KAIST Summer Session 2018

Module 1. Research Design for Data Analytics

The Art of Prescriptive Analytics

KAIST College of Business

Jiyong Park

12 July, 2018

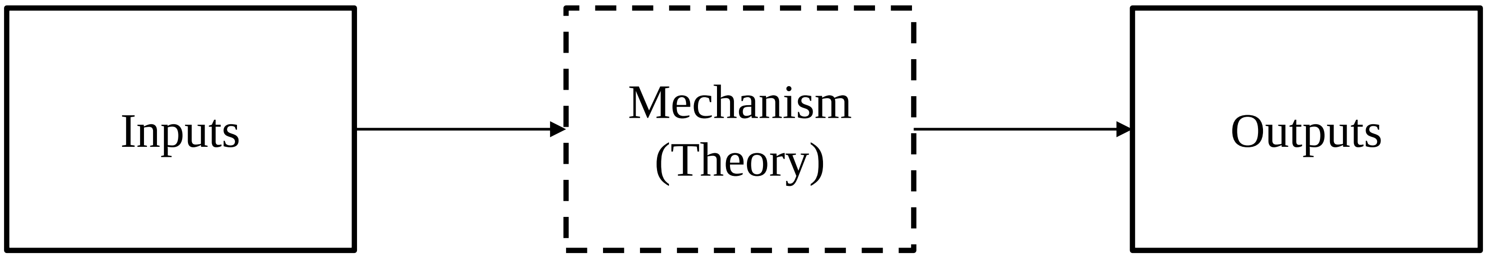
Predictive Analytics

$\neq$

Prescriptive Analytics

Framework of Prescriptive Analytics

Input-Output Framework



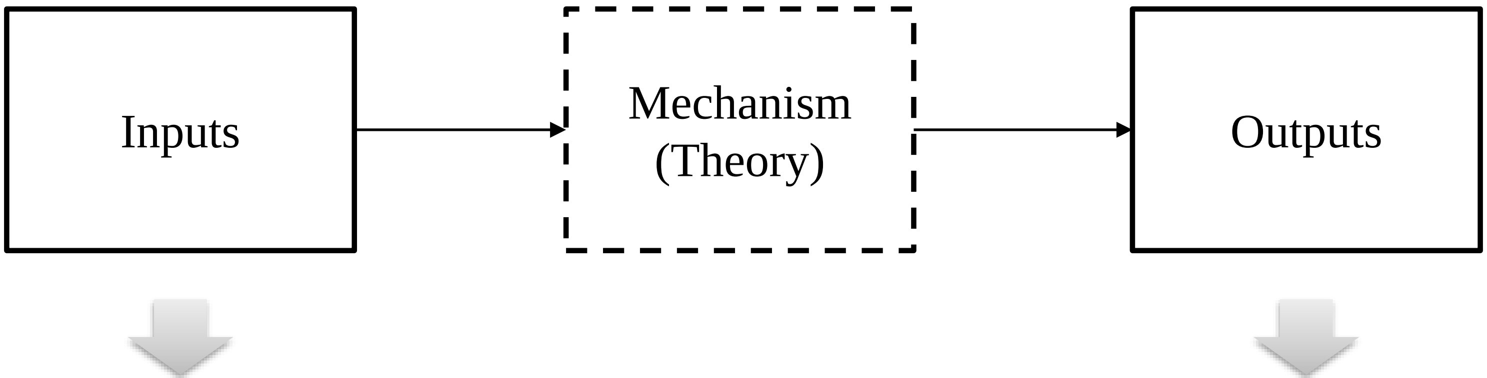
If you are interested in contemplating some intervention in the *inputs*, identification strategy for causal inference would be the right tool.

→ Intervention-oriented research

If you are interested in obtaining the best (or precise) *output*, predictive analytics would be the right tool.

→ Solution-oriented research

Input-Output Framework



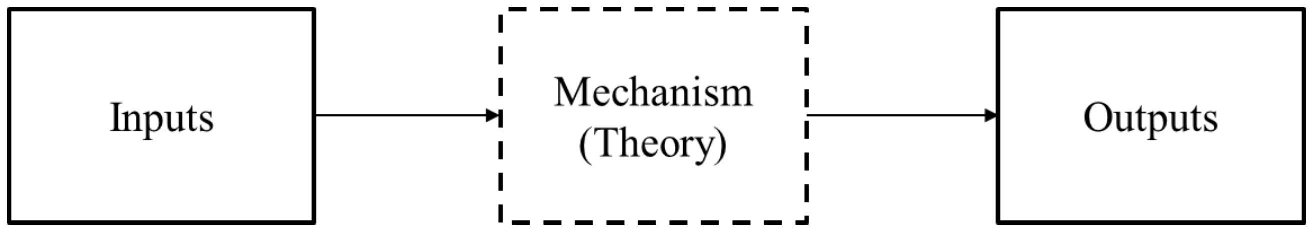
→ Intervention-oriented research

→ Solution-oriented research

Prescription is to design the system as a whole
(e.g., individuals, business organizations, society)

Framework of Prescriptive Analytics

- Input-Output Framework



  - (known) Payoff function: $\pi(Y, X)$
  - (unknown) Output variable: Y
  - (known) Input variable: X

- Prescription is to determine X in order to maximize $\pi(Y, X)$.

$$\frac{\partial \pi(Y,X)}{\partial X} = \frac{\partial \pi}{\partial X}(Y) + \frac{\partial \pi}{\partial Y}\frac{\partial Y}{\partial X}$$

Prediction *Causation*

Framework of Prescriptive Analytics

- Prescriptive analytics aims at maximizing the payoff function.
  - Causal inference aims at $\frac{\partial Y}{\partial X}$
  - Predictive analytics aims at Y

$$\frac{\partial \pi(Y,X)}{\partial X} = \frac{\partial \pi}{\partial X}(Y) + \frac{\partial \pi}{\partial Y}\frac{\partial Y}{\partial X}$$

- However, the resources are often limited in reality.
  - Firms want to maximize the sales through business analytics, but their budges are limited.
  - Hospitals want to maximize the survival rate of anticancer treatments through healthcare analytics, but the dose ranges are restricted for safety.

Framework of Prescriptive Analytics

- Prescriptive analytics aims at maximizing the payoff function, with resource constraints.
  - ➢ It requires the concept of optimization.

Prediction *Causation*

$$Max \quad \frac{\partial \pi(Y,X)}{\partial X} = \frac{\partial \pi}{\partial X}(Y) + \frac{\partial \pi}{\partial Y}\frac{\partial Y}{\partial X}$$

$$subject\ to \qquad X < c$$

$$\mathrm{Y}, X > 0$$

Optimization

- The Art of Prescriptive Analytics
  - ➢ Predictive Analytics + Causal Inference
  - ➢ Predictive Analytics + Optimization

Prescriptive Analytics

= Predictive Analytics + Causal Inference

Role of Causal Inference in Prescriptive Analytics

- Someone might advocate “the end of theory.” (Anderson 2008)

CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE

- Why should we understand the causal relationship and theory behind the observed patterns?
  (1) Guidance to prediction
  (2) Long-term stable prediction
  (3) Explanation and interpretability
  (4) Better system design

Chris Anderson. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *WIRED*
https://www.wired.com/2008/06/pb-theory/

(1) Causation Guides the Design Principle for Prediction

- Example: Predicting teacher value-added and hiring good teachers
  - Teacher value-added is defined as their effect on students' test scores.
  - Chalfin et al. (2016) show that schools can achieve the gain in students' learning (test score) from using machine learning to deselect the predicted bottom 10 percent of teachers and replace them with average quality teachers.
  - Before this prediction, however, we should answer the fundamental question: *Does teacher value-added really matter?*

