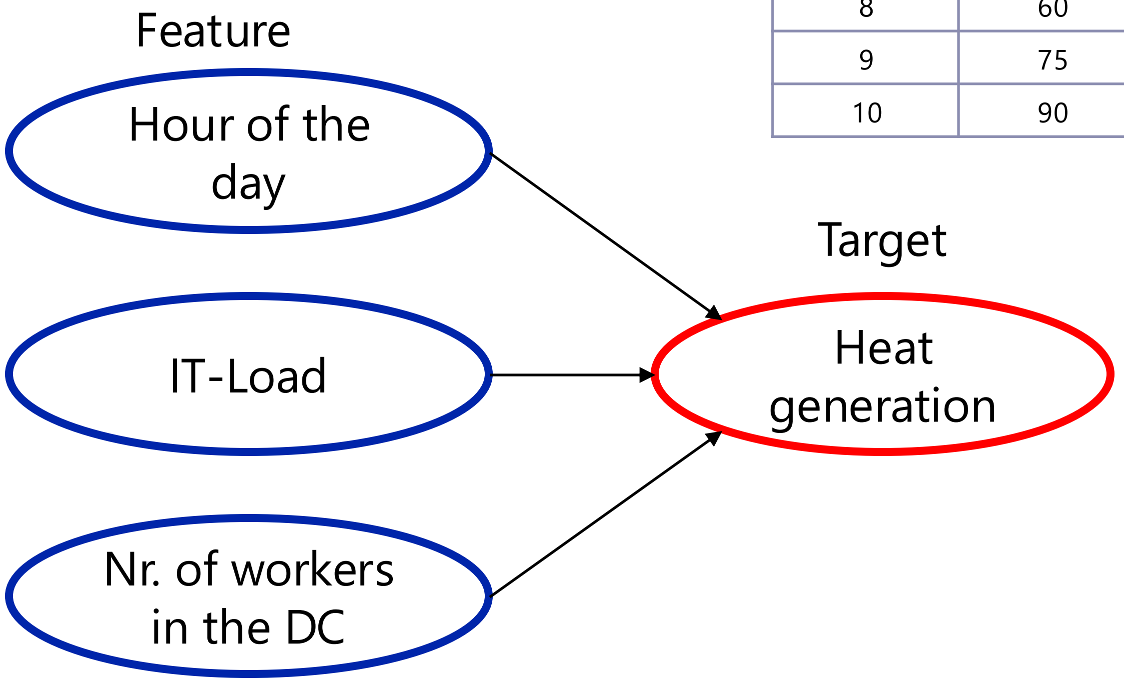
Machine Learning

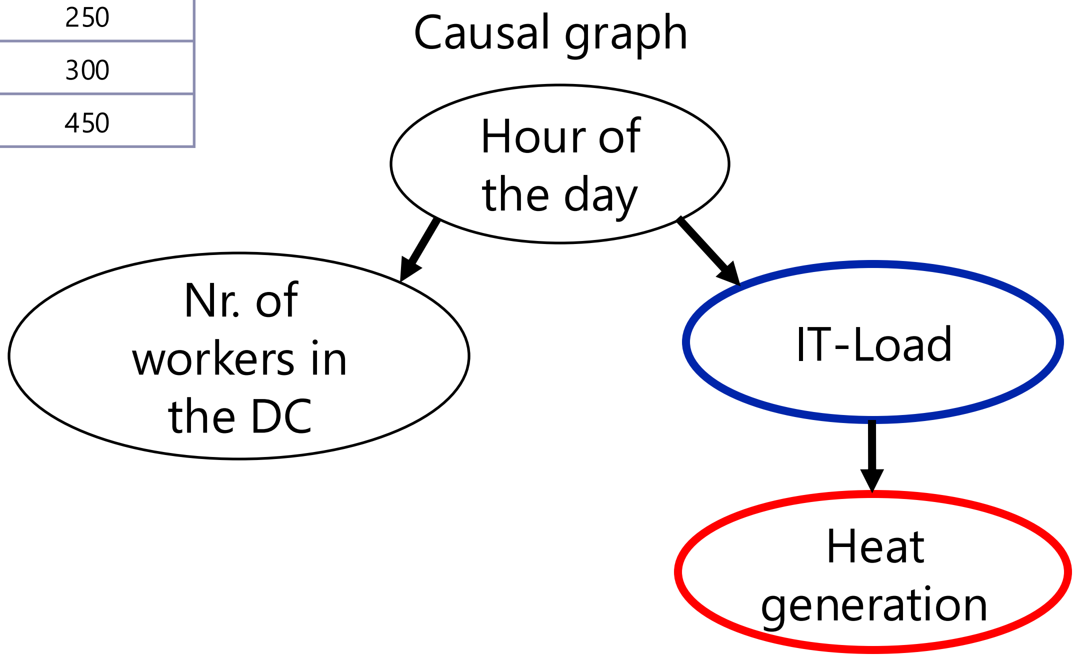


Causal Machine Learning

Background – Causal Graphs and Causal Discovery

Hour of the day	IT-Load	Nr. of workers in the DC	Heat generation
7	50	2	200
8	60	5	250
9	75	6	300
10	90	7	450

