The Impact of Machine Learning on Economics

CHEN Chen, LI Jiayue, WANG Xuanyi, XU Hongyi and ZHENG Xing

June 12, 2018

Introduction

About the author

- The Economics of Technology Professor
- Co-organizer of Productivity and Information Technology/Digitization, National Bureau of Economics Research
- John Bates Clark Medal, 2007
- 38 papers in top journals

4 Main Parts

- Definition: the past
- Review: the present use
- Outlook: the casual inference
- Conclusion: the exploration of new fields

5 Highlighted Themes

- Identification of model forms
- Flexible and data-driven model selection
- Outsourcing of some hard work
- An algorithm to be further modified
- Change in research field

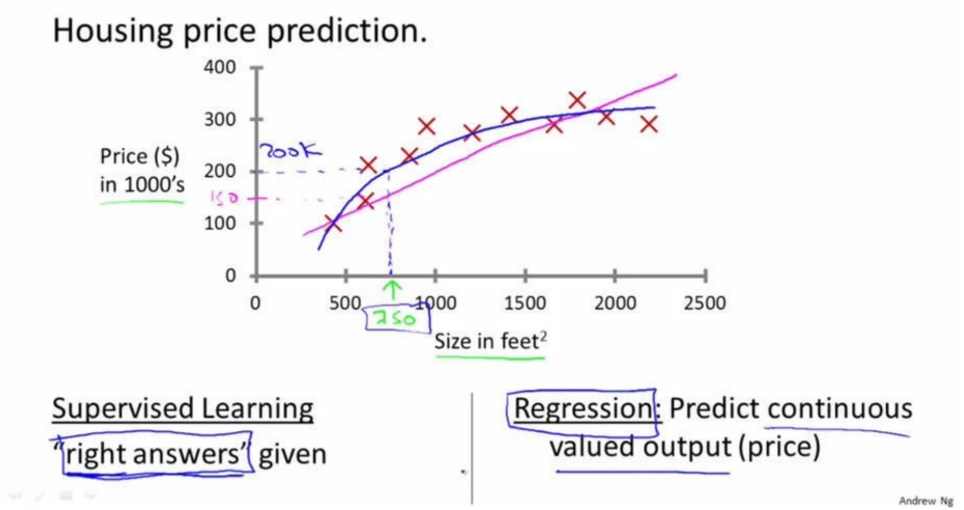
What is Machine Learning
What are Early Use Cases

What is Machine Learning?

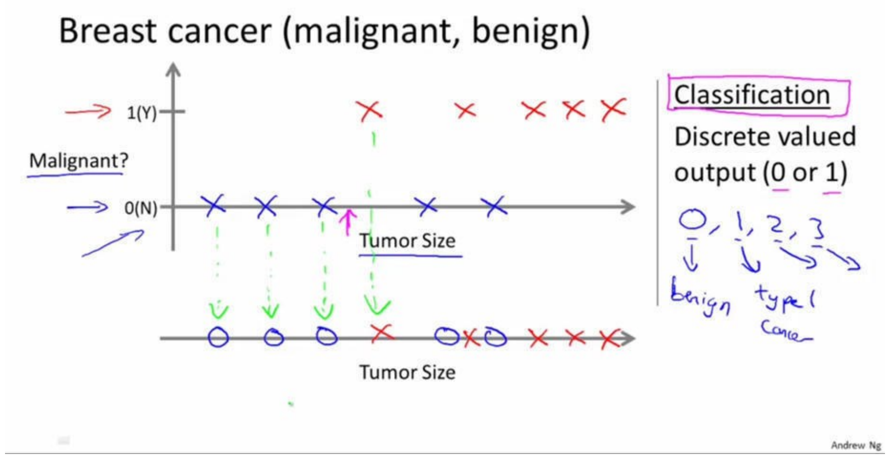
- It is harder than one might think to come up with an operational definition of ML.
- Indeed, one could devote an entire article to the definition of ML, or to the question of whether the thing called ML really needed a new name other than statistics, the distinction between ML and AI, and so on.
- Starting from a relatively narrow definition of machine learning, machine learning is a field that **develops algorithms designed to be applied to datasets**, with the main areas of focus being **prediction (regression), classification, and clustering or grouping tasks**.

What is Machine Learning? Supervised ML

- **Prediction**: Supervised machine learning typically entails **using a set of features or covariates (X) to predict an outcome (Y)**.
  - $\mu(x) = \mathbb{E}[Y|X = x]$



- **Classification**: the goal is to accurately classify observations.