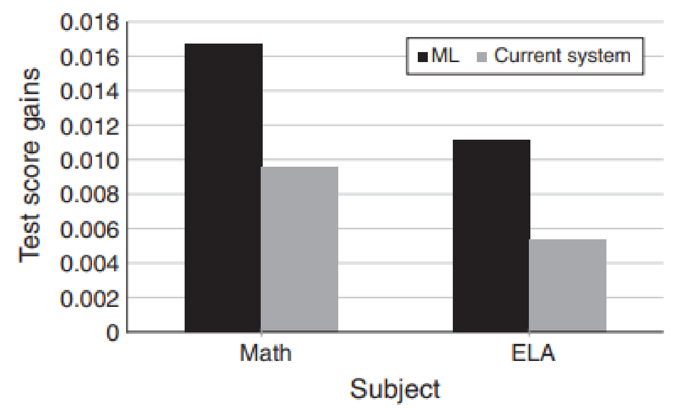
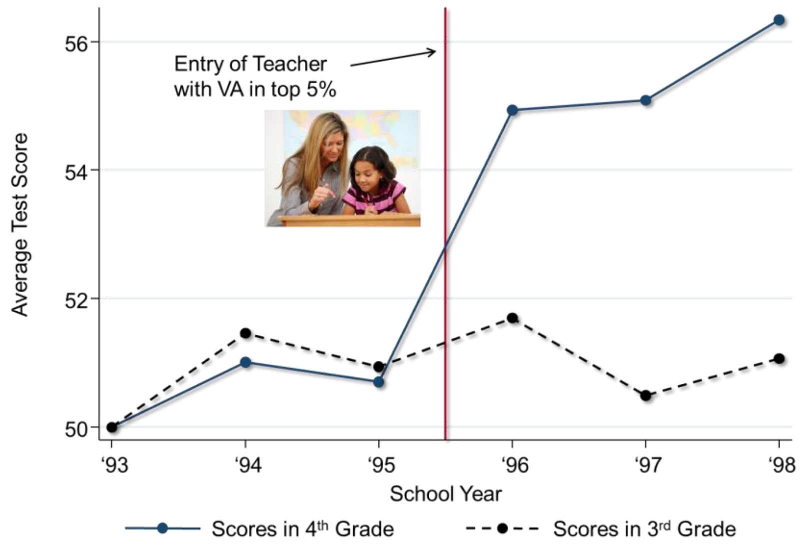


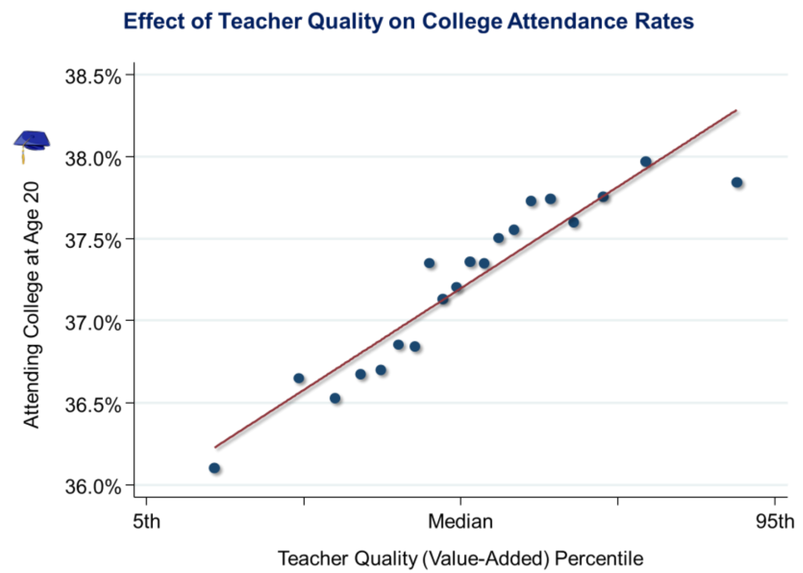
FIGURE 2. CHANGE IN TEST SCORES FROM ML VERSUS STATUS QUO SELECTION OF TEACHERS

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J. and Mullainathan, S., 2016. Productivity and Selection of Human Capital with Machine Learning. *American Economic Review*, 106(5), pp.124-127.

(1) Causation Guides the Design Principle for Prediction

- Example: Effect of teacher value-added on student long-term outcomes
  - ➢ Chetty et al. (2014) use a quasi-experiment of teachers turnover.
  - ➢ The authors find that assigning a student to a higher value-added (VA) teacher raises not just test scores but also long-term outcomes.



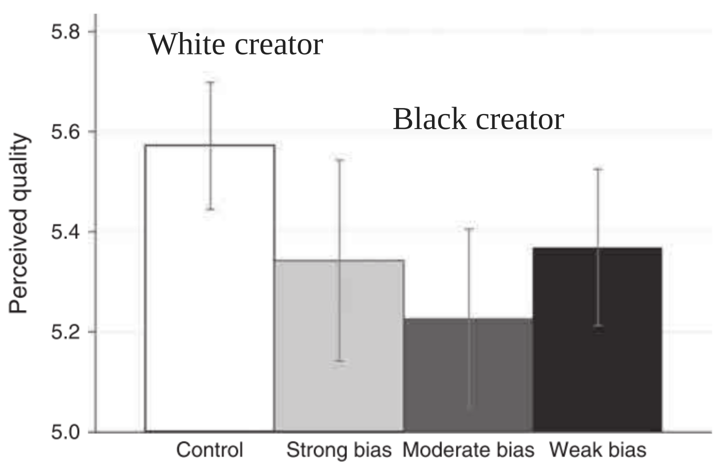


Chetty, R., Friedman, J.N. and Rockoff, J.E., 2014. Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9), pp.2633-79.

(1) Causation Guides the Design Principle for Prediction

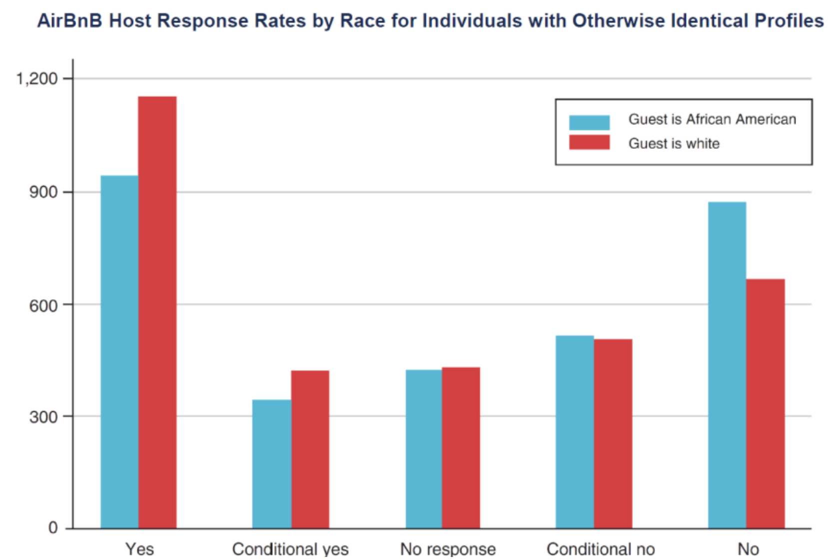
- Example: Racial discrimination on online platforms
  - [Theoretical background] Empirical evidence on racial discrimination in crowdfunding (Younkin and Kuppuswamy 2017) and sharing economy (Airbnb; Edelman et al. 2017)

Figure 3. Effect of Treatment on Quality



Note. Bars indicate 95% confidence intervals.

(Younkin and Kuppuswamy 2017)



(Edelman et al. 2017)

Younkin, P. and Kuppuswamy, V., 2017. The Colorblind Crowd? Founder Race and Performance in Crowdfunding. *Management Science*. forthcoming
Edelman, B., Luca, M. and Svirsky, D., 2017. Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics*, 9(2), pp.1-22.

