ECONOMIC FINDING • FEBRUARY 2023

Machine Learning as a Tool for Hypothesis Generation

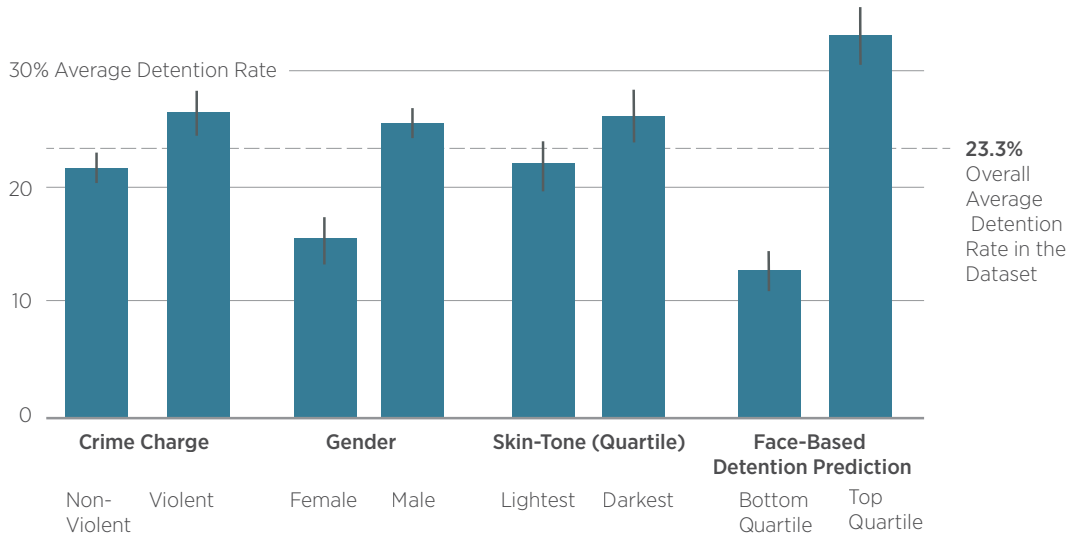
Based on BFI Working Paper 2023-28, "Machine Learning as a Tool for Hypothesis Generation," by Jens Ludwig, UChicago's Harris School of Public Policy; and Sendhil Mullainathan, Chicago Booth

By applying machine-learning algorithms to human behavior data, this paper offers a novel approach to hypothesis generation for scientific research; the authors apply this framework to reveal that judges rely on facial characteristics of defendants when making incarceration decisions.

For all of its empirical and theoretical rigor, science often begins with an intuition or inspiration. Long before a new idea becomes a paper that appears in an academic journal, for example, it begins as a hypothesis. These creative suppositions commence with "data" stored in a researcher's mind, which she then "analyzes" through a purely psychological process of pattern recognition. The "aha" moments that may follow are not necessarily inspiration, but rather the output of the researcher's brain-driven data analysis. In other words, scientific contributions often derive from a researcher's idiosyncratic and very human thought process.

Given the importance of scientific research and its many benefits, this raises a question: Is there a better way to generate hypotheses other than a reliance on personal analytical insight? This novel paper examines this question through the lens of machine-learning algorithms and the exploding

Figure 1 • Relationship Between Detention Rates and Defendant Characteristics



Note: This figure shows the average validation set detention rates for defendants by different defendant characteristics: crime charge is violent vs. non-violent (first panel), defendant is male vs. female (second panel), defendant is in the lightest (Q4) vs. darkest (Q1) skin tone shade according to independent subject ratings of mugshots (third panel), and defendant is in lowest quartile of predicted risk (Q1) vs. highest quartile (Q4) according to mugshot-based predictor of judge detention decision (final panel). 95% confidence intervals are shown at the top of each bar; overall average detention rate in the validation dataset is 23.3%.

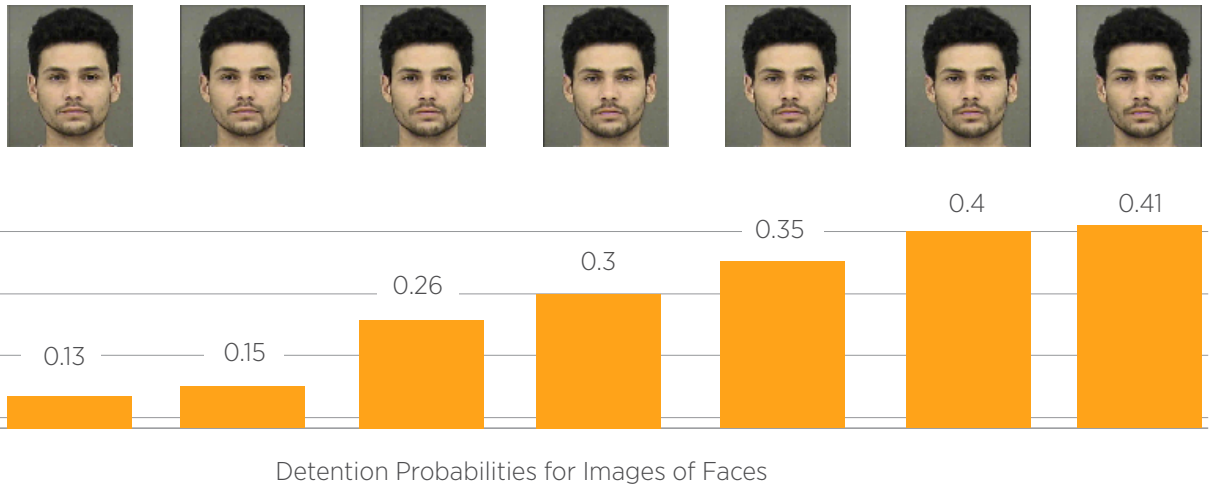
Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