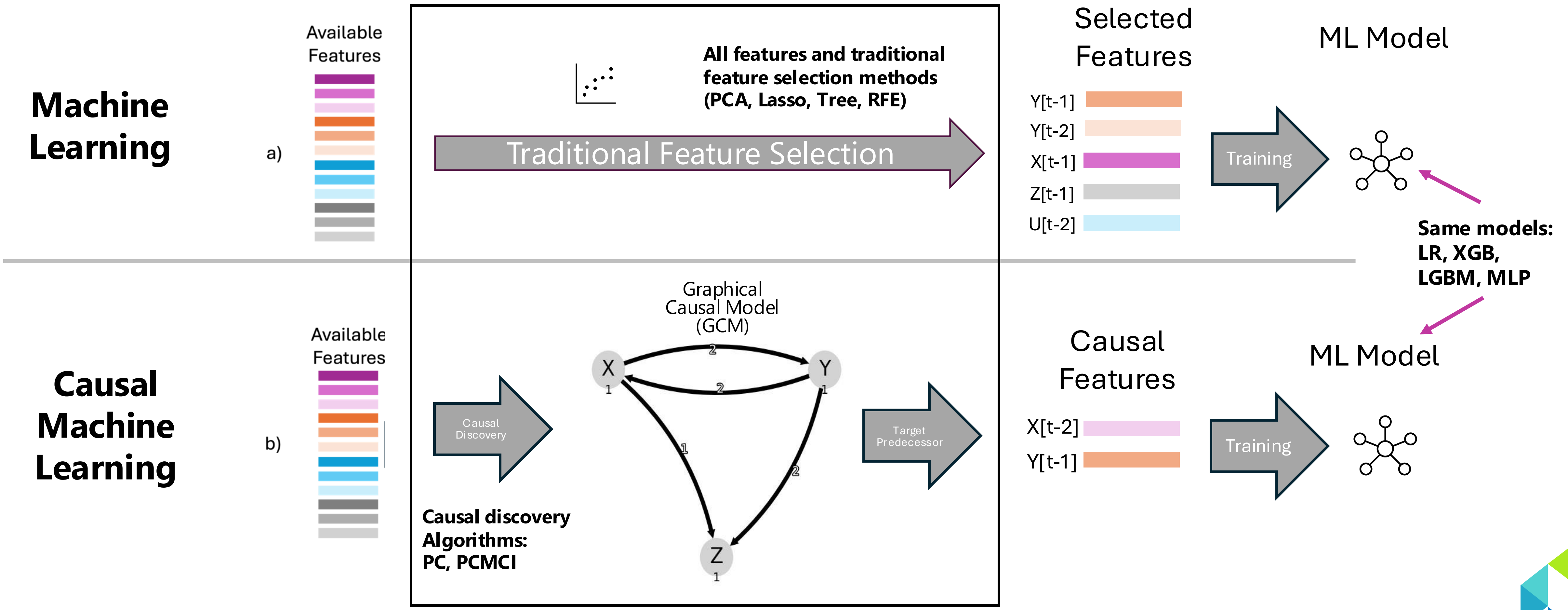




Machine Learning

Causal Machine Learning

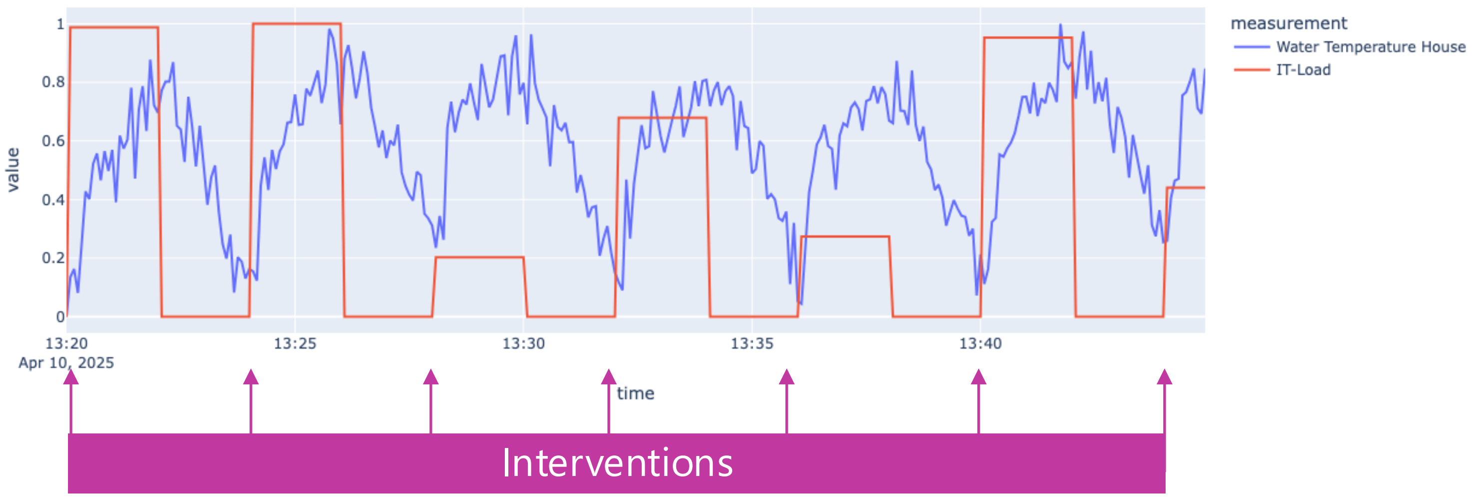
Methodology - Framework