(1) Causation Guides the Design Principle for Prediction

- Example: Racial discrimination on online platforms
  - ➢ [Theoretical background] Economic theory of discrimination (Park and Kim 2018)

Analytical model

Funder's utility function

$$U = F(q, N) - N$$

Information asymmetry and signaling

$$s = q + u \quad s.t.\ q \sim N(\alpha, \sigma_q^2), u \sim N(0, \sigma_u^2)$$

$$\widehat{U} = F(\hat{q}, N) - N \quad s.t.\ \hat{q} = E(q|s)$$

Racial discrimination on crowdfunding

$$\widehat{U} = F(\hat{q}, N) - N(1 + d_i)$$

$$s.t.\ \hat{q} = E(q|s), \qquad q \sim N(\alpha_i, \sigma_q^2)$$

Theoretical proposition

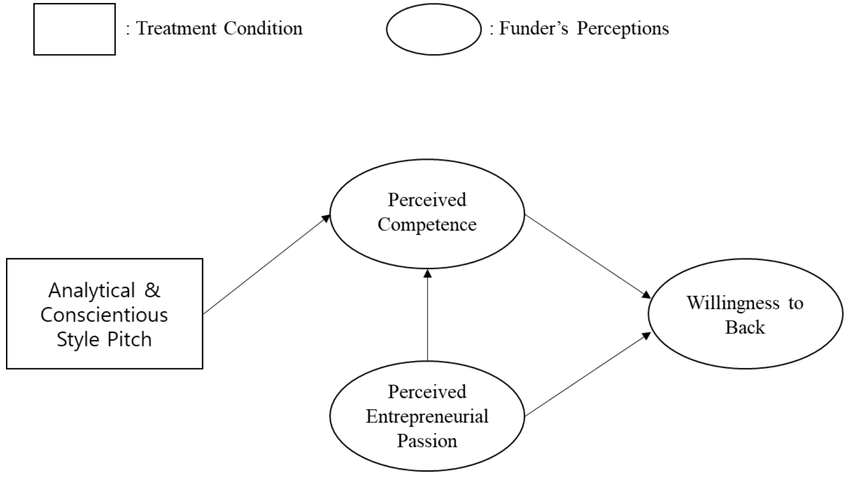
Proposition 1. *As the quality of signals increases (σ_u^2 decreases), the funder's racial bias will decrease in crowdfunding.*

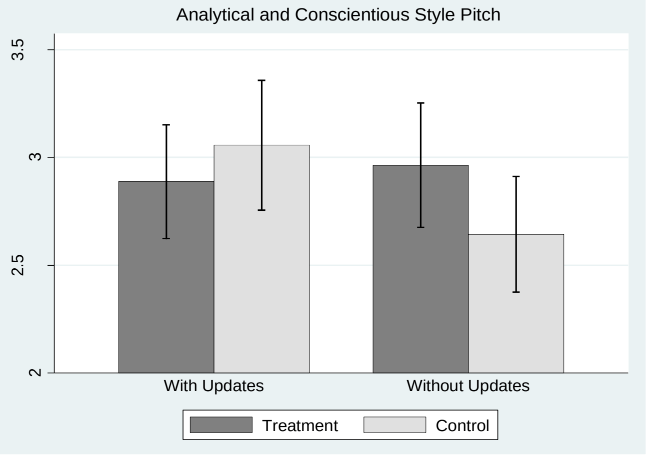
Proposition 2. *If the improvement in signal quality could turn away racial bias, statistical discrimination will be dominant in crowdfunding, rather than taste-based discrimination.*

Park, J. and Kim, J., 2018. Fixing Racial Discrimination through Analytics on Online Platforms: A Neural Machine Translation Approach. *KAIST Working Paper.*

(1) Causation Guides the Design Principle for Prediction

- Example: Racial discrimination on online platforms
  - ➢ [Causation-based insight for design principle] Park et al. (2018) find that the linguistic style plays a critical role in crowdfunding performance, using a combination of observational and experimental data.

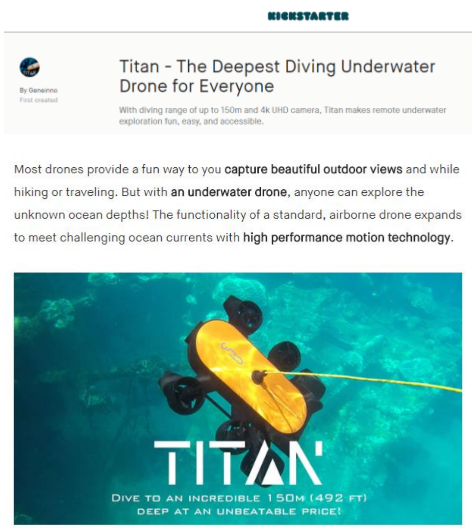




Park, J., Kim, J., Cho, D. and Lee, B., 2018. Pitching with Style: The Role of the Entrepreneur's Video Pitch on Online Crowdfunding. KAIST Working Paper.

(1) Causation Guides the Design Principle for Prediction

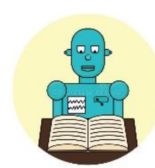
- Example: Racial discrimination on online platforms
  - [Prescriptive analytics] Leveraging machine learning techniques to fix racial discrimination on crowdfunding (Park and Kim 2018)
  - Basic intuition of our data-analytic framework: Style transfer in text



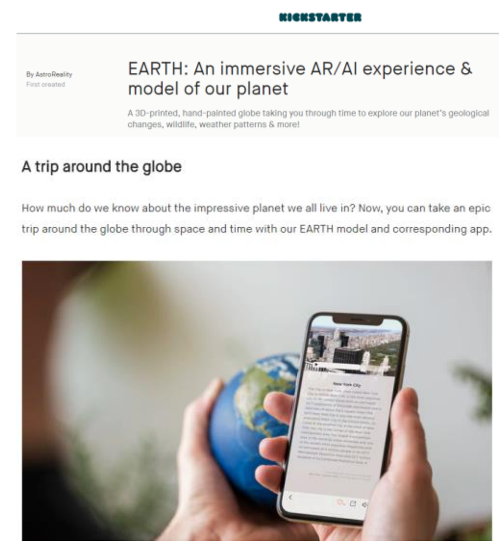
Fail-style text (Black creator)

Style transfer for augmenting signals

Learn



Predict

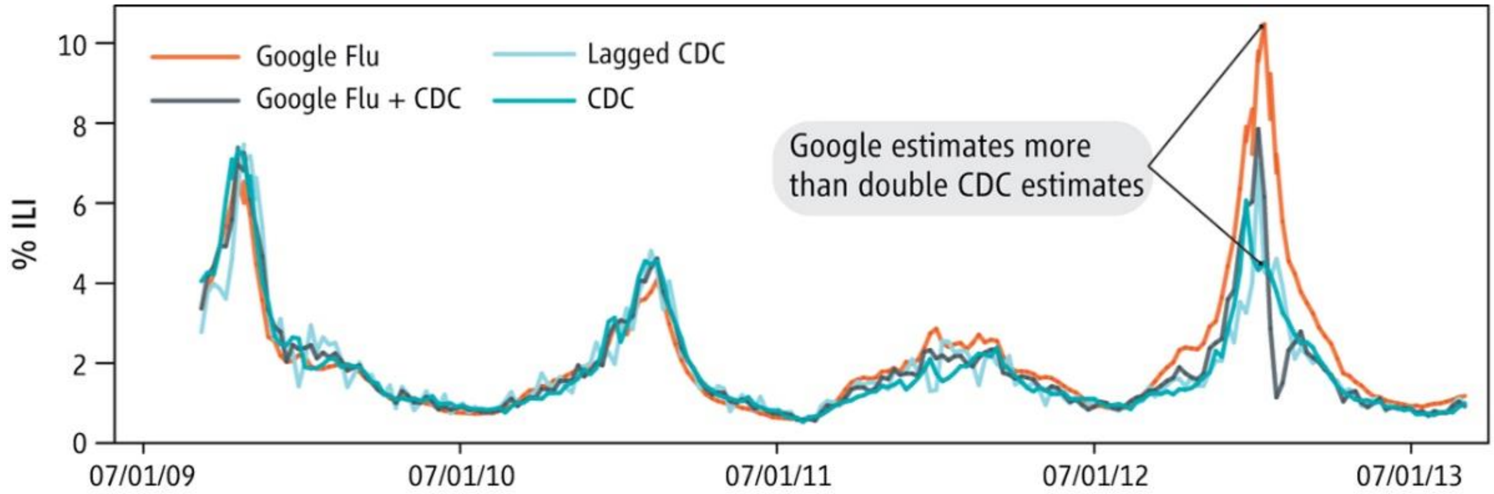


Success-style text (Black creator)

