UNIVERSITY OF MARYLAND

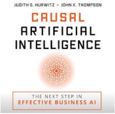
MSML610: Advanced Machine Learning

8.3: Causal AI in Business

Instructor: Dr. GP Saggese - gsaggese@umd.edu

References:

- Hurwitz, Thompson: Causal Artificial Intelligence, 2024



- ***Business Processes Around Data Modeling***
  - Modeling Processes
  - Roles

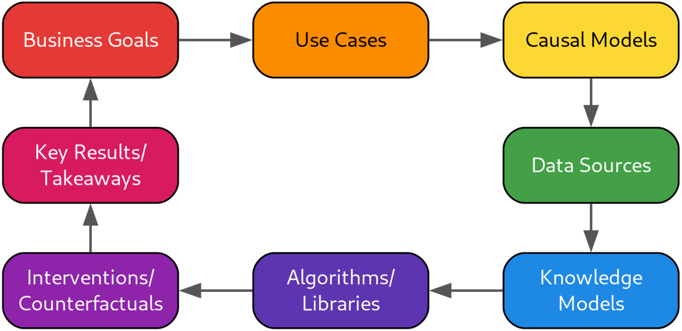
- Business Processes Around Data Modeling
    - ***Modeling Processes***
    - Roles

Digital Transformation in Business

- **Digital transformation** is the process of using digital technologies to:
  - Reduce costs
  - Improve revenues
- How?
  - **Integration of digital technology**
    - Embed digital tools (AI, cloud, IoT, automation) into all business areas
  - **Cultural and organizational change**
    - Encourage innovation, agility, and a data-driven mindset
  - **Customer-centric approach**
    - Use digital solutions (personalized experiences, AI-driven insights) to enhance customer engagement and satisfaction
  - **Process automation and optimization**
    - Streamline operations through AI, robots, and analytics
  - **Data-driven decision making**
    - Leverage big data, machine learning, and real-time analytics to make smarter, faster, and more strategic decisions

AI in Business

- How do you use (Causal) AI in Business?



SCIENCE ACADEMY

Step 1: What Are the Intended Outcomes?

- What is the **process** we are interested in analyzing?
- What will happen if a **course of action** is (or not) taken?
  - What outcomes are positive, negative, unacceptable, optimal?
- What **related data** can be combine / leverage?
- (Causal AI): What are the possible / feasible **interventions**?
  - What confounding factors might be correlated with outcomes and treatments?
  - What factors exist but we cannot accurately measure?

Step 2: What Are the Proposed Interventions?

- E.g., customer marketing for a product
- (Causal AI)
  - Should we introduce a **new product**?
  - Does bundling **multiple products** . . .
    - Improve sales?
    - Inhibit long-term sales?
  - Should advertisement focus on quality of our product vs other options?
  - Should we add more variations of the same product?
  - Should we discontinue the product?
  - . . .
  - Should we divest the product line?

Step 3: What Are the Proposed Interventions?

- **Identify confounders**
  - Difficult to understand the relationship between variables
  - Mute or inflate a relationship

Step 4: What Factors Create Changes?

- **Direct effect**
  - Effect introduced through intervention
- **Indirect effect**
  - Effect introduced by environment or as a byproduct of intervention
- **Total causal effect**
  - Effect of all factors in environment or model modifying outcome

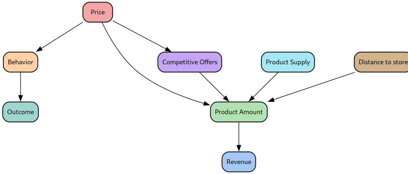
Step 5: Build Causal DAG

- **Causal models**
  - *Simplify complex systems* without losing key relationships
  - Focus on essential variables and their interactions
- **Visual models** (e.g., DAGs)
  - Help abstract complexity into interpretable formats
  - Highlight direction and strength of influence between variables
- **Simplicity**
  - Aids communication between technical and non-technical stakeholders
  - Promotes shared understanding and collaborative refinement
  - Reduces cognitive overload by excluding irrelevant details or noise
  - Guides data collection by identifying the most impactful variables
  - Supports hypothesis testing through counterfactual and intervention scenarios
- **Balance is key**
  - Too much simplicity loses insight
  - Too much complexity loses clarity

Marketing Example: Price Intervention

- Assume **price is our intervention**
- What happens to sales when we change the price often?
- What pricing interventions are optimal?
  - Should we increase the price and how much?
  - Should we decrease the price and how much?
  - Should price change in one-time or over time?
  - Should we develop individual pricing model for each customer?

- The Causal DAG is:



- Lots of confounders!

