BIG DATA

A tool for inclusion or exclusion?

AN FTC WORKSHOP
SEPT 15, 2014

Welcome

Opening Remarks

Chairwoman Edith Ramirez
Federal Trade Commission

Presentation: Framing the Conversation

Solon Barocas
Princeton University
Center for Information Technology Policy

Big Data: A Tool for Inclusion or Exclusion? Framing the Conversation

Solon Barocas

Princeton University

CENTER FOR INFORMATION TECHNOLOGY POLICY

CITP.PRINCETON.EDU

Big Data

Volume, Velocity, Variety

Observational → Transactional data
Self-Reported and User-Generated → Social media
Experimental → A/B testing

Data Mining

(Machine Learning)

Data Mining

- Automate the process of discovering useful patterns—regularities upon which subsequent decision-making can rely
  - “Learning”
- The accumulated set of discovered relationships in the dataset is commonly called a “model,” and these models can be employed to automate the process of:
  - classifying entities or activities of interest
  - estimating the value of unobserved variables or
  - predicting future outcomes

Data Mining as Discrimination

- By definition, data mining is *always* a form of statistical (and therefore seemingly rational) discrimination
- The very point of data mining is:
  - to provide a rational basis upon which to distinguish between individuals
  - to reliably confer to the individual the qualities possessed by those who seem statistically similar

How Data Mining Can Discriminate

1. Defining the target variable
2. Collecting and labeling the training data
3. Feature selection
4. Proxies
5. Masking

1. Target Variables

- Determine how to solve the problem at hand by translating it into a question about the value of some “target variable”
- Treats the target variable as a function of some other set of observed characteristics
  - A, B, C → X
  - D, E, F → Y

The "Art" of Data Mining

- The proper specification of the target variable is frequently not obvious, and it is the data miner's task to define it
- The definition of the target variable and its associated class labels will determine what data mining happens to find
- And it is possible to parse the problem and define the target variable in such a way that protected classes happen to be subject to systematically less favorable determinations

2. Training Data

- Data mining is really a way to learn by example
- The data that function as examples are known as "training data"—quite literally the data that train the model to behave in a certain way
- Two types of problems can arise
  - Data collection: skewed set of examples
  - Labeling of examples: setting a bad example

Data Collection

- Data mining is especially sensitive to statistical bias because it aims to extract *general* rules from a *particular* set of examples
  - These rules will only hold for future cases if the cases in the training data resemble those to which they are applied
- But data gathered for routine business purposes tend to lack the rigor of social scientific data collection
  - Firms tend to perform such analyses in order to *change* the composition of their customer base

Data Collection: Uncounted, Unaccounted, Discounted

- The quality and representativeness of records might vary in ways that correlate with class membership
  - less involved in the formal economy and its data-generating activities
  - unequal access to and less fluency in the technology necessary to engage online
  - less profitable customers or less important constituents and therefore less interesting as targets of observation
- The under- and over-representation of members of protected classes is not always evident

Data Collection: Limiting Future Contact

- Worse, skewed results may lead to decision procedures that limit the future contact companies have with specific groups, skewing still further the sample upon which subsequent analyses will be performed
- They would deny members of these populations the opportunity to prove that they buck the apparent trend

Labeling Examples

- Sometimes a rather straightforward affair
  - e.g., spam/not spam
- Sometimes a laborious process that is fraught with peril
  - e.g., good/bad customer

Labeling Examples: Reproduce Past Prejudice

- So long as prior decisions affected by some form of prejudice serve as examples of *correctly* rendered determinations, data mining will necessarily infer rules that exhibit the same prejudice
  - e.g., Fair Isaac's comments in the early debates about credit scoring
- Data mining can turn the conscious prejudice or implicit bias of individuals involved in previous decision-making into a formalized rule that would systematically discount all applicants in this way
  - For all their potential problems, the labels applied to the training data must serve as ground truth.