Park, J. and Kim, J., 2018. Fixing Racial Discrimination through Analytics on Online Platforms: A Neural Machine Translation Approach. *KAIST Working Paper.*

(2) Causation Might Guarantee the Long-term Prediction

- Predictive models may not be stable over time, possibly due to a range of volatility. (e.g., algorithmic change, sample change, unobserved confounding)
    - ➢ (Example) Google Flu and change in Google search engine (Lazer et al. 2014)



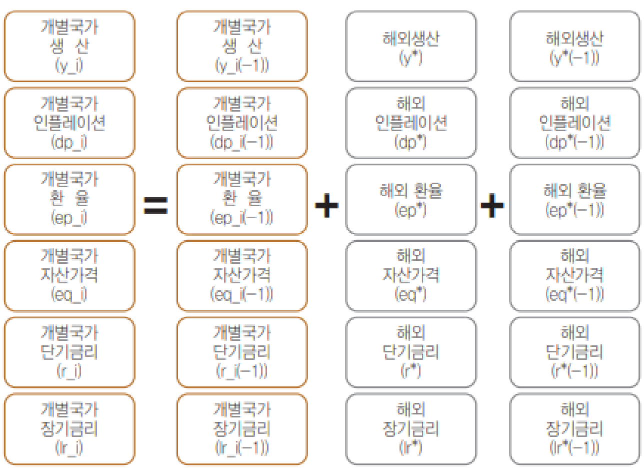
Lazer, D., Kennedy, R., King, G. and Vespignani, A., 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), pp.1203-1205.

KAIST COLLEGE OF BUSINESS

(2) Causation Might Guarantee the Long-term Prediction

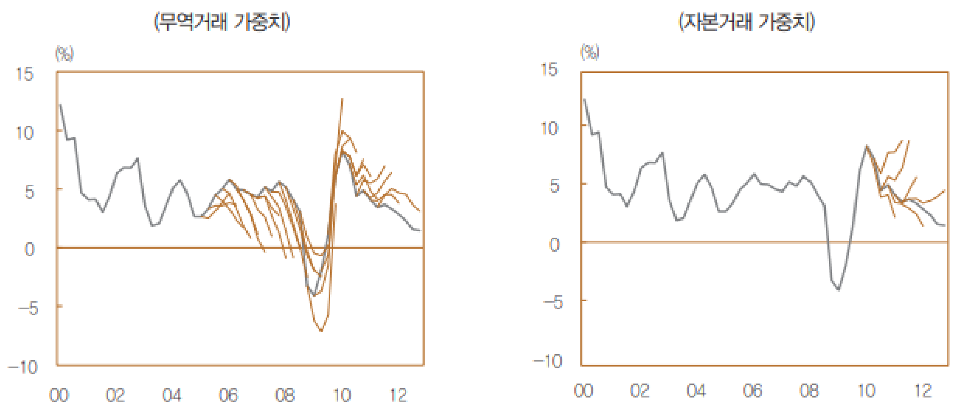
- Unlike correlates coming from unobserved confounds, causal factors drawing upon theoretical models are likely to stably influence the outcome over time.
  - ➢ (Example) GDP forecasting is based on (causal) macroeconomic factors.

개별국가별 VARX 모형 체계[1]



주: 1) *는 해외변수를 표시

〈그림 4〉 추정모형별 우리나라 GDP 성장률의 표본외 예측치[1]

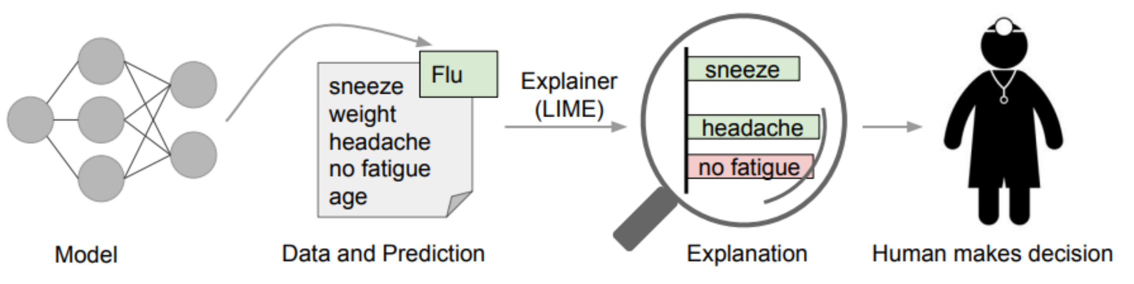


BOK 거시경제모형시스템의 VAR 모형블록 구축 결과, 2013, 한국은행 (http://dl.bok.or.kr/search/DetailView.ax?cid=612430)

(3) Prescription Needs Explanation

- The effectiveness of machine learning-based AI systems will be limited by the machine's inability to explain its thoughts and actions to human users.
  - ➢ (Example) Healthcare diagnose (Ribeiro et al. 2016)

    A "interpretable" model predicts that a patient has the flue, and highlights the symptoms in the patient's history that led to the prediction. Sneeze and headache are portrayed as contributing to the "flu" prediction, while "no fatigue" is evidence against it. With these explanations, a doctor can make an informed decision about whether to trust the model's prediction.



Ribeiro, M.T., Singh, S. and Guestrin, C., 2016. Why Should I Trust You?: Explaining the Predictions of Any Classifier. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* (pp. 1135-1144).

(3) Prescription Needs Explanation

- The effectiveness of machine learning-based AI systems will be limited by the machine's inability to explain its thoughts and actions to human users.
  - ➢ (Example) Mortgage loan

CIO FROM IDG

AI의 '블랙박스'를 어디까지 신뢰할 수 있을까?

