UNIVERSITY OF MARYLAND

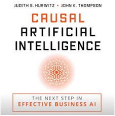
MSML610: Advanced Machine Learning

8.1: Intro to Causal AI

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

References:

- Hurwitz, Thompson: Causal Artificial Intelligence, 2024



- ***Causal AI***
  - Why Causal AI?
  - The Ladder of Causation
  - Correlation vs Causation Models

- Causal AI
  - ***Why Causal AI?***
  - The Ladder of Causation
  - Correlation vs Causation Models

Big Data and Traditional AI

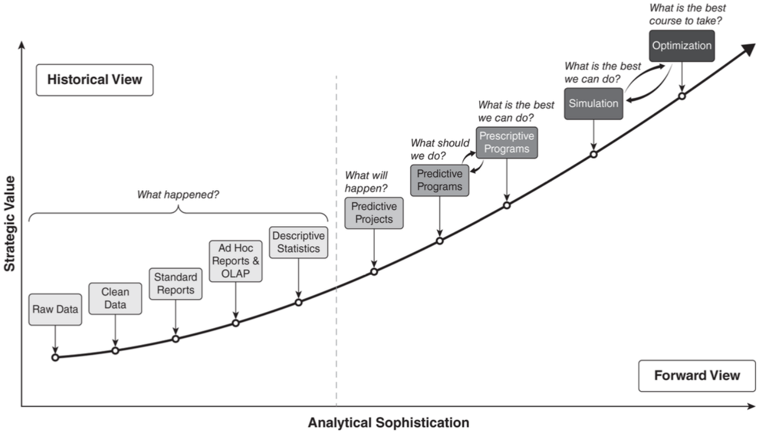
- For the past 10 years, **focus of analytics** on:
  - Organize and analyze massive amount of data
  - Data analytics (dashboards, models, reports)
  - Run machine learning on data
- Problems with **traditional AI**
  - Predicts based on observed correlations
  - Can't explain why an outcome occurred
- **AI in decision making**
  - Understand impact of decisions
  - E.g., *"What happens if a product price is reduced by 10%?"*
    - Will more customers buy?
    - If revenue decreases, what to do?
    - Why are customers leaving? Quality issue? Emerging competitor?

What Are Data Analytics?

- **Collections of data**
  - Aggregated, organized data sets for analysis
  - E.g., customer purchase histories in a CRM system
- **Descriptive statistics**
  - Summary metrics: mean, median, mode, standard deviation
  - E.g., average sales per quarter to understand trends
- **Historical reports**
  - Examination of *past performance*
  - E.g., monthly sales reports for past fiscal year
- **Dashboards**
  - Visual displays of key metrics for insights
  - E.g., dashboard showing quarterly revenue, expenses
- **Models**
  - Statistical representations to *forecast, explain phenomena*
  - E.g., model to anticipate customer churn based on behavioral data

Data Analytics Sophistication

Business Question	Methodology
What happened?	Descriptive statistics
What will happen?	Predictive models
What should we do?	Prescriptive programs
What's the best we can do?	Simulation + optimization



AI / ML Explainability

- **Regulators** require you to defend ML/AI decision results
  - E.g., hiring decisions, policy setup
- E.g., neural networks are "black boxes"
  - Lack of explainability
    - Humans can't understand input-to-conclusion process (easily or at all)
    - Can't explain decisions to shareholders
  - Bias
    - E.g., using age, race, sex as features can introduce bias
- Drawbacks: organizations can:
  - **Be fined** by authorities
  - **Face backlash** from customers and activists
- Solution:
  - **Explainable AI** allows users to:
    - Comprehend
    - Explain
    - Trust machine results

Correlation is Not Causation!