Experiments - Testbed

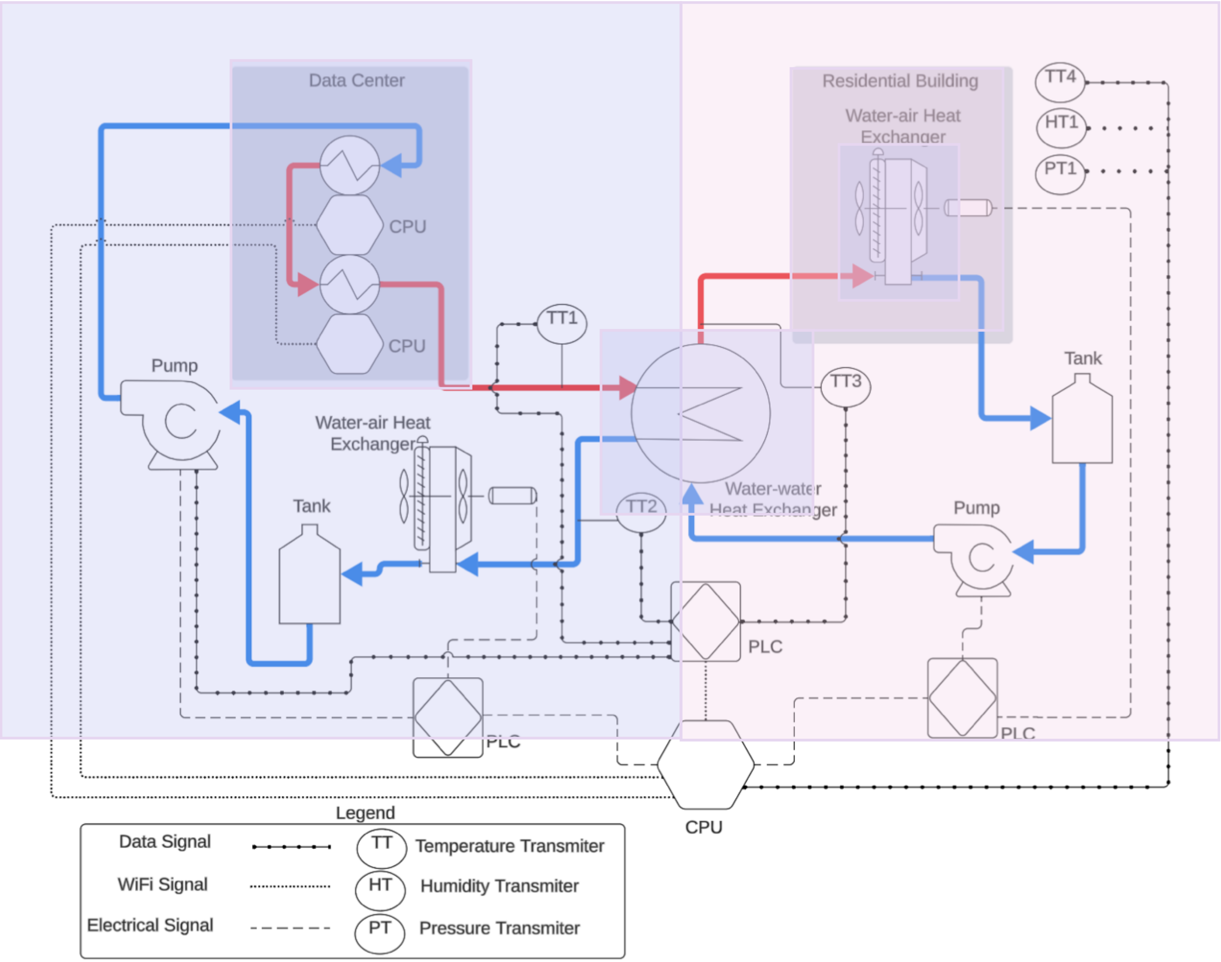
- ✓ Water cooled mini–DC with Heat exchanger network for heat recovery
- ✓ Generate data for 5 experiments with different interventions
- ✓ Train/test split
- ✓ 2 Evaluations

Selected Variables and Interventions (MinMaxScaled)



