

BROOKINGS

Report

Credit denial in the age of AI

Aaron Klein Thursday, April 11, 2019

Editor's Note:

This report is part of "[A Blueprint for the Future of AI](#)," a series from the Brookings Institution that analyzes the new challenges and potential policy solutions introduced by artificial intelligence and other emerging technologies.

Banks have been in the business of deciding who is eligible for credit for centuries. But in the age of artificial intelligence (AI), machine learning (ML), and big data, digital technologies have the potential to transform credit allocation in positive as well as negative directions. Given the mix of possible societal ramifications, policymakers must consider what practices are and are not permissible and what legal and regulatory structures are necessary to protect consumers against unfair or discriminatory lending practices.

In this paper, I review the history of credit and the risks of discriminatory practices. I discuss how AI alters the dynamics of credit denials and what policymakers and banking officials can do to safeguard consumer lending. AI has the potential to alter credit practices in transformative ways and it is important to ensure that this happens in a safe and prudent manner.

The history of financial credit

There are many reasons why credit is treated differently than the sale of goods and services. Since there is a history of credit being used as a tool for discrimination and segregation, regulators pay close attention to bank lending practices. Indeed, the term “redlining” originates from maps made by government mortgage providers to use the provision of mortgages to segregate neighborhoods based on race. In the era before

computers and standardized underwriting, bank loans and other credit decisions were often made on the basis of personal relationships and sometimes discriminated against racial and ethnic minorities.

People pay attention to credit practices because loans are a uniquely powerful tool to overcome discrimination and the historical effects of discrimination on wealth accumulation. Credit can provide new opportunities to start businesses, increase human and physical capital, and build wealth. Special efforts must be made to ensure that credit is not allocated in a discriminatory fashion. That is why different parts of our credit system are legally required to invest in communities they serve.

The Equal Credit Opportunity Act of 1974 (ECOA) represents one of the major laws employed to ensure access to credit and guard against discrimination. ECOA lists a series of protected classes that cannot be used in deciding whether to provide credit and at what interest rate it is provided. These include the usual—race, sex, national origin, age—as well as less common factors, like whether the individual receives public assistance.

The standards used to enforce the rules are disparate treatment and disparate impact. Disparate treatment is relatively straight forward: Are people within a protected class being clearly treated differently than those of nonprotected classes, even after accounting for credit risk factors? Disparate impact is broader, asking whether the impact of a policy treats people disparately along the lines of protected class. The Consumer Financial Protection Bureau defines disparate impact as occurring when:

“A creditor employs facially neutral policies or practices that have an adverse effect or impact on a member of a protected class unless it meets a legitimate business need that cannot reasonably be achieved by means that are less disparate in their impact.”

The second half of the definition provides lenders the ability to use metrics that may have correlations with protected class elements so long as it meets a legitimate business need, **and** there are no other ways to meet that interest that have less disparate impact.

In a world free of bias, credit allocation would be based on borrower risk, known simply as “risk-based pricing.” Lenders simply determine the true risk of a borrower and charge the borrower accordingly. In the real world, however, factors used to determine risk are almost always correlated on a societal level with one or more protected class. Determining who is likely to repay a loan is clearly a legitimate business impact. Hence, financial institutions can and do use factors such as income, debt, and credit history, in determining whether and at what rate to provide credit, even when those factors are highly correlated with protected classes like race and gender. The question becomes not only where to draw the line on what can be used, but more importantly, how is that line drawn so that it is clear what new types of data and information are and are not permissible.

AI and credit allocation

How will AI challenge this equation in regard to credit allocation? When artificial intelligence is able to use a machine learning algorithm to incorporate big datasets, it can find empirical relationships between new factors and consumer behavior. Thus, AI coupled with ML and big data, allows for far larger types of data to be factored into a credit calculation. Examples range from social media profiles, to what type of computer you are using, to what you wear, and where you buy your clothes. If there are data out there on you, there is probably a way to integrate it into a credit model. But just because there is a statistical relationship does not mean that it is predictive, or even that it is legally allowable to be incorporated into a credit decision.

“If there are data out there on you, there is probably a way to integrate it into a credit model.”

Many of these factors show up as statistically significant in whether you are likely to pay back a loan or not. A recent paper by Manju Puri et al., demonstrated that five simple digital footprint variables could outperform the traditional credit score model in predicting who would pay back a loan. Specifically, they were examining people shopping online at Wayfair (a company similar to Amazon but much larger in Europe) and applying for credit to complete an online purchase. The five digital footprint variables are simple, available immediately, and at no cost to the lender, as opposed to say, pulling your credit score, which was the traditional method used to determine who got a loan and at what rate:

- Borrower type of computer (Mac or PC).
- Type of device (phone, tablet, PC).
- Time of day you applied for credit (borrowing at 3am is not a good sign).
- Your email domain (Gmail is a better risk than Hotmail).
- Is your name part of your email (names are a good sign).

An AI algorithm could easily replicate these findings and ML could probably add to it. Each of the variables Puri found is correlated with one or more protected classes. It would probably be illegal for a bank to consider using any of these in the U.S, or if not clearly illegal, then certainly in a gray area.

