June 1, 2021

Dave Uejio
Acting Director
Consumer Financial Protection Bureau
1700 G Street, NW
Washington, DC 20552

Re: Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning

Dear Acting Director Uejio:

Consumer Reports (CR) writes today in response to the Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning. Consumer Reports is an expert, independent, non-profit organization whose mission is to work for a fair, just, and safe marketplace with and for all consumers and to empower consumers to protect themselves.\(^1\) We applaud the Bureau for collecting information on artificial intelligence (AI) and machine learning (ML) because these products have the potential to discriminate in financial services as well as create significant financial consumer harms. Our concerns about the use of AI and ML in financial services are not unique to technology. They are about fairness. AI, when training data is biased, or when algorithms are flawed due to human biases, can reproduce and further entrench existing harms, or create new ones. As AI becomes more integrated into financial services, it is important for the Bureau to set clear rules for its use.

We recommend that CFPB ensure increased transparency to consumers, testability and explainability requirements, and fairness in the algorithm design process. Specifically:

\(^1\) CR works for pro-consumer policies in the areas of financial services and marketplace practices, antitrust and competition policy, privacy and data security, food and product safety, telecommunications and technology, travel, and other consumer issues in Washington, DC, in the states, and in the marketplace. Consumer Reports is the world's largest independent product-testing organization, using its dozens of labs, auto test center, and survey research department to rate thousands of products and services annually. Founded in 1936, Consumer Reports has over 6 million members and publishes its magazine, website, and other publications.
• The use of algorithms should be transparent to the end users. When algorithms make decisions about consumers the individual should have notice that an algorithm was used.

• Algorithmic decision-making should be testable for errors and bias. Algorithms should be able to be tested by outside researchers and investigators.

• Algorithms should be designed with fairness and accuracy in mind. Companies should not simply rely on outsiders to detect problems with their algorithms; instead, companies should be required to plan for and design to avoid adverse consequences at all stages of the development of algorithms. This includes but is not limited to: training data usage, model design, and testing procedures.

• Algorithmic decision-making should avoid the use of data sets as proxies for protected attributes. Algorithms can only serve to address the question posed to them. Where possible, algorithms should avoid using factors that can serve as proxies for prohibited factors such as race. Factors such as zip code and education have been found to serve as such proxies, so should only be used in algorithms where information about education or address are essential. As the use of alternative data rises, it is important that these new data points are carefully chosen to avoid acting as prohibited factors.

• Algorithmic decision-making processes that could have significant consumer consequences should be explainable. In some cases, algorithms are programmed to learn or evolve over time, such that a developer might not know why certain inputs lead to certain results. This could lead to unfair results if there is no meaningful accountability for how decisions are made. If an algorithm is (1) used for a significant purpose, like the determination of a credit score and (2) cannot be sufficiently explained, then the process should not be used.

• If the use of AI or algorithms in financial services leads to less fair or inclusive outcomes, it should not be used. Much of the innovation in the financial space attempts to include underserved populations in services previously unavailable to them. If these algorithms result in further discriminatory impacts to the same or different populations, they should not be used.

Background

AI is a broad term that means using data to make predictions or classifications about future data points. An algorithm is simply a set of instructions to make these predictions and classifications. Data is used to train an algorithm so that it can make more accurate decisions. The key to all of this: the algorithm can only be as good as the quality of the data it is fed. Certain kinds of algorithms have been around for decades and have been commonly used for statistical modeling.

However, newer types of algorithms tend to be less explainable to the public and even to the engineers who design them, particularly when the task is more complicated — like trying to classify an image as containing a certain object or even predicting human behavior. In financial services, algorithms are routinely used to determine auto insurance rates, creditworthiness, and willingness to pay, to name just a few examples.

Proponents advocate for the use of artificial intelligence in financial services, claiming it can “reduce human biases and errors.” Algorithms are often positioned to consumers, regulators, and financial institutions as expanding access to financial services and/or decreasing bias in the provision or pricing of services. Claims of objectivity and proof notwithstanding, algorithms can and sometimes do exacerbate bias or have unexpected discriminatory effects, as numerous examples have demonstrated. While there are laws that prohibit discrimination, there are not laws in place that ensure sufficient transparency, testing or accountability of algorithms.

Algorithmic discrimination occurs when an automated decision system repeatedly creates unfair or inaccurate outcomes for a particular group. CR has documented numerous areas in which the use of algorithms raises questions of discrimination. For example, CR research shows that when auto insurers use factors such as education, job status, and zip code to price policies, consumers of color pay higher prices than risk can explain.