- **Correlation** is a statistical method for understanding relationships between data
  - Pros
    - Use past outcomes to predict future outcomes by finding patterns and anomalies
  - Cons
    - Doesn't explain the cause
    - Variables may move together due to coincidence or a hidden factor
- **Causation** explains how changing one variable influences the other
  - Cannot be concluded from correlation alone
- **Data does not understand causes and effects**
  - Only humans can identify variables and relationships based on context
  - Without causation, you can't make intelligent decisions

Causal AI

- **Understands the why**
  - Determines cause-and-effect between variables
  - E.g., whether a marketing campaign increased sales
- **Identify interventions**
  - Identifies variables and interventions to change outcomes
  - E.g., which lifestyle changes reduce blood pressure
- **Predicting counterfactuals**
  - Hypothesizes outcomes under different circumstances
  - E.g., student grades if they attended a different school
- **Avoiding bias**
  - Traditional AI biased by training data and ignored variables
  - Ensure fairness by accounting for confounding variables
- **Improving decision-making**
  - Provides understanding of relationships for better decisions
  - E.g., improve supply chain by understanding impact of decisions on logistics

Causal AI vs Traditional AI

- **Current AI** uses correlation to:
  - Analyze data
  - Identify patterns
  - Make predictions
- **Models depend on data quality**
  - Biased or unclean data $\implies$ poor model
- *"The next revolution of data science is the science of interpreting reality, not of summarizing data"* (Judea Pearl, 2021)

- Causal AI
  - Why Causal AI?
  - ***The Ladder of Causation***
  - Correlation vs Causation Models

The Ladder of Causation

- Pearl provided a 3-level framework for understanding causality
  1. Association
  2. Intervention
  3. Counterfactuals

Level	**Symbol**	**Activity**	**Typical Questions**
1. Association	$\Pr(Y\|X)$	Observing	What is?
2. Intervention	$\Pr(Y\|do(X), Z)$	Intervening	What if?
3. Counterfactuals	$\Pr(Y_X\|x', y')$	Imagining	Why?

Rung 1: Association

- **Question**: *"How would seeing X change our belief in Y?"*
- **In math terms**: $\Pr(Y|X)$
- **Activity**
  - It is just "passive observation" to determine if X, Y are related
  - Traditional AI and ML is based on this
  - In the best case, it's Bayesian update
  - In the worst case, it's some hacked up overfitted model
- **Example**
  - *"The tree has green leaves during spring"*
  - *"What does a symptom tell you about a disease?"*
  - *"What does a survey tell you about the election results?"*

Rung 2: Intervention

- **Question**: *"What happens to Y if you do X?"*
- **In math terms**: $\Pr(Y|do(X), Z)$
- **Activity**
  - Understand the impact of an action X on Y under conditions Z
  - Association is just about observations
  - Interventions involve "doing something" and need a causal model
- **Example**
  - *"Spring makes tree leaves turn green"* (vs *"tree has green leaves"*)
  - *"Why did the headache go away?"*
    - "Because the pain reliever" or "Because you ate food after skipping lunch"
  - *"If you take aspirin, will your headache be cured?"*
  - *"What if you ban sodas?"*

Rung 3: Counterfactuals

- **Question**: *"Was X that caused Y?"*
- **Symbol**: $\Pr(Y_X | x', y')$
- **Activity**:
  - Imagine what will happen if facts were different
  - Predicting an outcome is the highest form of reasoning
    - It requires to understand relationships between cause and effect
- **Example**
  - Scientific experiments: *"What if we give a child an adult dose of a drug?"*
  - Litigation: *"What would the jury conclude?"*
  - Marketing: *"Why did my marketing campaign fail to generate sales?"*

- Causal AI
  - Why Causal AI?
  - The Ladder of Causation
  - ***Correlation vs Causation Models***

Correlation vs Causation Model

- **Correlation** = identify how variables are related to each other
- **Causality** = determine whether one variable causes another variable
  - Both:
    - Accept inputs and transform them to compute predictions
    - Identify how variables are related to each other
  - Correlation-based AI works well when there is abundant historical and observational data
  - Causal-based AI first creates a business-focused model before integrating data

Correlation-Based Model Process

- **Correlation-based AI** is “data first”
  - The more data collected the better
- **Modeling process**
  - Acquire data
  - Integrate and clean data
  - Exploratory data analysis (EDA)
  - Feature engineering
  - Build and test models
  - Deploy models in production
- **Many AI projects fail because**
  - Cultural and organizational issues
  - Models are opaque and lack explainability
  - Spurious correlations
  - Missing articulating “what's the goal of doing ML?”

Causation-based Model Process

- **Causal AI** is "model first"
  - Understand business question before ingest and transform the data
- **Modeling process**
  - What is the intended outcome?
  - What is the proposed intervention?
  - What are the confounding factors?
  - What are the effecting factors?
  - Create a model graph or diagram
  - Data acquisition
  - ...