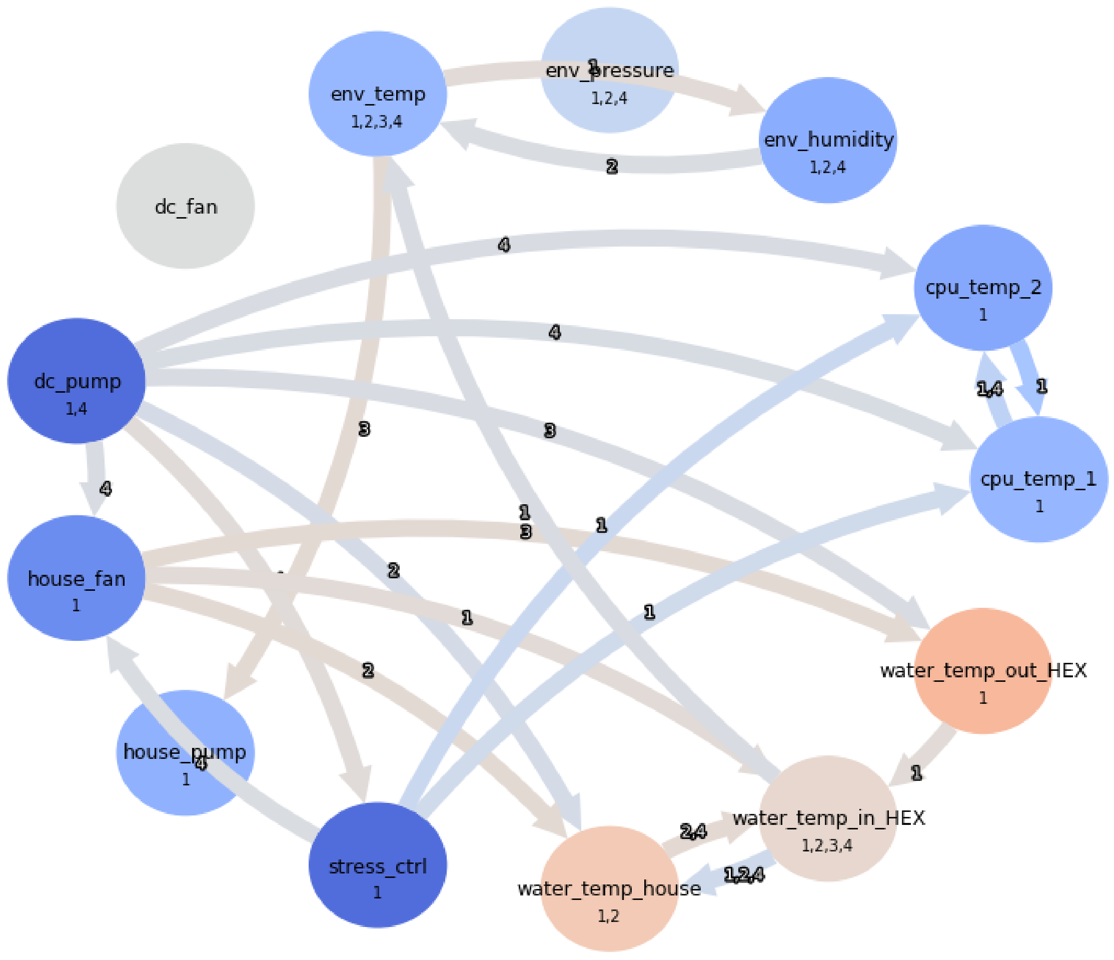
PADERBORN UNIVERSITY

Experiments - Testbed



Variable	Description
cpu_temp_1	Pi 1 CPU (°C)
cpu_temp_2	Pi 2 CPU (°C)
env_humidity	Room Humidity (HT1)
env_pressure	Room Pressure (PT1)
env_temp	Room Temperature (°C) (TT4)
dc_fan	DC fan state (%)
dc_pump	DC pump state (%)
house_fan	House fan state (%)
house_pump	House pump state (%)
stress_ctrl	Stress control CPU (%)
water_temp_in_HEX	Water in HEX (°C) (TT1)
water_temp_out_HEX	Water out HEX (°C) (TT2)
water_temp_house	House water (°C) (TT3)

