

Did artificial intelligence deny you credit?

March 14 2017, by Anupam Datta



Credit: AI-generated image ([disclaimer](#))

People who apply for a loan from a bank or credit card company, and are turned down, are owed an explanation of why that happened. It's a good idea – because it can help teach people how to repair their damaged credit – and it's a federal law, the [Equal Credit Opportunity Act](#). Getting an answer wasn't much of a problem in years past, when humans made those decisions. But today, as artificial intelligence systems increasingly assist or replace people making credit decisions,

getting those explanations has become much more difficult.

Traditionally, a loan officer who rejected an application could tell a would-be borrower there was a problem with their income level, or employment history, or [whatever the issue was](#). But computerized systems that use complex machine learning models are difficult to explain, even for experts.

Consumer credit decisions are just one way this problem arises. [Similar concerns](#) exist in health care, online marketing and even criminal justice. My own interest in this area began when a research group I was part of discovered [gender bias in how online ads were targeted](#), but could not explain why it happened.

All those industries, and many others, who use machine learning to analyze processes and make decisions have a little over a year to get a lot better at explaining how their systems work. In May 2018, the new [European Union General Data Protection Regulation](#) takes effect, including a section giving people a right to get an explanation for automated decisions that affect their lives. What shape should these explanations take, and can we actually provide them?

Identifying key reasons

One way to describe why an automated [decision](#) came out the way it did is to identify the factors that were most influential in the decision. How much of a credit denial decision was because the applicant didn't make enough money, or because he had failed to repay loans in the past?

My research group at Carnegie Mellon University, including PhD student Shayak Sen and then-postdoc Yair Zick created a way to [measure the relative influence](#) of each factor. We call it the Quantitative Input Influence.

In addition to giving better understanding of an individual decision, the measurement can also shed light on a group of decisions: Did an algorithm deny credit primarily because of financial concerns, such as how much an applicant already owes on other debts? Or was the applicant's ZIP code more important – suggesting more basic demographics such as race might have come into play?

Capturing causation

When a system makes decisions based on multiple factors it is important to identify which factors cause the decisions and their relative contribution.

For example, imagine a credit-decision system that takes just two inputs, an applicant's debt-to-income ratio and her race, and has been shown to approve loans only for Caucasians. Knowing how much each factor contributed to the decision can help us understand whether it's a legitimate system or whether it's discriminating.

An explanation could just look at the inputs and the outcome and observe correlation – non-Caucasians didn't get loans. But this explanation is too simplistic. Suppose the non-Caucasians who were denied loans also had much lower incomes than the Caucasians whose applications were successful. Then this explanation cannot tell us whether the applicants' race or debt-to-income ratio caused the denials.

Our method can provide this information. Telling the difference means we can tease out whether the system is unjustly discriminating or looking at legitimate criteria, like applicants' finances.

To measure the influence of race in a specific credit decision, we redo the application process, keeping the debt-to-income ratio the same but changing the race of the applicant. If changing the race does affect the

outcome, we know race is a deciding factor. If not, we can conclude the algorithm is looking only at the financial information.

In addition to identifying factors that are causes, we can measure their relative causal influence on a decision. We do that by randomly varying the factor (e.g., race) and measuring how likely it is for the outcome to change. The higher the likelihood, the greater the influence of the factor.

Aggregating influence

Our method can also incorporate multiple factors that work together. Consider a decision system that grants credit to applicants who meet two of three criteria: credit score above 600, ownership of a car, and whether the applicant has fully repaid a home loan. Say an applicant, Alice, with a [credit score](#) of 730 and no car or home loan, is denied credit. She wonders whether her car ownership status or home loan repayment history is the principal reason.

An analogy can help explain how we analyze this situation. Consider a court where decisions are made by the majority vote of a panel of three judges, where one is a conservative, one a liberal and the third a swing vote, someone who might side with either of her colleagues. In a 2-1 conservative decision, the swing judge had a greater influence on the outcome than the liberal judge.

The factors in our credit example are like the three judges. The first judge commonly votes in favor of the loan, because many applicants have a high enough [credit](#) score. The second judge almost always votes against the loan because very few applicants have ever paid off a home. So the decision comes down to the swing judge, who in Alice's case rejects the loan because she doesn't own a car.

We can do this reasoning precisely by using [cooperative game theory](#), a

system of analyzing more specifically how different factors contribute to a single outcome. In particular, we combine our measurements of relative causal influence with the [Shapley value](#), which is a way to calculate how to attribute influence to multiple factors. Together, these form our Quantitative Input Influence measurement.

So far we have evaluated our methods on decision systems that we created by training common machine learning algorithms with real world data sets. Evaluating algorithms at work in the real world is a topic for future work.

An open challenge

Our method of analysis and explanation of how algorithms make decisions is most useful in settings where the factors are readily understood by humans – such as debt-to-income ratio and other financial criteria.

However, explaining the decision-making process of more complex algorithms remains a significant challenge. Take, for example, an image recognition system, like ones that [detect and track tumors](#). It is not very useful to explain a particular image's evaluation based on individual pixels. Ideally, we would like an explanation that provides additional insight into the decision – such as identifying specific tumor characteristics in the image. Indeed, designing explanations for such automated decision-making tasks is keeping many researchers [busy](#).

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: Did artificial intelligence deny you credit? (2017, March 14) retrieved 25 April 2024 from <https://phys.org/news/2017-03-artificial-intelligence-deny-credit.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.