Skadden

Why Does the Brand of My Phone Affect My Credit Rating?

To capitalize on the promise of artificial intelligence and alternative data, boards need to anticipate and mitigate various risks.

– Hidden biases need to be prevented.

– Neither regulators nor the public will be satisfied with "black box" decisions.

– Reputational risk must be weighed alongside legal requirements.

The Potential

Alternative data and artificial intelligence (AI) have generated tremendous excitement in the business world. The technology offers the potential for faster, more efficient and more reliable decisions. Banks and fintech platforms already use them to make credit decisions, and they show promise in other areas, from fraud prevention to hiring decisions.

But the surprising predictive success of AI-based decision models is precisely what makes them tricky legally. Often, the developers of such models cannot explain the relation between a variable and its predictive value. Anything from where you shop to the type of mobile phone you use or the first letter of your last name

may prove to be a predictor of, say, the likelihood you will default on a loan or that you will perform well in a job you are applying for.

To the companies that use such models, this may seem like brilliant data mining — unearthing nonobvious predictors that outperform conventional ones. But to regulators and those harmed by AI-based decisions whose rationales cannot be fully explained, the process may seem capricious.

AI models are the subject of particularly lively debate in the lending sphere, both because lending is a heavily regulated activity and because the technology may turn out to be more reliable than traditional credit bureau factors (number of tradelines, average balance, debt-to-income, etc.). The latter have proven

less effective in predicting defaults over the past year, particularly during the pandemic.

Alternative data may also benefit consumers who have not established the kind of borrowing track record typically relied on by credit bureaus. It could therefore expand the population qualifying for credit. Similar benefits and problems arise in recruiting and other areas where AI models based on alternative data are being explored.

Hidden Biases

The appeal of AI is that the technology can make predictions using offbeat data that humans cannot make or explain. But the fact that AI uncovers new and surprising predictors by poring through hundreds of types of data poses a basic problem: The most valuable variables may have no obvious relation to the thing being predicted, such as a borrower's ability to pay its debts.

In the worst cases, outcomes of these models may be both surprising and problematic. For example, one AI-based recruiting model for software developers relied on the success of previous hires. They were almost entirely male, however, and the model turned out to have a strong bias against women — so strong that

it excluded any candidate from two women's colleges. That could violate employment nondiscrimination laws in jurisdictions where it's not necessary to show discriminatory intent.

This sort of unforeseen bias is on the minds of bank regulators, because it could violate fair lending rules. This hiring example was cited by Federal Reserve Bank Governor Lael Brainard in a recent speech about AI in financial services. Regulators may demand proof that a similar nonobvious variable is not a proxy for another, forbidden factor such as the race or gender of the applicant.

The hiring example also underscores that companies cannot blindly accept AI-based recommendations as technical wizardry. The predictions based on novel data may be quite explainable if you dig deep enough.

Predictors That Aren't Understood

An AI model's "black box" quality itself poses a problem, apart from any biases. To illustrate this, assume one variable in a lender's underwriting model is whether the applicant uses an Apple or a Samsung mobile phone, because that (hypothetically) has been shown to be highly predictive of an applicant's risk of default.

If predictive value were the only factor, the brand variable might satisfy "safety and soundness" bank rules. But regulations often require more than predictive value. U.S. banking regulators require that financial models be "conceptually sound." And banking rules in the U.S., EU and Hong Kong all generally require lenders to be able to explain to an applicant why credit was denied.

Hence, "explainability" — the ability to articulate the relationship between a variable and the attribute being predicted — has become a buzzword in AI, and is particularly central to fintech regulation and the growth of AI in finance. In the U.S., regulators may also ask if a prospective borrower could anticipate that a lender would consider a certain factor in its decision, so the applicant can take action to avoid being denied credit. If the model uses, say, the first letter of an applicant's surname, the consumer would have few options. (And, of course, if the phone brand proved to be correlated with race, gender or some other factor lenders cannot consider, that would pose fair lending and other problems.)

Potential for Bad Publicity

Finally, reputational risk needs to be weighed. If it becomes public that an institution makes decisions based on

complex "black box" models relying on puzzling alternative data, it could lead to bad publicity. Several years ago, a lender drew criticism for scoring applicants based in part on the chain stores where they shopped. As a result, the company stopped using that factor.

Explain Yourself

In many cases, the best approach will be the common sense one: Make sure your business can explain the relationship between each type of data used and the decisions that result. That will be necessary to satisfy regulators, customers and the public at large.

Authors

Brian D. Christiansen / Washington, D.C.

Ken D. Kumayama / Palo Alto

Sven G. Mickisch / New York

Sonia K. Nijjar / Palo Alto

Simon Toms / London

Austin K. Brown / Washington, D.C.

This article is from Skadden's *The Informed Board*.

View past editions
You can find all *Informed Board* articles here.

A Checklist for Boards

Key steps companies can take to identify and mitigate the risks of using alternative data and artificial intelligence models:

☐ Make sure the company has reviewed each variable used in the model through a compliance lens. For model variables that are not intuitively associated with the decision at issue, management should challenge modelers to explain why the variable is predictive.

☐ Conduct statistical analyses of models to determine whether some variables serve as a proxy for prohibited biases (race, gender, ethnicity, etc.), and explore less discriminatory alternatives.

☐ Confirm the accuracy of alternative data points on a regular basis and validate that the model is operating as expected with respect to the data to avoid "model drift."

☐ Perform periodic risk assessments, with a focus on the impact of such data on outcomes for groups that are protected by law.

☐ Document the results of these analyses and any action plans that grow out of them.

☐ Make sure that the use of alternative data does not violate data privacy laws, contractual restrictions on its use (*e.g.,* nondisclosure agreements) or intellectual property rights. For instance, "scraping" data from public websites without permission may infringe intellectual property rights or violate terms of use.