Incorporating new data raises a bunch of ethical questions. Should a bank be able to lend at a lower interest rate to a Mac user, if, in general, Mac users are better credit risks than PC users, even controlling for other factors like income, age, etc.? Does your decision change if you know that Mac users are disproportionately white? Is there anything inherently racial about using a Mac? If the same data showed differences among beauty products targeted specifically to African American women would your opinion change?

“Should a bank be able to lend at a lower interest rate to a Mac user, if, in general, Mac users are better credit risks than PC users, even controlling for other factors like income or age?”

Answering these questions requires human judgment as well as legal expertise on what constitutes acceptable disparate impact. A machine devoid of the history of race or of the agreed upon exceptions would never be able to independently recreate the current system that allows credit scores—which are correlated with race—to be permitted, while Mac vs. PC to be denied.

With AI, the problem is not only limited to overt discrimination. Federal Reserve Governor Lael Brainard pointed out an actual example of a hiring firm’s AI algorithm: “the AI developed a bias against female applicants, going so far as to exclude resumes of graduates from two women’s colleges.” One can imagine a lender being aghast at finding out that their AI was making credit decisions on a similar basis, simply rejecting everyone from a woman’s college or a historically black college or university. But how does the lender even realize this discrimination is occurring on the basis of variables omitted?

A recent paper by Daniel Schwarcz and Anya Prince argues that AIs are inherently structured in a manner that makes “proxy discrimination” a likely possibility. They define proxy discrimination as occurring when “the predictive power of a facially-neutral characteristic is at least partially attributable to its correlation with a suspect classifier.” This argument is that when AI uncovers a statistical correlation between a certain behavior of an individual and their likelihood to repay a loan, that correlation is actually being driven by two distinct phenomena: the actual informative change signaled by this behavior and an underlying correlation that exists in a protected class. They argue that traditional statistical techniques attempting to split this impact and control for class may not work as well in the new big data context.

Policymakers need to rethink our existing anti-discriminatory framework to incorporate the new challenges of AI, ML, and big data. A critical element is transparency for borrowers and lenders to understand how AI operates. In fact, the existing system has a safeguard already in place that itself is going to be tested by this technology: the right to know why you are denied credit.

Credit denial in the age of artificial intelligence

When you are denied credit, federal law requires a lender to tell you why. This is a reasonable policy on several fronts. First, it provides the consumer necessary information to try and improve their chances to receive credit in the future. Second, it creates a record of decision to help ensure against illegal discrimination. If a lender systematically denied people of a certain race or gender based on false pretext, forcing them to provide that pretext allows regulators, consumers, and consumer advocates the information necessary to pursue legal action to stop discrimination.

This legal requirement creates two serious problems for financial AI applications. First, the AI has to be able to provide an explanation. Some machine learning algorithms can arrive at decisions without leaving a trail as to why. Simply programming a binary yes/no credit decision is insufficient. In order for the algorithm to be compliant, it must be able to identify the precise reason or reasons why a credit decision was made. This is an added level of complexity for AI that might delay adoption.

The second problem is what happens when the rationale for the decision is unusual. For example, one of the largest drivers of personal bankruptcy and default is divorce. An AI algorithm may be able to go through a person's bank records and web search history, and determine with some reasonable accuracy if they are being unfaithful. Given that is a leading cause of divorce, it would probably be a relevant factor in a risk-based pricing regime, and a good reason to deny credit.

Is it acceptable for a bank to deny an application for credit because a machine suspects infidelity? If so, the next step would be whether it is right for the bank to inform the consumer directly that is the reason why. Imagine if the bank sent a letter to the

consumer's home with that finding. The use of AI to determine credit coupled with the requirement that written notice be given with the rationale for credit denial raises a host of privacy concerns.

“The use of AI to determine credit coupled with the requirement that written notice be given with the rationale for credit denial raises a host of privacy concerns.”

If it is not acceptable, then who determines what acceptable grounds are? While marital status is a protected class under the ECOA (you cannot discriminate against someone for being single), it is not clear that lenders concerned with changes to marital status would be prohibited from using that information. As AI determines new metrics that interact with existing protected classes, it will be incumbent upon financial regulators, courts, and eventually lawmakers to set new policy rules that govern this brave new world.

Where do we go from here?

AI has the power to transform consumer lending. On the positive side, it could help identify millions of good credit risks who are currently being denied access to credit. My Brookings colleague Makada Henry-Nickie, highlights innovative ways AI can help promote consumer protection. On the negative side, AI could usher in a wave of hidden discrimination, whereby algorithms deny credit or increase interest rates using a host of variables that are fundamentally driven by historical discriminatory factors that remain embedded in society. Traditional safeguards like requiring disclosures of the rationale for credit denial themselves will be challenged by new AI thinking that will raise privacy concerns.

The core principle of risk-based lending will be challenged as machines uncover new features of risk. Some of these are predictive in ways that are hard to imagine. Some are predictive in ways that are difficult to disclose. And some are repetitions resulting from a

world in which bias and discrimination remain. Unlocking the benefits from this data revolution could help us escape the cycle using credit as the powerful tool for opportunity that it is. However, the existing 1970s era legal framework is going to need a reboot to allow us to get there.

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.

Report Produced by **Center for Technology Innovation**