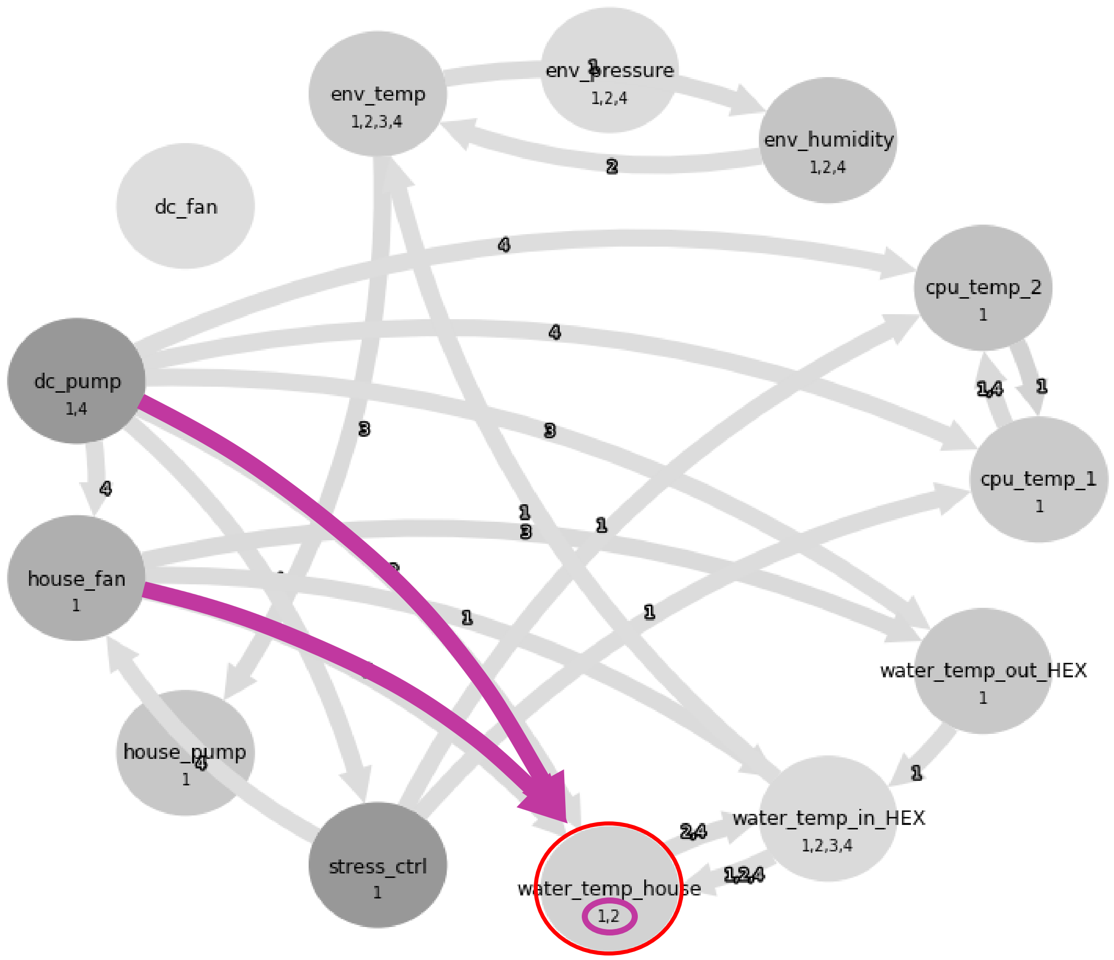
Experiment – Causal Discovery



- ✓ Summary causal graph from causal discovery for time-series data. Colours represent the partial and autocorrelations
- ✓ The numbers indicate when a causal effect appears
- ✓ A feature is a variable and its lag (delays between observations)

In this case:
13 Variables x 4 lags = 52 Features

Experiment – Causal Discovery



For predicting water Temperature in district heating:

- ✓ For ML: with all predictors (52 features)
- ✓ For causal ML: only causal links (4 Features)