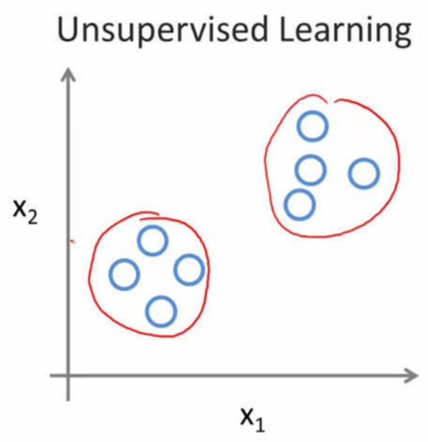


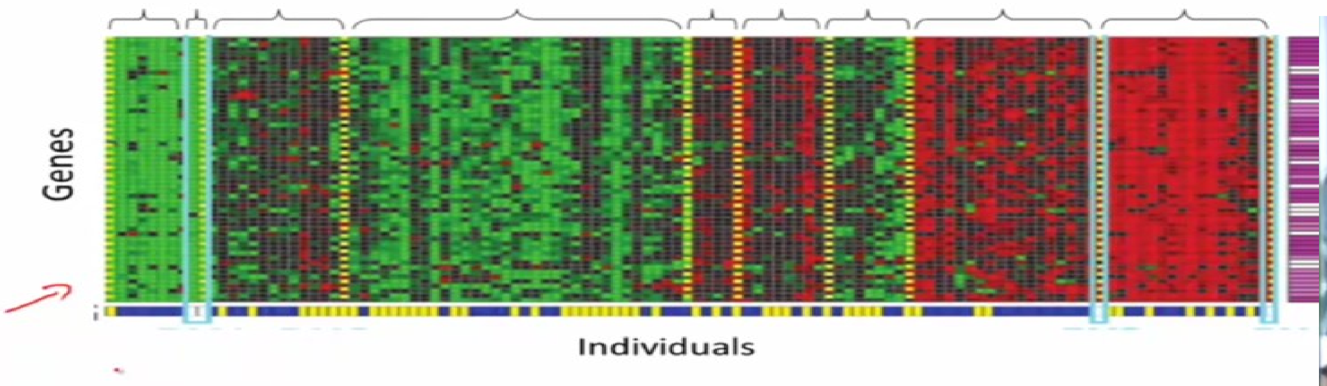
- $\Pr(Y = k|X = x)$ for each of $k = 1, \cdots, L$ possible realizations of K.

A variety of ML methods for supervised learning: regularized regression, random forest, regression trees, support vector machines, neural nets, matrix factorization, and many others, such as model averaging.

What is Machine Learning? Unsupervised ML

- It is commonly used for video, images and text.
- There are a variety of **techniques** available for unsupervised learning, including k-means clustering, topic modeling, community detection methods for networks, and many more.
- Goals: Clustering & Dimensionality Reduction
- Examples: Genes / “I discovered cats!”





[Source: Su-In Lee, Dana Pe’er, Aimee Dudley, George Church, Daphne Koller]

Andrew Ng

What is Machine Learning? Unsupervised ML

- These tools are very useful as an **intermediate step in empirical work** in economics.
    - They provide a data-driven way to find similar newspaper articles, restaurant reviews, etc., and thus **create variables that can be used in economic analyses**. For example, if an analyst wishes to estimate a model of consumer demand for different items, it is common to **model consumer preferences** over characteristics of the items.
    - Unsupervised learning can also be used to **create outcome variables**. For example, Athey *et al*. (2017d) examine the impact of Google's shutdown of Google News in Spain on the types of news consumers read.

What Are Unique Features of Cross-Sectional Econometrics vs. Other Branches of Statistics?

- **Framework and language for causality**
- **Causal inference from observational data**
  - Theory and PRACTICE
- **Structural models to do counterfactuals for environments that have never been observed**
- **Emphasis on interpretable (~causal) models**
- **Relatively little emphasis on systematic model selection in applied micro-econometrics**
  - Even in environments where theory does not motivate functional forms
- **Emphasis on standard errors for a pre-specified models**
  - Estimators must have established properties

What We Say vs. What We Do (Econometrics)

- **What We Say**
  - Causal inference and counterfactuals
  - God gave us the model
  - We report estimated causal effects and appropriate standard errors
  - Plus a few additional specifications for robustness

- **What we do**
- Run OLS or IV regressions
  - Try a lot of functional forms
  - Report standard errors as if we ran only one model
  - Have research assistants run hundreds of regressions and
  - pick a few “representative” ones
- Use complex structural models
  - Make a lot of assumptions without a great way to test them

What We Say vs. What We Do (Econometrics)

- **What We Say**
  - ML = Data Science, statistics
  - Is there anything else?
  - Use language of answering questions or solving problems, e.g. advertising allocation, salesperson prioritization
  - Aesthetic: human analyst does not have to make any choices
  - All that matters is prediction
- **What we do**
  - Use predictive models and ignore other considerations, e.g. causality
  - Wonder/ worry about interpretability/reliability/robustness/adaptability, but have little way to conceptualize or ask algos to optimize for it
  - Limited conceptual framework for feedback effects, equilibrium, etc.

