

26 JAN 2023

AI-based credit default risk assessments require regulation

Katja Langenbucher: The EU is still struggling to find an appropriate regulation. The issues at stake go beyond current anti-discrimination law



Artificial intelligence (AI) allows us to predict future events that until now “only God knew about,” as Mikio Okumura, president of the Japanese insurance group Sampo, has recently put it. The company had collected data on the lifestyle choices of residents of nursing homes and examined them for their correlation with the onset of dementia. This way, he claimed, we not only achieve progress in medical prevention, but also improve the risk-adjusted calculation of insurance premiums.

The world of consumer lending, which has been conquered by fintechs, is already making use of these kinds of predictions. Less surprisingly, data on account movements, purchasing and payment behavior provide information about how the cash flow of a potential borrower is developing. However, age, gender or place of residence can also supply relevant information for assessing credit default risk. The same is true for educational or migration history, religious denomination, or party affiliation. Platforms that deal with the assessment of credit default risks do not stop there. Relevant correlations may also result from friends on social media, taste in music, or how people spend their free time. The same can be said about the brand of cell phone someone uses, the number of typos in text messages, or the time it takes to fill out an online form.

There are benefits for both sides

Lenders as well as borrowers benefit from this combination of machine learning and access to “Big Data.” For lenders, an important information asymmetry is mitigated. Borrowers can receive a contract that would not have existed based on a conventional evaluation. Last year the German private credit bureau Schufa had made plans for an account data-based “Check Now” procedure, which was designed to make it easier for customers with poor credit ratings to sign mobile communications contracts. However, the procedure was immediately discontinued because consumer advocates raised data protection concerns.

The European legislator is aware of such developments. The [draft EU regulation](#) on artificial intelligence (AI) is one important piece of legislation. A second piece of legislation is the reform of the Consumer Credit Directive. Recital (37) of the proposed AI Regulation maintains that the use of AI merits “special consideration” in the case of “access to and enjoyment of certain essential private and public services.” The EU legislator regards this as particularly relevant if AI systems are used to “evaluate the credit score or creditworthiness of natural persons” because they determine those persons’ access not only to financial resources. The results from algorithmic decision-making might also affect the provision of “essential services such as housing, electricity, and telecommunication services”. Those AI systems have been classified as “high-risk” due to their central impact on citizens’ access to the safeguards of modern civil society.

In this context, the AI Regulation applies only to the developer and the user of the AI system. The situation of the end user warrants the label of a high-risk system; however, it is outside of its scope. Among other things, this leads to special compliance requirements: Like product regulation, quality assurance measures are defined, certification procedures introduced, and the responsibilities of supervisory authorities specified.

Risks of discrimination

The recital mentioned above assumes that AI systems used to assess creditworthiness pose a risk of discrimination. Drawing on research findings in computer science, historical bias is said to perpetuate experiences of discrimination “based on racial or ethnic origins, disabilities, age, or sexual orientation.” This is due to the peculiarities of machine learning. An AI system learns based on whatever empirical data its developer provides it. Thus, it creates the profile of a successful borrower based on people who have been able to service their loans in the past. Put differently: members of a group of people who have had difficulties obtaining credit in the past will also initially be classified as high-risk applicants by the AI system. Certain members of these very same groups may nonetheless benefit from the AI assessment (invisible primes). A potential borrower whose nontraditional parameters are more similar to those of classically successful candidates has a higher chance of receiving a loan due to this match, despite having a nontypical profile. The example of an immigrant or a young person illustrates this. Their credit history may not (yet) match that of a candidate evaluated according to conventional models. However, other variables that the AI has identified as relevant may align with a previously successful profile. They may use a cell phone from a high-priced manufacturer, have a prestigious university education, or be particularly quick with online forms. Stepping into precisely this gap in the market, some fintech companies have been able to help expand the volume of loans to previously disadvantaged groups.

However, not all people benefit to the same extent from this form of inclusion. In the United States, several empirical studies have been able to show that historically disadvantaged groups benefit from the expansion of loans to a much lesser extent. This may be due to the AI biases described above, if these groups do not exhibit any of the parameters that characterize the traditionally successful group. Sometimes the AI emphasizes characteristics that are found in most of the successful groups. In such cases, other characteristics that would be more significant for the minority may not be weighted sufficiently (majority bias). Again, illustrated by an example: Loans that have been successfully repaid in the past have a positive effect on creditworthiness. But whether properly serving a buy-now-pay-later-agreement or even a private credit agreement leads to the same effect depends on whether the model's metric is programmed accordingly. This may disadvantage younger or inexperienced borrowers.

Right to human intervention

As to the legal classification of discriminatory credit practices, the AI Regulation refers back to the jurisdictions of Member States. However, the reform of the Consumer Credit Directive explicitly addresses this risk. Art. 6 makes sure that “consumers legally resident in the Union are not discriminated against on ground of their nationality, residence or on any ground as referred to in Article 21 of the Charter when requesting (...) a credit agreement.” Moreover, the directive stipulates that consumers have the right to request human intervention when credit scoring involves automated processing. Additionally, Art. 18 mandates “a meaningful explanation of the assessment of creditworthiness.”

