

FinRegLab and researchers from the Stanford Graduate School of Business release new research on the explainability and fairness of machine learning in consumer lending

WASHINGTON, D.C., April 27, 2022 - FinRegLab and Professors Laura Blattner and Jann Spiess of the Stanford Graduate School of Business have released new research on "[Machine Learning Explainability and Fairness: Insights from Consumer Lending.](#)"

This research focuses on the implications of using machine learning for extending consumer credit by analyzing tools for diagnosing and managing questions about the transparency and fairness of lending algorithms. Machine learning models have the potential to increase credit access by more accurately identifying applicants who are likely to repay loans and to reduce the number of people given loans that they are unlikely to repay. But because these models can be more complicated to analyze and manage, model transparency has become a critical threshold question for both lenders and regulators.

The research empirically evaluates model diagnostic tools to help lenders address these transparency challenges and manage machine learning underwriting models as required by law. Specifically, the research assesses how these tools generate disclosures explaining why someone was rejected for credit or charged higher prices and what factors in the model drive differences in model predictions among different demographic groups.

The project evaluated model diagnostic tools from seven technology companies—[ArthurAI](#), [H2O.ai](#), [Fiddler AI](#), [Relational AI](#), [Solas AI](#), [Stratyfy](#), and [Zest AI](#)—as well as several open-source tools on a spectrum of credit underwriting models.

The research found some of the diagnostic tools can identify factors in the model that drove different aspects of the credit model's behavior, although no one tool performed the best across all tasks. The results suggest that it is important to interpret tool outputs carefully, particularly given that many features in the model are correlated.

"Our research suggests that some of these diagnostic tools can help lenders address transparency challenges associated with machine learning underwriting models," says FinRegLab CEO Melissa Koide. "But the results also emphasize that responsible use of these tools adds an important dimension to the many consequential decisions that lenders must make—and will be accountable for—when they adopt machine learning to extend credit."

The research used consumer credit as a case study in part because federal regulatory requirements already make model transparency a key consideration for any lender that wants to use machine learning underwriting models. The analyses are also potentially useful to other sectors where machine learning predictive models are being used to make important decisions, such as medicine, criminal justice, and employment.