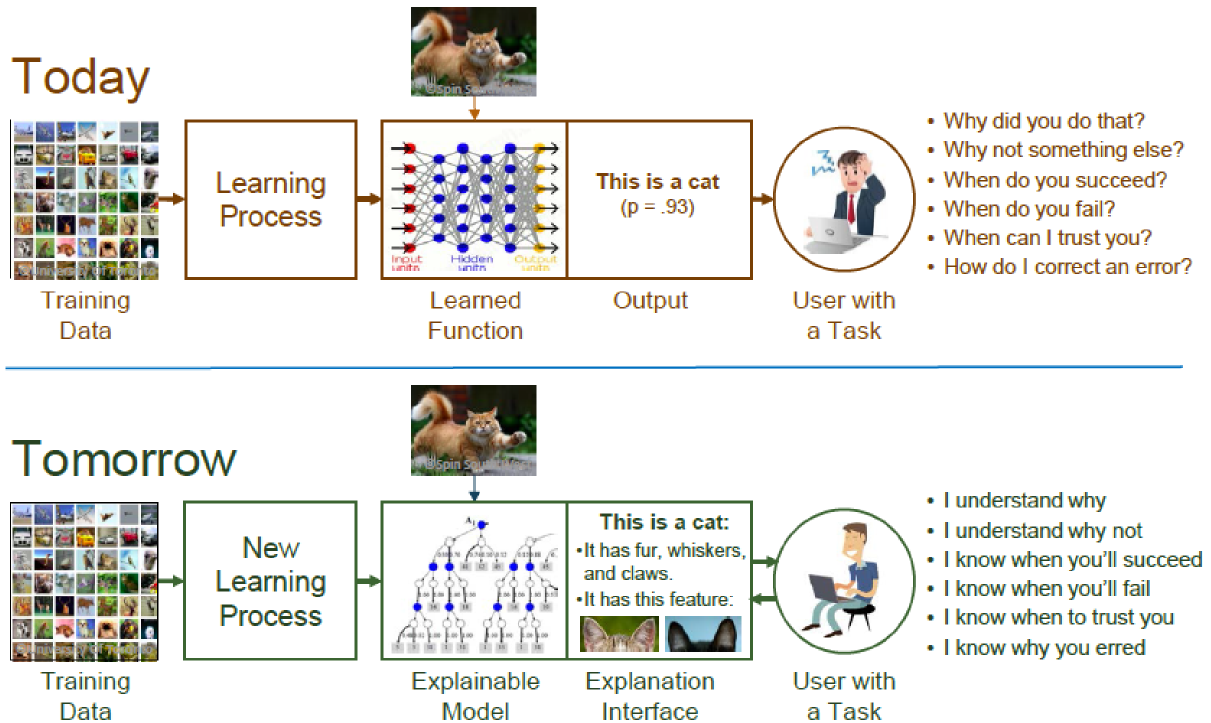
페니매(Fannie Mae)의 운영 및 기술 책임자인 브루스 리에 따르면, 이는 규제 당국에게 AI가 내린 결정을 더 정확히 설명할 수 있어야 하는 페니매 같은 모기지 회사들이 풀어야 할 질문이다. 네스트(Nest) 온도조절기 같은 제품을 사용해 전기 요금을 관리하는 주택 소유주가 모기 상황에 쓸 현금 흐름에 더 여유가 있다고 논리적으로 추론할 수 있을지 모른다. 그러나 AI가 이를 자격 조건으로 포함시키도록 만드는 것은 규제 당국의 시각에서는 문제의 소지가 있다.

리는 "공정한 대출에 대한 규칙에서 표본의 편향에 대한 문제 없이, 네스트를 사용하는 사람들에게 더 나은 이자율의 대출 상품을 제공할 수 있을까? (분명히 AI가 활용될 분야인) 신용 및 대출 결정에 AI를 활용하기 위해 넘어야 할 규제적 장애물이 아주 많다. 우리가 하고 있는 일 중에는 부적절한 편향이 발생하지 않도록 만들고, 주택 시장 인프라에 이익이 발생하도록 만드는 철저한 검증과 관련된 일들이 많다. 특히 AI의 경우 설명이 가능해야 한다"라고 강조했다.

"AI의 '블랙박스'를 어디까지 신뢰할 수 있을까?" (CIO, 2017.7.10)
http://www.ciokorea.com/t/22001/%EB%A8%B8%EC%8B%A0%EB%9F%AC%EB%8B%9D%20%7C%20%EB%94%A5%EB%9F%AC%EB%8B%9D/34812

(3) Prescription Needs Explanation

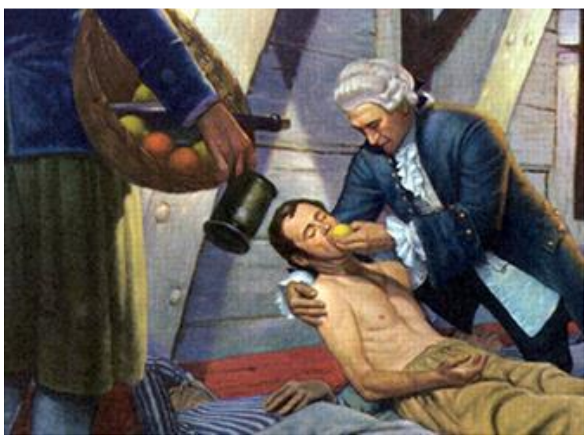
- Explainable artificial intelligence (XAI)
  - ➢ In my opinion, explainable artificial intelligence would be a promising area where computer science and social science (empirical research) meet together.



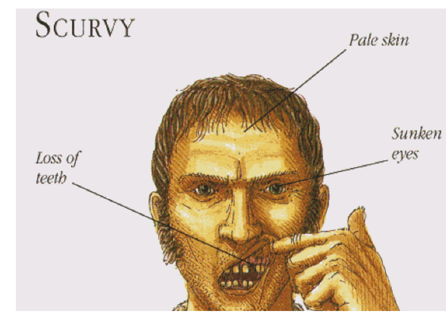
Source: http://explainablesystems.comp.nus.edu.sg/wp-content/uploads/2018/03/XAI%20for%20IUI%202018.pdf

(4) Causation Guides the Better System Design

- Example: Dr. James Lind's solution for scurvy
  - ➢ Dr. James Lind found the effect of lemon on scurvy, but did not understand why.
  - ➢ Lind mistakenly believed that the acidity of the fruit was the cure and tried to create a less-perishable remedy by heating the citrus juice into a concentrate, which destroyed the vitamin C.



Dr. James Lind (1716 – 1794)

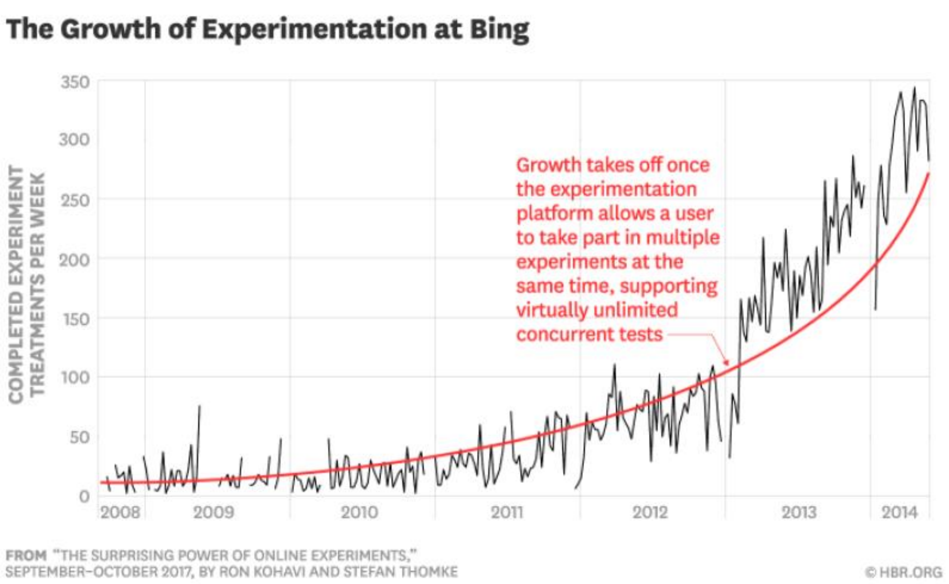


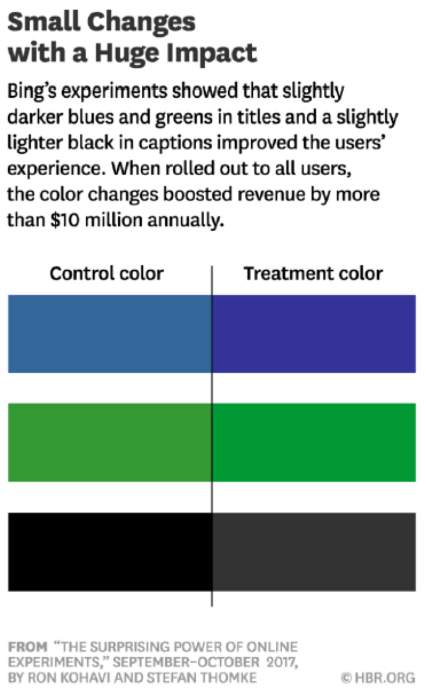
Underlying mechanism

The Surprising Power of Online Experiments, 2017. Harvard Business Review. https://hbr.org/2017/09/the-surprising-power-of-online-experiments

(4) Causation Guides the Better System Design

- Example: Online experiments (A/B tests)
  - ➢ Microsoft has conducted A/B tests on their websites more than hundreds per week in order to improve functions and interfaces, as well as to inform investment decisions.