Differences: Causal Effect Estimation vs. Prediction

- ML (prediction) :
  - There is typically a tradeoff between **expressiveness** of the model and risk of **over-fitting**, which occurs when the model is too rich relative to the sample size.
- Econometrics and empirical work in economics:
  - the tradition has been that the researcher **specifies one model, estimates the on the full dataset**, and relies on statistical theory to estimate **confidence intervals** for estimated **parameters**.
- Eg. IV estimation

- Key concerns in different approaches
- A second concern is whether the assumptions required to “identify” a causal effect are satisfied, where in econometrics we say that a parameter is identified if we can learn it eventually with infinite data
- Economists also build **more complex models** that incorporate both behavioral and statistical assumptions in order to estimate the impact of counterfactual policies that have never been used before.

Applications of Prediction Methods to Policy in Economics

- Stop-and-frisk
- Bail
- Allocate inspector resources

Future Application of Prediction to Policy in Economics

- Fairness and nondiscrimination
- Stability and robustness
- Manipulability

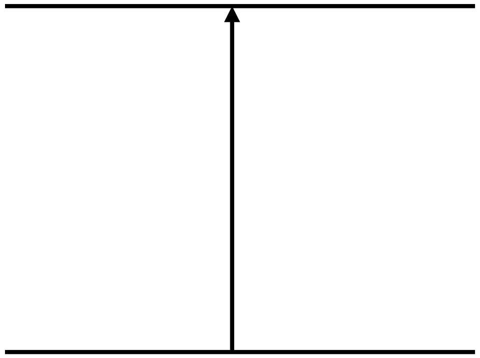
Causal Inference

Neyman-Rubin Counterfactual Framework

$$\text{Observed Effect} = \mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 0]$$

$$\text{Treatment Effect} = \mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 1]$$

