Applied AI Seminar Series: Causal ML for Healthcare Applications

Washington University in St. Louis

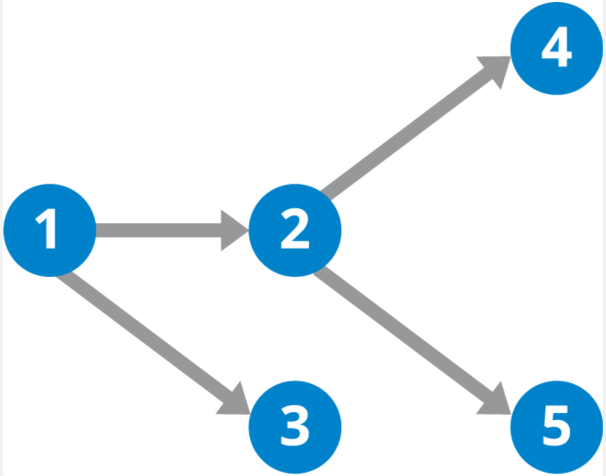
School of Medicine

Causal Chains in Healthcare Modeling

- Identifying Treatment Efficacy: Causal learning discerns effective treatments from patient data, distinguishing successful interventions from coincidental recoveries.
- Predicting Outcomes: Forecast patient responses to treatments, enabling personalized healthcare plans based on predicted outcomes.
- Enhancing Clinical Decision-Making: Provides evidence-based insights, supporting clinicians in making informed decisions about patient care.
- Understanding Disease Progression: Helps unravel the complex causal pathways of diseases, facilitating early intervention and preventive strategies.
- Policy and Resource Allocation: Informs healthcare policy and resource allocation by identifying causal factors that significantly impact public health outcomes.

LiNGAM for Causality

- Non-Gaussian Assumption: LiNGAM identifies causal relationships using data that is assumed to be non-Gaussian, distinguishing it from other causal discovery methods.
- Directed Acyclic Graph (DAG): Utilizes a DAG to represent causal structures, with nodes as variables and edges indicating direct causal influences.
- Causal Discovery: Employs statistical techniques, such as independent component analysis, to uncover the causal order of variables without needing prior knowledge.
- Healthcare Application: In healthcare, LiNGAM can analyze patient data to reveal causal factors affecting outcomes, aiding in treatment and policy decisions.
- Advantages and Limitations: Offers a unique approach to causal inference, particularly where traditional models falter, though its effectiveness depends on the data's adherence to its non-Gaussian and no-latent-confounder assumptions.



DoWhy for Causality

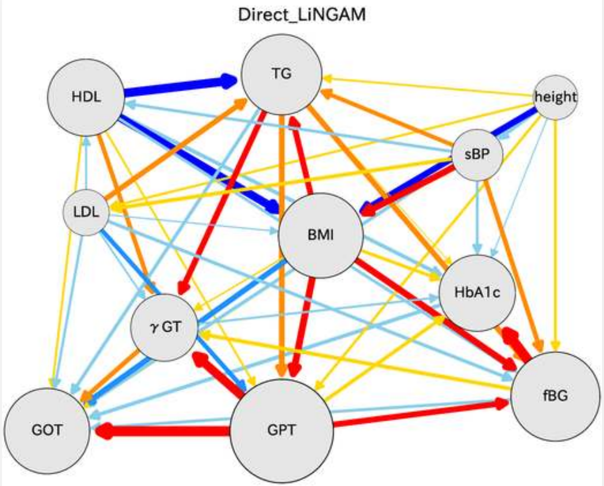
- Causal Inference Framework: DoWhy is a Python package designed to facilitate causal inference by providing a unified interface for modeling, estimating, and interpreting causal effects. It helps users to define and validate causal assumptions in their analyses.
- Four-Step Process: The package follows a four-step process: Model, Identify, Estimate, and Refute. This structured approach ensures rigorous causal analysis by explicitly modeling assumptions and systematically verifying results.
- Graphical Models: DoWhy uses graphical causal models (causal graphs) to represent assumptions visually. These graphs help in identifying potential sources of bias and in determining the appropriate methods for estimating causal effects.
- Multiple Estimation Methods: The package supports various causal effect estimation methods, including propensity score matching, instrumental variables, and regression discontinuity designs, allowing flexibility and robustness in analysis.
- Refutation and Validation: DoWhy emphasizes the importance of validating causal claims. It provides tools to refute causal estimates through robustness checks, such as placebo tests and sensitivity analyses, ensuring the credibility of the results.

Comparing Methods

Feature	DoWhy	LiNGAM
Purpose	Causal inference framework	Linear Non-Gaussian Acyclic Model for causal discovery
Approach	Four-step process: Model, Identify, Estimate, Refute	Exploits non-Gaussianity to identify causal structure
Graphical Models	Uses causal graphs to represent assumptions	Constructs causal graphs based on data
Estimation Methods	Supports multiple methods (e.g., propensity scores, instrumental variables)	Focuses on linear models with non-Gaussian noise
Validation	Emphasizes robustness checks and sensitivity analyses	Provides statistical tests for causal ordering
Flexibility	Broad range of estimation techniques	Specific to linear non-Gaussian models
Use Case	General-purpose causal inference	Suitable for discovering causal structures in datasets with non-Gaussian variables
Ease of Use	Unified interface for causal analysis	Requires understanding of linear models and non-Gaussianity
Assumptions	Explicit modeling of causal assumptions	Assumes linearity and non-Gaussianity
Applications	Healthcare, social sciences, economics	Neuroscience, genomics, econometrics

Evaluating Causal Models

- Graph Completeness: Ensure the causal graph includes all relevant variables and relationships to accurately represent the domain.
- Directional Arrows: Verify that the arrows correctly represent the causal direction between variables based on domain knowledge or prior studies.
- Conditional Independencies: Check that the graph encodes the correct conditional independencies, aligning with the data and theoretical expectations.
- Plausibility of Assumptions: Assess the plausibility of the causal assumptions made, such as no unmeasured confounders, using domain expertise and supporting evidence.
- Sensitivity Analysis: Conduct sensitivity analyses to test how changes in the graph structure or assumptions affect the causal conclusions, ensuring robustness.



Let's Look at LiNGAM

Let's a go!