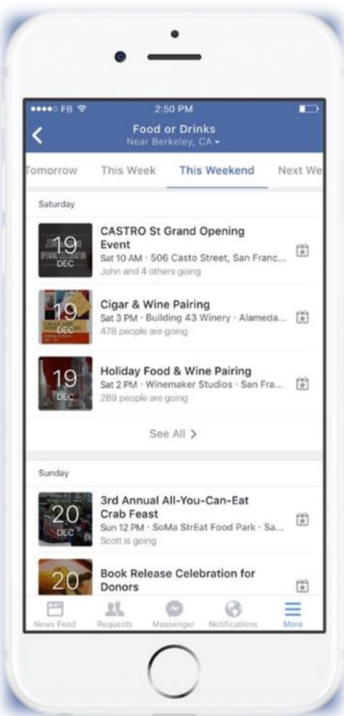


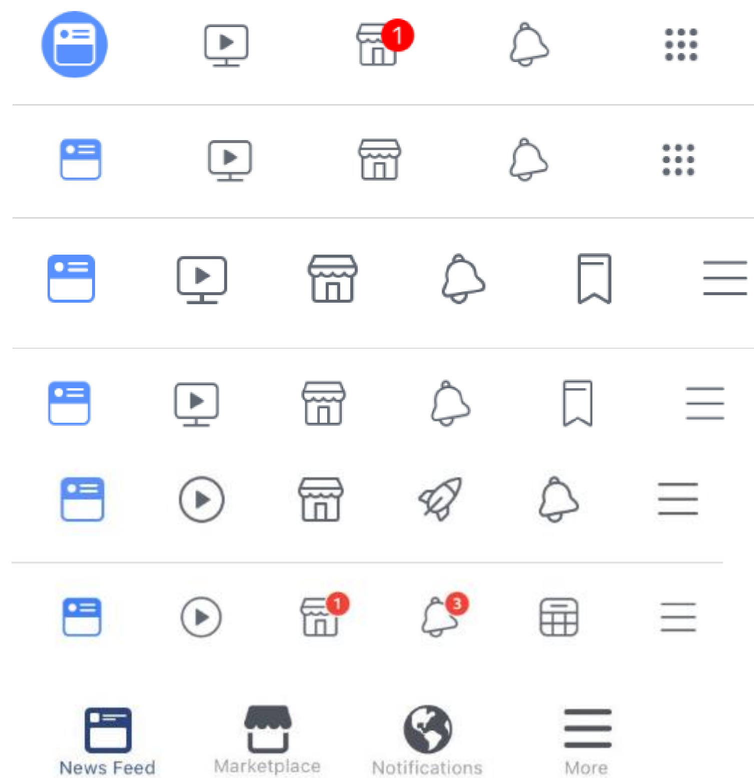


The Surprising Power of Online Experiments, 2017. Harvard Business Review. https://hbr.org/2017/09/the-surprising-power-of-online-experiments

(4) Causation Guides the Better System Design

- Example: Online experiments (A/B tests)
  - ➢ You may see the Facebook App interface different from your friend's.





Source: https://docs.google.com/document/d/19BoeoynEBtk0lQZREyrhHHemaMT9zP3CZL31xr6Xtvs/edit

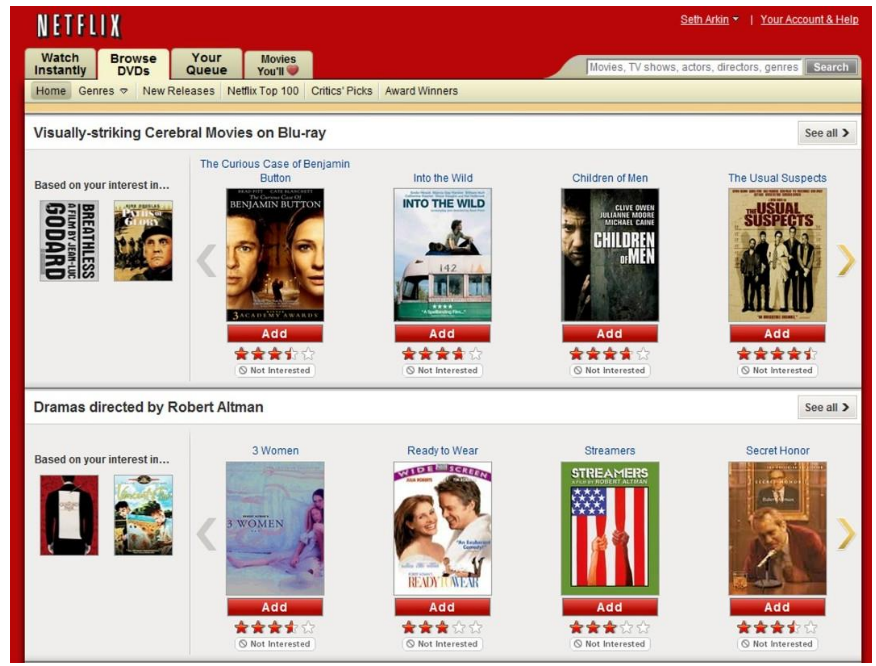
Prescriptive Analytics

= Predictive Analytics + Optimization

Same Predictive Algorithms, Different Goals

amazon

NETFLIX

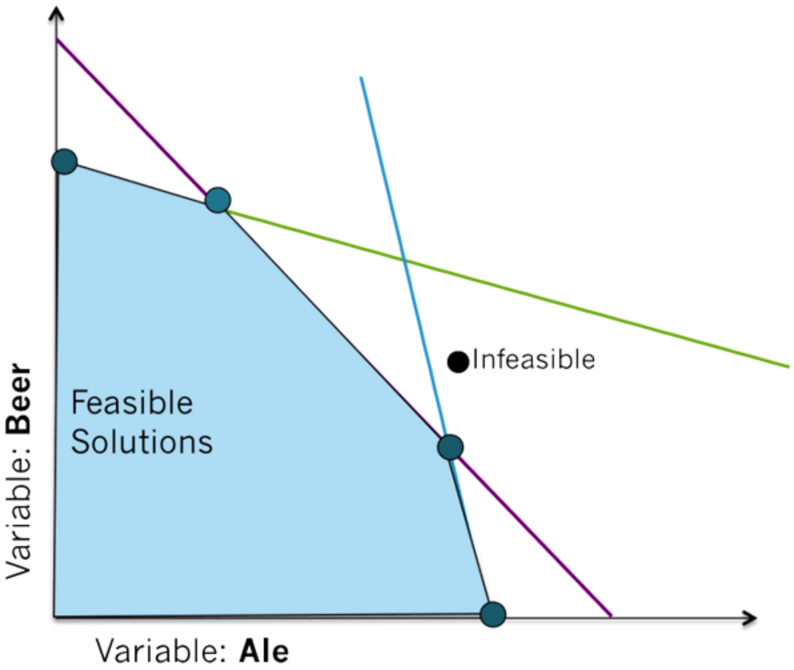


What is Optimization?