Figure 2 • Illustration of Morphed Faces Along Detention Gradient

0.13 Detention Probability

0.41 Detention Probability

Transformations of the Face Along Selected Steps of the Morphing Process



Note: This figure shows the result of selecting a random point on the generative adversarial network (GAN) latent face space for a white Hispanic male defendant, then using the authors' new morphing procedure to increase the predicted detention risk of the image to 0.41 (at left) or reduce the predicted detention risk down to 0.13 (at right). The overall average detention rate in the validation data set of actual mugshot images is 0.23 by comparison. The row of faces shows the different intermediate images between the two endpoints, while the bar chart underneath shows the predicted detention risk for each of the images.

availability of human behavior data. Second-by-second price and volume data in asset markets, high-frequency cellphone data on location and usage, CCTV camera and police "bodycam" footage, news stories, children's books,[1] the entire text of corporate filings, and so on, is now machine readable. What was once solely mental data in the service of hypothesis generation is increasingly becoming actual data.

The authors posit that these changes can fundamentally change how science gets done, which they demonstrate by developing a procedure that applies machine learning algorithms on rich data sets to generate novel hypotheses. Please see the working paper for more details, but broadly described, the authors extend the human process of generating data-driven correlations to supervised machine learning. Their new approach not only yields far more correlations than a human researcher, but it has the capacity to notice correlations that a human might never discern. This is especially true in high-dimensional data applications, potentially opening the door to research that would otherwise go unexplored.

To illustrate their procedure, which is applicable across disciplines, the authors study a high-stakes issue with profound implications: How pre-trial judges decide which defendants awaiting trial are incarcerated and which are sent free. These decisions are supposedly based on predictions of a defendant's risk, but mounting evidence from recent research suggests that judges make these decisions imperfectly. In this case, when the authors build a deep learning model of the

[1] See "What We Teach About Race and Gender: Representation in Images and Text of Children's Books," by Anjali Adukia, et al., for a BFI Research Brief and links to the paper, an interactive tool, and video presentation.

judge—one that predicts whether the judge will detain a given defendant—a single factor has large explanatory power: the defendant's face. They find the following:

- A predictor that uses only the pixels in the defendant's mugshot explains from one-quarter to nearly one-half of the predictable variation in detention.
- Defendants whose mugshots fall in the bottom quartile of predicted detention are 20.4 percentage points (pp) more likely to be jailed than those in the top quartile.
- By comparison, the difference in detention rates between those arrested for violent versus non-violent crimes is 4.8 pp.

It is important to note what this work does not reveal. The authors do not claim that a mugshot predicts a defendant's behavior; rather their analysis reveals that mugshots predict a judge's behavior: A defendant's appearance correlates strongly with a judge's decision to jail or not. And what are those appearance characteristics? There are two, and the authors label the first as "well-groomed" (e.g., tidy, clean, groomed, vs. unkempt, disheveled, sloppy look), with the second as "heavy-faced" (e.g., wide facial shape, puffier face, rounder face, heavier). Importantly, these features are not just predictive of what the algorithm sees, but also of what judges actually do:

- Both well-groomed and heavy-faced defendants are more likely released.
- Detention rates of defendants in the top and bottom quartile of well-groomedness differ by 5.5 pp (24% of the base rate) while the top vs. bottom quartile difference in heavy-facedness is 7 pp (about 30% of the base rate).
- Both differences are larger than the 4.8 pp detention rate difference between those arrested for violent vs. non-violent crimes.

Again, please see the working paper for the authors' careful consideration of the many factors involved in their analysis, including their discussion of the ways that psychological and economic research over the past century has informed our understanding of people's reaction to faces, including on such factors as race. The point here is that the authors' algorithm seems to have found something new, beyond what scientists have previously hypothesized, and beyond human capability.

Bottom line: Hypothesis generation matters, and developing new ideas need not remain an idiosyncratic or nebulous process. The authors' framework reveals that combining supervised machine learning with rich human behavior data can open the scientific world to new and fruitful lines of research. The authors' example of judicial decision-making, along with other applications that they describe, are just the beginning. Future research will help move hypothesis generation from a pre-scientific to a scientific activity.

READ THE WORKING PAPER

NO. 2023-28 · FEBRUARY 2023

Machine Learning as a Tool for Hypothesis Generation

bfi.uchicago.edu/working-paper/2023-28

ABOUT OUR SCHOLARS

Jens Ludwig
Edwin A. and Betty L. Bergman Distinguished Service Professor, Pritzker Director of UChicago's Crime Lab, Codirector of the Education Lab, Harris School of Public Policy
harris.uchicago.edu/directory/jens-ludwig

THE UNIVERSITY OF CHICAGO
HARRIS SCHOOL OF PUBLIC POLICY

Sendhil Mullainathan
Roman Family University Professor of Computation and Behavioral Science, Chicago Booth
chicagobooth.edu/faculty/directory/m/sendhil-mullainathan

CHICAGO BOOTH
The University of Chicago Booth School of Business