Labeling Examples: Reflect Current Prejudice

- Not only can data mining inherit *prior* prejudice through the mislabeling of examples, it can also reflect *current* prejudice through the ongoing behavior of users taken as inputs to data mining
  - When relying on data mining to cater to the demonstrated preferences of users, companies may unintentionally adopt the prejudices that guide users' behavior

3. Feature Selection

- The process of settling on the specific string of input variables
- Protected classes may find that they suffer a disproportionate cost of errors
  - Coarseness and comprehensiveness of the criteria that permit statistical discrimination and the uneven rates at which different groups happen to be subject to erroneous determinations
  - Artifacts of statistical reasoning rather than decision-maker prejudice or bias in the dataset

At What Cost?

- Obtaining information that is sufficiently rich to draw precise distinctions can be expensive. Even marginal improvements in accuracy may come at significant practical costs, and may justify a less exacting and encompassing analysis.
  - Does the relatively higher costs involved in gaining more data about marginalized groups justify subjecting them to higher error rates?
  - Should these groups bear the disproportionate burden of erroneous determinations, even if this means that the majority enjoys greater accuracy in decision-making?

4. Proxies

- The very same criteria that correctly sort individuals according to their predicted profitability, for example, may also sort individuals according to class membership
  - Decision-makers' reasonable priorities as profit-seekers unintentionally recapitulate the inequality that happens to exist in society
- These discoveries reveal the simple fact of inequality, but they also reveal the fact that these are inequalities in which members of protected classes are frequently the groups in the relatively less favorable position
  - Better data will simply expose the exact extent of inequality

“Redundant Encodings”

- In many instances, making accurate determinations will mean considering factors that are somehow correlated with proscribed features
- There is no obvious way to determine how correlated a relevant attribute must be with class membership to be worrisome, nor is there a self-evident way to determine when an attribute is sufficiently relevant to justify its consideration, despite the fact that it is highly correlated with class membership.

5. Masking

- Data mining could also breathe the new life into traditional forms of intentional discrimination because decision-makers with prejudicial views can mask their intentions by exploiting each of the mechanisms enumerated above
  - Also possible to infer class membership
- Unintentional discrimination is likely to be *far* more common than the kinds of discrimination that could be pursued intentionally

Conclusion

Unintentionality

Exacerbate Inequality

No Ready Answer

Solon Barocas

sbarocas@princeton.edu

solon.barocas.org

@s010n

Solon Barocas and Andrew Selbst, "Big Data's Disparate Impact"

http://ssrn.com/abstract=2477899

Panel 1: Assessing the Current Environment

- **Kristin Amerling**, U.S. Senate Committee on Commerce, Science and Transportation
- **danah boyd**, Microsoft Research & New York University
- **Mallory Duncan**, National Retail Federation
- **Gene Gsell** , SAS
- **David Robinson**, Robinson + Yu
- **Joseph Turow**, University of Pennsylvania

Break

Panel 2: What's On the Horizon With Big Data?

- **Alessandro Acquisti**, Carnegie Mellon University
- **Pamela Dixon**, World Privacy Forum
- **Cynthia Dwork**, Microsoft Research
- **Mark MacCarthy**, Software Information Industry Association
- **Stuart Pratt**, Consumer Data Industry Association
- **Nicol Turner-Lee**, Minority Media and Telecommunications Council

Lunch

Remarks

Commissioner Julie Brill
Federal Trade Commission

Presentation: Digging Into the Data

LaTanya Sweeney &
Jinyan Zang
Federal Trade Commission

Panel 3: Surveying the Legal Landscape

- **Leonard Chanin**, Morrison Foerster
- **Carol Miaskoff**, Equal Employment Opportunity Commission
- **Montserrat Miller**, Arnall Golden Gregory
- **C. Lee Peeler**, Council of Better Business Bureaus
- **Peter Swire**, Georgia Institute of Technology

Break

Panel 4: Considerations on the Path Forward

- **Christopher Calabrese**, American Civil Liberties Union
- **Daniel Castro**, Information Technology and Innovation Foundation
- **Jeanette Fitzgerald**, Epsilon
- **Jeremy Gillula**, Electronic Frontier Foundation
- **Michael Spadea**, Promontory Financial Group
- **Christopher Wolf**, Future of Privacy Forum

Closing Remarks

Jessica Rich
Bureau of Consumer Protection
Federal Trade Commission