Step 7: Data Acquisition and Integration

- Use **data collected** for correlation-based ML
- Collect **data specifically** for causal AI
    - Treat, condition, transform data

Step 8: Model Modification

- **After initial DAG** is designed
  - Refine to reduce bias and improve reliability
  - Clarify variables as confounders, mediators, or outcomes
- **Avoid common pitfalls**
  - Do not control for mediators or effects, which can distort results
  - Control for direct and indirect confounders to prevent biased estimates
- **Test models** against technical and business objectives
- Determines necessary controls to **isolate causal effects**
  - Ensure causal assumptions align with data and domain logic
  - Improve model interpretability and predictive power

Step 10: Data Transformation

- Prepare data to match refined causal model
  - Clean, normalize, align data with model assumptions
- **Transformation data**
  - Map observed variables to DAG nodes
  - Encode categorical variables
  - Handle missing data (imputation or exclusion)
  - Normalize or scale values
- Control for bias and confounding
- **Goal**: Ensure data structure supports causal estimation

Step 11: Preparing for Deployment in Business

- Transition **from experimentation to decision-making processes**
  - Validate model against real-world business outcomes
  - Ensure stakeholders understand and trust causal logic and assumptions
- **Develop user-friendly interfaces** for business users
  - Automate data pipelines for timely updates and monitoring
- **Set up monitoring**:
  - Establish metrics for performance tracking and drift detection
  - Create feedback loops to refine and improve models post-deployment
- **Document and train**:
  - Provide clear model documentation for auditors and users
  - Train stakeholders on interpreting results and making informed decisions
- **Goal**: Deliver measurable business value

- Business Processes Around Data Modeling
    - Modeling Processes
    - ***Roles***

Why ML / AI Projects Fail?

- **AI projects fail** because they approach problems only from a ML perspective
  - Data scientists:
    - Focus on model accuracy and technical metrics
    - Work in isolation from domain experts, business users, and data engineers
  - Lack of understanding of problem context leads to irrelevant solutions
- **Black-box models** fail to solve real-world problems
  - Models are difficult to interpret or explain to stakeholders
  - Lack of transparency reduces trust and adoption
- **Data issues** undermine model performance
  - Poor quality, missing, or biased data
- **Deployment and integration challenges**
  - Models not easily integrated into production systems
  - Real-time constraints or infrastructure limitations
- **Lack of feedback loops**
  - No process for monitoring model performance post-deployment
- **Misalignment with business goals**
  - Models solve the wrong problem or answer irrelevant questions

How to Make ML / AI Project Succeed

1. **Create a hybrid team**
   - Organizations are complex
   - Single group lacks skills for difficult problems
   - Hybrid team:
     - Represents all business aspects
     - Uses collaborative framework
     - Communicates with a single language (e.g., DAGs)
     - Team size varies by company and project complexity
2. **Meet regularly to ensure continuity**
3. **Find an executive sponsor**
   - Understands project's goals and potential
4. **Initial pilot**
   - Small team for targeted problem
   - Demonstrate AI approach merit

Roles in Hybrid Teams

Role	Responsibilities
Business Strategists	Align modeling with business goals Sponsor projects Communicate insights to stakeholders
Subject-Matter Experts	Provide domain expertise Identify relevant variables Validate DAGs
Data Experts	Source and clean data Map data to model variables Handle missing values
Data Scientists	Construct and validate DAGs Apply causal inference methods Simulate decisions
Software Developers	Build tools and interfaces Create data pipelines
IT Professionals	Provide infrastructure and governance Ensure model execution Integrate with enterprise systems
Project Managers	Coordinate collaboration and timelines Manage documentation Ensure alignment with strategic goals

Steps for a Hybrid Team Project

- **Establish a phased / collaborative approach**
  - Align strategic goals with technical efforts
  - Emphasize early stakeholder engagement and shared ownership
- **Define team goals**
  - Use SMART objectives aligned with business strategy
- **Target a project**
  - Choose a bounded, feasible, high-value use case
  - Prioritize early wins for trust and momentum
- **Define the hypothesis**
  - Build a preliminary DAG with experts' input
- **Incremental model development**
  - Build the model in small stages
  - Iterate with regular feedback

The Importance of Explainability

- **Understanding AI-based models** is growing in importance
  - How do ML models make decisions?
  - Can they be trusted?
  - Are they biased?
- AI systems need to foster trust, transparency, and user confidence
- Management faces
  - Loss of trust, regulation violations, fines, lawsuits
  - E.g., false negative in medical screening for cancer
- **Humans in the loop**
  - Ensure interpretation and context are considered

Techniques for Interpretability

- **Local Interpretable Model-agnostic Explanations** (LIME)
  - Focus on a single prediction
  - Approximate the model locally with an interpretable model
  - Perturb input data and observe changes in predictions
- **Partial Dependence Plots** (PDP)
  - Show marginal effect of one or two features on predicted outcome
  - Vary value of one feature while keeping others constant
- **Individual Conditional Expectation** (ICE)
  - Show relationship between a feature and prediction for individual instances
- **SHapley Additive exPlanations** (SHAP)
  - Quantify the contribution of each feature to a specific prediction

Causal AI in Interpretable AI

- Causal AI helps understand causes, effects, and solutions
  - Causal DAGs present variables, relationships, strengths
  - Counterfactual analysis predicts outcomes of actions
  - Model output is understandable to humans, non-experts
  - Hybrid teams (technical + domain experts) enhance context awareness, reduce blind spots
  - Removing confounding variables prevents skewed causal estimates

The Future of Causal AI

- **Causal AI**:
  - Moves from academia to commercial use
  - Departs from purely data-driven AI systems
  - Reflects human thinking, analysis, decision-making
- Causal, traditional AI, deep learning, and generative techniques will merge