- Optimization is a method to achieve the best outcome (such as maximum profit or lowest cost) in a mathematical model whose requirements (constraints) are represented by linear or non-linear relationships.

$$\text{Maximize}: \quad P = 90x_1 + 25x_2$$

$$\begin{aligned} \text{Subject to}: \quad & 8x_1 + 2x_2 \le 400 \\ & 2x_1 + x_2 \le 120 \\ & x_1,\ x_2 \ge 0 \end{aligned}$$



Prescription = Prediction + Optimization

- Example: Online ad bidding (Ren et al. 2018)
  - ➢ The goal of prescriptive analytics

    Spending the campaign budget on the most effective ad opportunities to achieve high profits and more click-through rate (CTR)
  - ➢ Analytics approach

    (1) Predicting the CTR and cost estimation (winning price of the given bid request)

    (2) Based on the estimated CTR and cost, there is a need to seek for the optimal bidding function along with other considerations including the campaign budget and the auction volume etc.

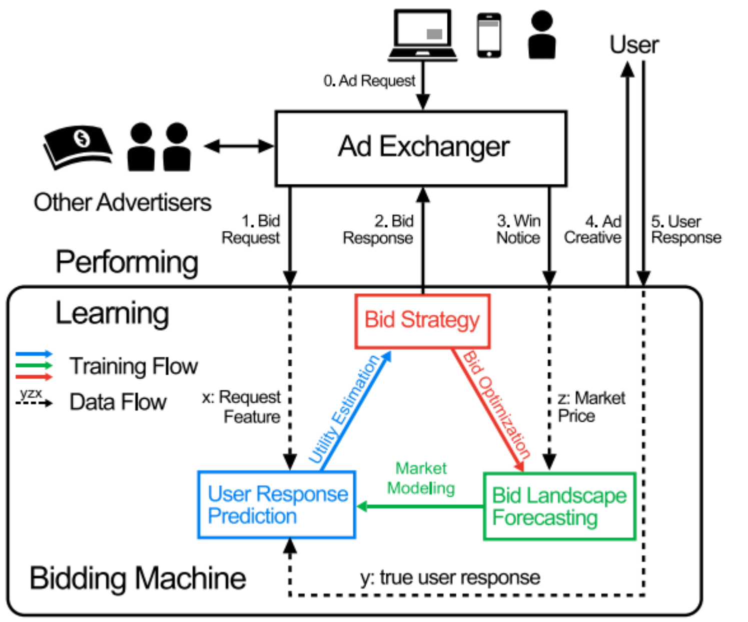


Fig. 1. The joint learning framework of bidding machine.

Ren, K., Zhang, W., Chang, K., Rong, Y., Yu, Y. and Wang, J., 2018. Bidding Machine: Learning to Bid for Directly Optimizing Profits in Display Advertising. *IEEE Transactions on Knowledge and Data Engineering*, 30(4), pp.645-659.

Prescription = Prediction + Optimization

- Example: Medical therapy design (Bertsimas et al. 2016)
  - ➢ The goal of prescriptive analytics

    Improving the quality of chemotherapy regimens for cancer, in the range of acceptable toxicity
  - ➢ Analytics approach

    (1) Predicting clinical trial efficacy and toxicity outcomes

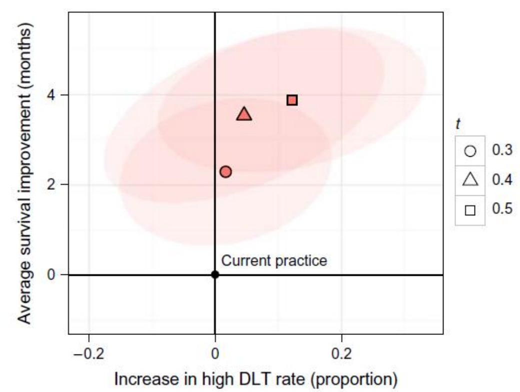
    (2) Integer optimization using our statistical models to select novel chemotherapy regimens with high predicted efficacy and acceptable predicted toxicity

Maximizing the predicted overall survival of the selected chemotherapy regimen

$$\max_{\mathbf{b},\mathbf{i},\mathbf{a}} \ \{(\hat{\boldsymbol{\beta}}^{\mathbf{b}}_{\mathrm{OS}} + \boldsymbol{\Gamma}\mathbf{u})'\mathbf{b} + (\hat{\boldsymbol{\beta}}^{\mathbf{i}}_{\mathrm{OS}})'\mathbf{i} + (\hat{\boldsymbol{\beta}}^{\mathbf{a}}_{\mathrm{OS}})'\mathbf{a} + (\hat{\boldsymbol{\beta}}^{\mathbf{x}}_{\mathrm{OS}})'\mathbf{x}\}$$

$$\begin{aligned}
\text{subject to} \quad & (\hat{\boldsymbol{\beta}}^{\mathbf{b}}_{\mathrm{DLT}})'\mathbf{b} + (\hat{\boldsymbol{\beta}}^{\mathbf{i}}_{\mathrm{DLT}})'\mathbf{i} + (\hat{\boldsymbol{\beta}}^{\mathbf{a}}_{\mathrm{DLT}})'\mathbf{a} + (\hat{\boldsymbol{\beta}}^{\mathbf{x}}_{\mathrm{DLT}})'\mathbf{x} \le t, \\
& \sum_{d=1}^{n} b_d \le N, \\
& \mathbf{A}\mathbf{b} \le \mathbf{c}, \\
& (\mathbf{b}, \mathbf{i}, \mathbf{a}) \notin P, \\
& (b_d, i_d, a_d) \in \boldsymbol{\Omega}_d, \quad d = 1, \ldots, n, \\
& b_d \in \{0, 1\}, \quad d = 1, \ldots, n.
\end{aligned}$$

Predicted toxicity is bounded by a constant *t*.



Bertsimas, D., O'Hair, A., Relyea, S. and Silberholz, J., 2016. An Analytics Approach to Designing Combination Chemotherapy Regimens for Cancer. *Management Science*, 62(5), pp.1511-1531.

Prescription = Prediction + Optimization

- Example: Outpatient appointment scheduling (Samorani and LaGanga 2015)
  - ➢ The goal of prescriptive analytics

    Maximizing the revenue made from seeing patients, based on predicted probability of patient's no-show (related to costs of patient waiting time, in the presence of overbooking)
  - ➢ Analytics approach

    (1) Predicting clinical trial efficacy and toxicity outcomes

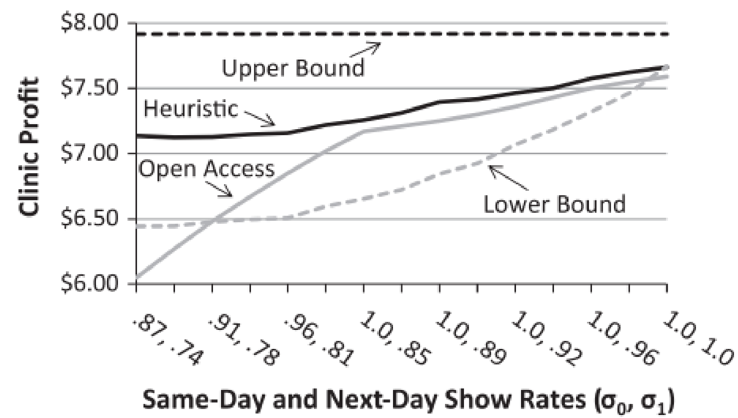
    (2) Integer optimization using our statistical models to select novel chemotherapy regimens with high predicted efficacy and acceptable predicted toxicity