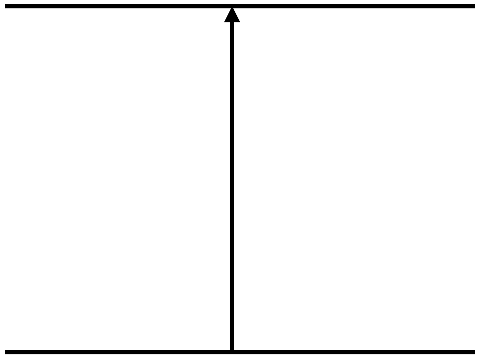
Hotel Example



Occupancy rate UP

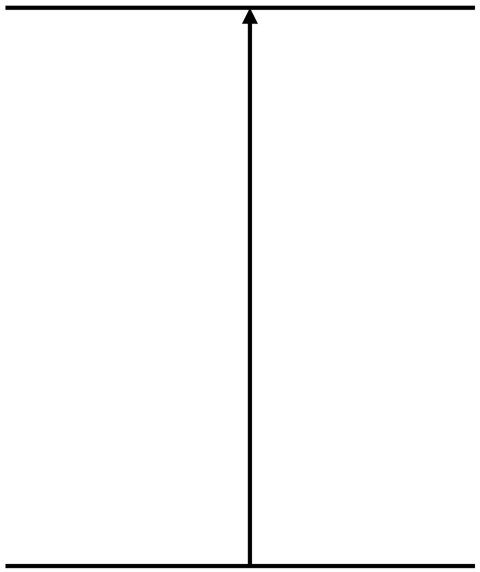
Price UP

Hotel Example



Occupancy rate UP

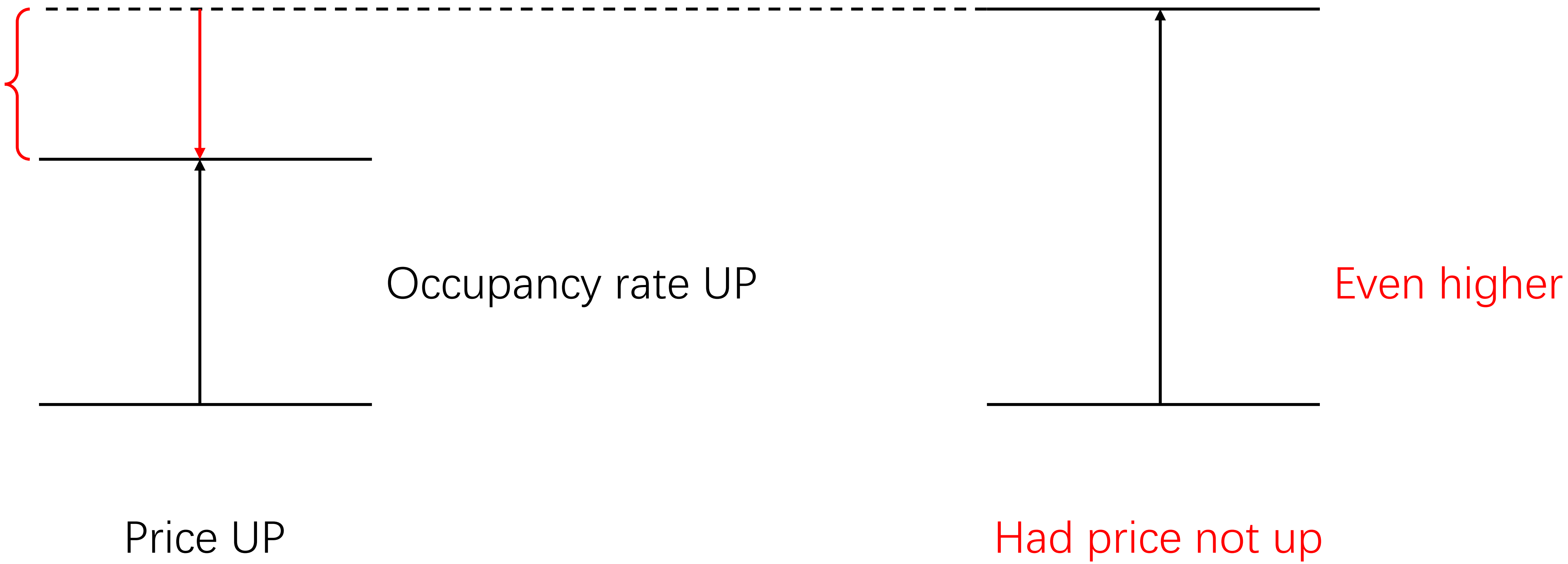
Price UP



Even higher

Had price not up

Hotel Example



Average Treatment Effect

- 1900s: Semi-parametric methods

  e.g. Propensity Score Matching

- 1900s: Semi-parametric methods

  e.g. Propensity Score Matching

- ML: Regularised regression
  - Belloni, Chernozhukov and Hansen (2014)
  - Athey, Imbens and Wager (2016)
  - Chernozhukov *et al*. (2017)

Heterogeneous Treatment Effects

- Simple patterns of heterogeneity
  - Causal Trees: sample partition
  - Application: Davis and Heller (2017)

- Simple patterns of heterogeneity
- Prediction based on covariates
  - $\mu(x) = \mathbb{E}[Y^{(1)} - Y^{(0)} | X = x]$
  - Causal Forest, Local GMM, Neural Nets $\cdots$

Optimal Policies

- Contextual Bandits: mapping from covariates to treatment
  - State: Covariates
  - Action: Treatment assignment
  - Reward: Heterogeneous treatment effect

Robustness and Supplementary Analysis

Data-driven Model Selection Placebo Test

Panel Data and Difference-in-Difference

“ML-based” DID

- DID: Algorithm average
- Synthetic Control Method: Weighted average
- Doudchenko and Imbens (2016): Generalised linear function

Matrix Completion

Y	Time	Covariates	Treatment
5	1	1	0
6	1	5	0
7	1	2	0
30	2	7	1
31	2	8	1
35	2	4	1

Matrix Completion

Y	Time	Covariates	Treatment
5	1	1	0
6	1	5	0
7	1	2	0
???	2	7	0
???	2	8	0
???	2	4	0

Structural Model

- Markov Chain Monte Carlo
- Factorisation of Matrix (Ruiz, Athey and Blei, 2017; Athey *et al.*, 2017)

Expectation

Data Scale

- Declination in Data Cost

  --“A perfect research assistant”

- Increased Interdisciplinary Research

- Increased Interdisciplinary Research
- Economists Engagement with Firms
    - “Economists as Engineers”

- Research Organization
    - Returns to Scales in Data Processing

- Research Organization
    - Returns to Scales in Data Processing
- Education