Experiment – Evaluation 1: Complete Test Set

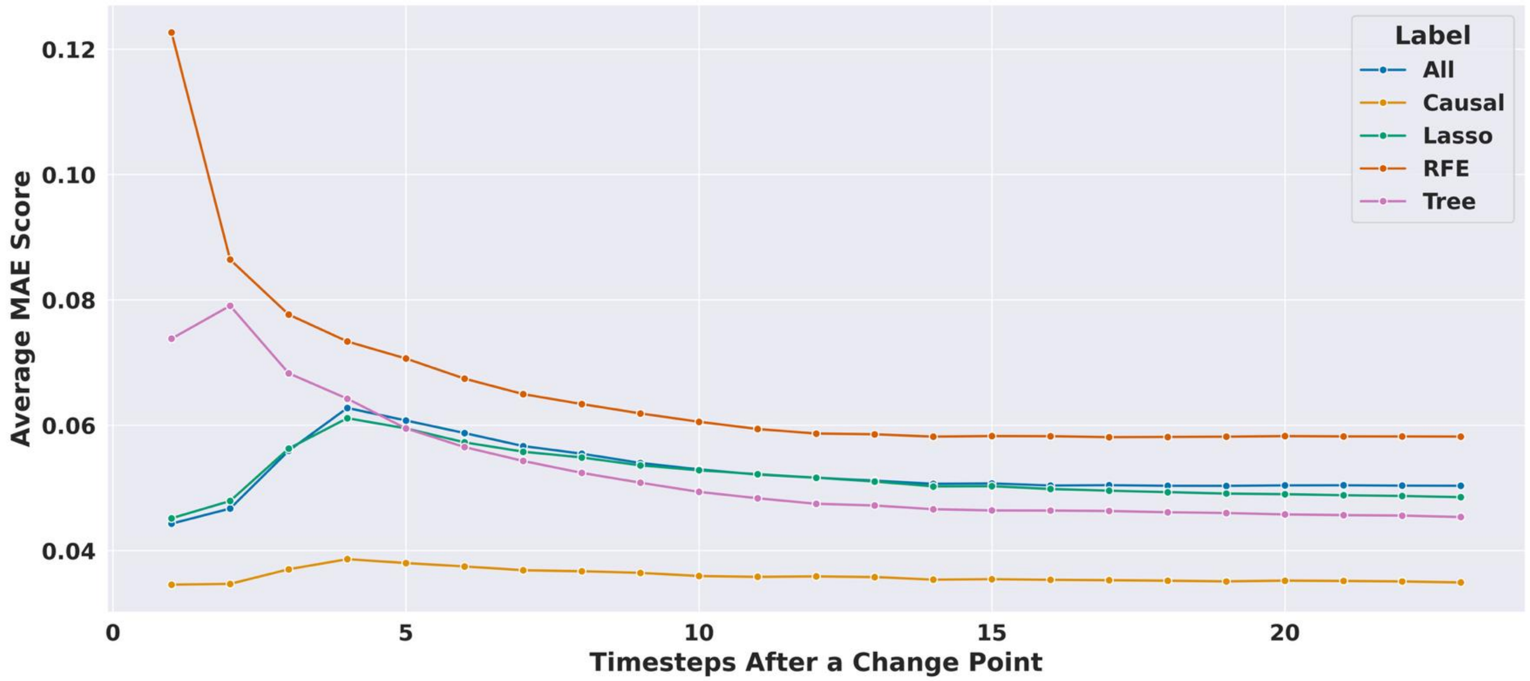
Exp.	Feat. Selec.	Model	F. Nº	MAE	MSE	MAPE
1	Causal	LR	20	0.0357	0.0021	0.1581
	Tree	LR	2	0.0375	0.0023	0.1660
	RFE	LR	26	0.0376	0.0023	0.1665
	Lasso	LR	18	0.0409	0.0028	0.1808
	All	LR	52	0.0413	0.0028	0.1823
	PCA	XGB	2	1.1451	1.5389	5.0212
2	Causal	LR	16	0.0329	0.0017	0.1417
	Lasso	LR	16	0.0381	0.0025	0.1639
	Tree	LR	6	0.0408	0.0032	0.1755
	RFE	LR	26	0.0455	0.0047	0.1963
	All	RF	52	0.0518	0.0043	0.2223
	PCA	RF	2	0.5349	0.3802	2.2828
3	Causal	LR	4	0.0319	0.0016	0.1411
	All	LR	52	0.0331	0.0019	0.1458
	Tree	LR	9	0.0358	0.0020	0.1580
	Lasso	LR	16	0.0359	0.0022	0.1585
	RFE	LR	26	0.0517	0.0047	0.2265
	PCA	ENet	2	1.6007	2.7418	7.0543
4	Causal	RF	19	0.0354	0.0020	0.1521
	Lasso	LGBM	19	0.0355	0.0020	0.1528
	Tree	LR	9	0.1289	3.5293	0.5507
	All	RF	52	0.1409	0.0873	0.6034
	RFE	LR	26	0.1621	2.8015	0.6931
	PCA	MLP	3	1.2977	1.7766	5.5705
5	Causal	LR	18	0.0342	0.0019	0.1481
	Tree	LR	9	0.0348	0.0020	0.1509
	All	LR	52	0.0443	0.0031	0.1911
	RFE	ENet	26	0.0506	0.0040	0.2184
	Lasso	LR	17	0.0743	0.0078	0.3192
	PCA	MLP	2	1.1619	1.6752	4.9813