Maximizing the revenue and minimize waiting costs

$$\text{Max} \quad P = \sum_{d=0\ldots h-1} P(i) = \sum_{d=0\ldots h-1} [R_d - O_d - W_d]$$

$$r_d = \tilde{\pi} - d\delta + \pi \cdot s_d, \quad \forall d = 0, \ldots, h-1$$

Predicting whether the appointment will result in a show or no-show

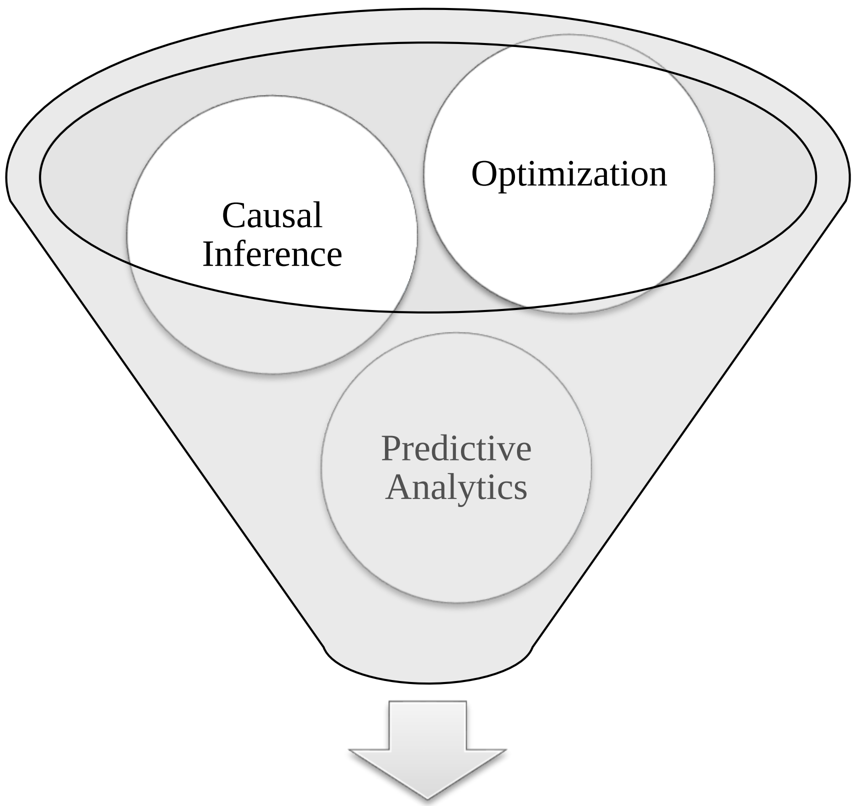


Samorani, M. and LaGanga, L.R., 2015. Outpatient Appointment Scheduling Given Individual Day-Dependent No-Show Predictions. *European Journal of Operational Research*, 240(1), pp.245-257.

KAIST COLLEGE OF BUSINESS

Conclusion

The Art of Prescriptive Analytics



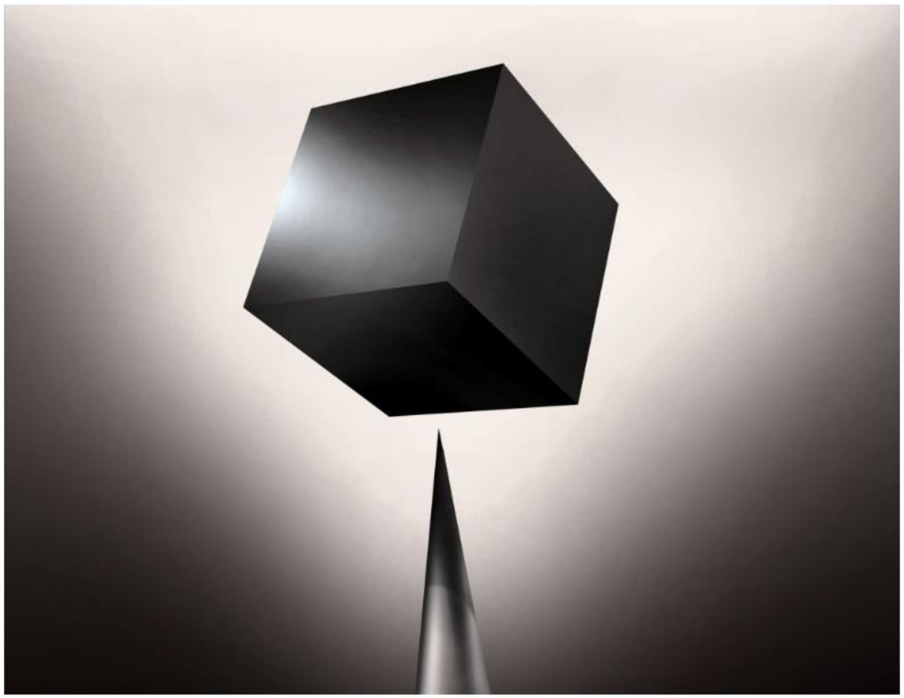
Prescriptive Analytics

We Don’t Want to Just Rely on a Black Box Model

MIT Technology Review

Log in / Create an account   Search   Subscribe

Topics+   The Download   Magazine   Events   More+



KEITH RANKIN

Intelligent Machines

The Dark Secret at the Heart of AI

No one really knows how the most advanced algorithms do what they do. That could be a problem.

by Will Knight   April 11, 2017

https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/

KAIST COLLEGE OF BUSINESS

Causality Should Play a Critical Role in Prescriptive Analytics

- Data-driven prescriptive analytics is not yet established in scholarship and practice. We have still limited understanding about the role of causal inference in prescriptive analytics.
  - ➢ I believe this area will be most promising in business analytics.

informs
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–1
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Call for Papers

Management Science—Special Issue on Data-Driven Prescriptive Analytics

Coeditors: Kay Giesecke,[a] Gui Liberali,[b] Hamid Nazerzadeh,[c] J. George Shanthikumar,[d] Chung Piaw Teo[e]

[a] Stanford University, Stanford, California 94305; [b] Erasmus University, 3062 PA Rotterdam, Netherlands;
[c] University of Southern California, Los Angeles, California 90089; [d] Purdue University, West Lafayette, Indiana 47907;
[e] National University of Singapore, Singapore 119245

Contact: giesecke@stanford.edu (KG); liberali@rsm.nl (GL); nazerzad@marshall.usc.edu (HN); shanthikumar@purdue.edu (JGS); bizteocp@nus.edu.sg (CPT)

Deadline for submission is March 31, 2019.

Alternately, the papers could develop innovative solution based on novel methods that harness advances in areas such as machine and reinforcement learning, optimization, computation, and algorithms; and validate results empirically in actual application settings. Some examples include the following:

- Novel prescriptive methods based on causal inference about the role of individual- and social-media behavior in population-level causata such as voting, product ratings, and opinion aggregation

Causality Should Play a Critical Role in Machine Learning

- Data-driven prescriptive analytics is not yet established in scholarship and practice. We have still limited understanding about the role of causal inference in prescriptive analytics.
  - ➢ I believe this area will be most promising in business analytics.

Call for Papers

5th Workshop on Fairness, Accountability, and Transparency in Machine Learning (FAT/ML 2018)

Co-located with 35th International Conference on Machine Learning (ICML 2018)

15 July 2018, Stockholm, Sweden

This year, the workshop is co-located with ICML, and will consist of invited talks, invited panels, contributed talks, as well as a poster session. We welcome paper submissions from researchers and practitioners that address any issue of fairness, accountability, and transparency related to machine learning. In particular, we will place a special emphasis on causal inference to address questions of fairness, and to create recommendation systems directed at altering causal factors. We will also focus on issues surrounding the collection, measurement, and mitigation of biased data.

https://www.fatml.org/schedule/2018/page/call-for-papers-2018

End of Document