However, the ban on discriminatory credit practices might create more problems than it solves. So far, European law recognizes two forms of discrimination: direct and indirect discrimination. Direct discrimination occurs when an individual is treated less favorably because of a protected characteristic than a person without that characteristic in a comparable situation. If a lender determines, for example, that women statistically exhibit a higher credit risk than men, they are nevertheless precluded from training their AI model in such a way that, for the sake of simplicity, all women are assigned a risk premium.

Question of comparability

Indirect discrimination is much more relevant in practice. It takes place when a neutral feature, such as part-time employment, leads to a discriminatory output, for instance between men and women, if part-time employees are predominantly female. Of course, the neutral variable is not prohibited per se. However, the discriminatory output entails follow-up questions, for instance whether the discriminated set of persons is similarly situated to the rest and whether there is an objective reason for the unequal treatment.

Indirect discrimination examines output, which is triggered by a seemingly neutral variable. If using this variable cannot be justified, the law requires to stop using the variable. However, for algorithmic decision-making, this might not help. AI models redundantly encode information. Blocking one neutral criterion will often see the AI use a different variable to fill in, allowing for the same prediction. For example, not only does working a part-time job correlate with gender, but also a certain height, first name, taste in music, or how to spend free time. A sophisticated AI system will discover such variables and establish correlations between entire sets of them.

Put differently: If the system developer restricts input, it is highly likely that substitute variables (proxies) will be found. A possible solution would be to perform many rounds in which more and more variables are deleted from the dataset. Usually, this would result in lower quality of the model's prediction. The inclusive potential of AI scoring, albeit limited, would shrink, and models that use broader data to make more accurate predictions would be penalized.

The suitability of anti-discrimination doctrine for algorithm-based lending is doubtful. This has to do with redundant encoding but extends to labelling of protected groups. In the future, imbalances may arise among entirely unexpected groups, such as people who update software frequently or hardly ever, who are present on social media or refrain from doing so, who charge their cell phones carefully, or whose batteries frequently die. Bans on discrimination only extend to these groups- by chance - if the group correlates with a protected characteristic.

Challenges of private enforcement

Consumers find themselves in a situation that appears Kafkaesque: they do not know which data are relevant for their assessment, hence, cannot respond by changing their behavior. The user of the AI has no incentive to disclose relevant variables. For one, these are usually proprietary information, and for another, a change in the borrower's behavior could reduce the informative value of the variables in question. If the customer learns that installing a dating app is detrimental to his credit score, while using a trading app is beneficial, he may delete the former and download the latter (gaming the system). If his behavior remains unchanged, the assessment will be wrongfully adjusted.

Accordingly, a regulatory design for algorithmic credit scoring will reach beyond anti-discrimination law. Core components include quality control, as outlined in the AI Regulation. This concerns the quality of the AI system and the reliability of the data. Data collected on social media is prone to misunderstandings and errors. The means provided by the General Data Protection Regulation (GDPR) to demand access and, if necessary, rectification of inaccurate data require that consumers be provided with clearly understandable information, as well as efficient legal enforcement procedures.

Starting the discussion

Scholarly and public discussion of these issues is still in its infancy. The Consumer Credit Directive stipulates that “[p]ersonal data, such as personal data found on social media platforms or health data, including cancer data, should not be used when conducting a creditworthiness assessment” (Recital 47). However, if a sophisticated AI can readily reconstruct such information by use of substitute variables, such as frequent Google searches for specific diseases or medications, consumers will hardly benefit from this.

Another example is AI-based price discrimination. The Consumer Credit Directive, surprisingly, seems open to personalized pricing (Recital 40). Granted, in some situations this may create efficiency gains for society. However, it is important to remember that AI systems can detect not only “invisible prime” candidates, but also particularly vulnerable potential borrowers. Inexperienced borrowers who are poorly educated in financial matters, or individuals who are in particularly dire need of credit, will appear to the (morally agnostic) AI as an attractive market opportunity for a high-cost loan. Somewhat reassuringly, the directive includes the requirement for Member States to cap interest rates, APR, and total cost of the loan (Art. 31), curbing predatory lending strategies.

Katja Langenbucher is Professor of Civil Law, Commercial Law, and Banking Law at the House of Finance of Goethe University Frankfurt and coordinates the LawLab – Fintech & AI as SAFE Bridge Professor.

This contribution was first published in the German newspaper "Börsen-Zeitung".

Blog entries represent the authors' personal opinion and do not necessarily reflect the views of the Leibniz Institute for Financial Research SAFE or its staff.



Prof. Dr. Katja Langenbucher
Coordinator LawLab – Fintech & AI



BACK

NEWS & LATEST

All News

SAFE Finance Blog

Events

Videos

SAFE in the media

Newsletter

Research Magazine Points

Distribution List

In this Section:

FOLLOW US



CONTACT

Leibniz Institute for
Financial Research SAFE

Theodor-W.-Adorno-Platz 3
60323 Frankfurt am Main

Phone: +49 69 798 30080

Fax: +49 69 798 30077

Email: info@safe-frankfurt.de

SAFE

[About SAFE](#)

[Job Offers](#)

[How to find us](#)

POLICY CENTER

[SAFE Senior Fellows](#)

[Policy Publications](#)

[SAFE Finance Blog](#)

[Policy Center Team](#)

RESEARCH

[Research Departments](#)

[Researchers](#)

[Data Center](#)

NEWS & EVENTS

[All News](#)

[Press](#)

[Expert List](#)