Results

- ✓ They were the best method in all experiments
- ✓ Causal approaches achieved on average 0.3% to 14.6% better performance than traditional ML methods

PADERBORN UNIVERSITY

Experiment – Evaluation 2: Focus on Interventions



Results

✓ Causal approaches achieved on average 20%-50% better performance than traditional ML methods

PADERBORN UNIVERSITY

Discussion

Testbed

Testbeds are important for evaluating new approaches of ML for prediction and control of dynamical systems, such as Data Centers with heat recovery

Practical implications

- ✓ First determine causal structure and then model with machine learning
- ✓ Interventions in real data centers are crucial for generating relevant data for causal approaches
- ✓ When the number for sensors increase, their values can be aggregated
- ✓ More features are not necessarily better

Performance

Is the performance of ML models trained with causal features competitive in comparison to other traditional feature selection approaches?

- ✓ Causal ML was better in both evaluations. If the focus is only prediction, traditional ML has also good performance, if the focus is on interventions, CML is clearly better

Limitations and future work

- ✓ The performance of the causal approaches depends heavily on the quality of the causal models:
  - ✓ Number of interventions, complexity and data quantity
- ✓ Use these causal models as base for training reinforcement learning agent or model predictive control instead of simulations

Thank you

PADERBORN UNIVERSITY

References

- Athavale, J., Yoda, M., & Joshi, Y. (2019). Comparison of data driven modeling approaches for temperature prediction in data centers. *International Journal of Heat and Mass Transfer*, *135*, 1039-1052.
- Pearl, J. (2009). *Causality*. Cambridge university press.
- Gamella, J. L., Peters, J., & Bühlmann, P. (2025). Causal chambers as a real-world physical testbed for AI methodology. *Nature Machine Intelligence*, 1-12.
- Runge, J., Gerhardus, A., Varando, G., Eyring, V., & Camps-Valls, G. (2023). Causal inference for time series. *Nature Reviews Earth & Environment*, *4*(7), 487–505.
- Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*(7), 075310.
- Gillis, T. B., & Spiess, J. L. (2019). Big data and discrimination. *The University of Chicago Law Review*, *86*(2), 459-488.
- Lundberg, S., Dillon, E., LaRiviere, J., Roth, J., & Syrgkanis, V. (2021). *Be careful when interpreting predictive models in search of causal insights; SHAP latest documentation—Shap.readthedocs.io*.
- Wiesner, Philipp, Ilja Behnke, Paul Kilian, Marvin Steinke, und Odej Kao. „Vessim: A Testbed for Carbon-Aware Applications and Systems". arXiv, 2023. https://doi.org/10.48550/ARXIV.2306.09774.
- Christoph Böhm (SAP), Fabian Buda (Data Center Group), Günter Eggers, (NTT Global Data Centers), Helmut Göhl (Ingenieurbüro GTS), Dr. Ralph Hintemann, (Borderstep Insitut), Stephan Hülskamp (COOLtec Systems), Michael Jakobi (contagi), Johannes Krafczyk (T-Systems), Uwe Müller (InfraOpt), David Schöler (MVV Enamic), und Nick Skala (Siemens), Mirjam Wilhelm (NTT Global Data Centers). „Energieeffizienzgesetz für Rechenzentren". Bitkom e.V., 2024. https://www.bitkom.org/sites/main/files/2024-01/bitkom-leitfaden-energieeffizienzgesetz-fuer-rechenzentren.pdf.