There are many sources of bias when designing an algorithm to complete a certain task, but many of them revolve around human error. Non-inclusive datasets (datasets that may not fully represent the populations the algorithm is trying to make decisions for) or biased data collection methods can lead to poor outcomes in algorithmic decision making for those who are underrepresented in the training data. Other types of error can arise from the specific type of model being used as well as

---


6 See for example, LendUp: “We consider all types of credit history. Just because your credit score may be "not-so-great" doesn't mean you can't get approved.” [https://www.lendup.com/](https://www.lendup.com/)

7 “Artificial intelligence (AI) presents an opportunity to transform how we allocate credit and risk, and to create fairer, more inclusive systems.” Aaron Klein, Brookings Institution, Reducing bias in AI-based financial services, [https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/](https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/).


the attributes of the data the engineer chooses as being important to the final outcome. Overall, while many companies claim that their algorithms lack bias and are significantly better than humans for decision making, there are still many steps in the process of designing algorithms that require human intervention. While many algorithmic technologies are improving, they will likely never be perfect.¹⁰

A major reason why algorithms can perpetuate discrimination against minorities is due to biases that often stem from societal inequities. For example, Black borrowers are more likely to be sued on debts,¹¹ and are therefore overrepresented in debt collection data. Due to systemic racism, Black and Latinx Americans are more likely to have damaged credit or a lower credit score compared to their white counterparts,¹² and are more likely to be sold high-cost, unmanageable loans.¹³ Black and Latinx Americans are more likely than white or Asian Americans to lack a credit score, also known as being “credit invisible.”¹⁴

Algorithmic bias will likely be a crucial area of policy in the near future as these technologies become more common in everyday life. CR is committed to racial justice, fairness and greater transparency in addressing bias in algorithms. The current lack of regulation surrounding algorithms has created a "Wild West" for many companies using AI and has the ability to do major damage to marginalized communities and consumers in general.

Against this backdrop, it is urgent that the Bureau take action.

**Risks of AI and ML in Financial Services**

Financial institutions have begun to use algorithmic decision-making and artificial intelligence for a variety of applications including credit scoring, lending decisions, fraud detection, and personalized financial advice. Unfortunately, some of these applications have the potential to be discriminatory towards certain communities due to algorithmic bias. Because many companies treat their algorithms as trade secrets and there are few laws regulating the transparency of algorithms in the financial policy space, many of the methods and practices of these companies are not being scrutinized or regulated appropriately. The data collection processes, model usage, and

---


testing procedures for mitigating bias and ensuring accuracy are often opaque to both policymakers and consumers. We will discuss how risks and biases arise in algorithms, why they can be discriminatory or otherwise problematic, and appropriate legal and practical mitigation strategies.

A. Biases in Data

Bias in artificial intelligence and other automated decision-making can stem from a variety of factors. Training data, the data that algorithms use to learn from, can often be one of the culprits. Datasets that are not inclusive of the population the algorithm is trying to make decisions about can contribute to skew in algorithmic decisions, and this can come from biased data collection methods or lack of data on certain populations which can often stem from societal inequities.\(^\text{15}\)

Another aspect of the algorithm design process is feature selection, when the engineer chooses what kinds of data should inform the output. For example, an engineer at Company A can design an algorithm that attempts to predict risk for auto insurance pricing such that it takes an input of someone's credit history and their driving record. Often, feature selection can be rather arbitrary — an engineer at Company B that designs an algorithm with the same goal can choose, for example, car make/model and current income as inputs for their algorithm. Ultimately, the data that ends up being used as predictors really depends on the algorithm designers as well as the availability of data. Feature selection can contribute to discriminatory outputs if not chosen thoughtfully, while also making sure that the features chosen are not proxies for protected classes (if prohibited) like race, gender, etc.\(^\text{16}\) An individual might get rated as lower risk by one company compared to another, simply because of the data being used to make this decision — for example, a someone with a good credit score and driving record but with an older car might be rated more highly by Company A than they would by Company B since their older car might hurt their risk score, and Company B disregards credit history and driving record.

Overfitting is another issue that could lead to bias or discriminatory outcomes in financial models. Overfitting occurs when a model is trained on data that is not representative of the larger population the model is designed to make decisions about; therefore, the model is not as accurate for other segments of the population that were not included in the training data.\(^\text{17}\) For example, if an app uses artificial intelligence to advise consumers on how to invest in stocks and only uses data from 2007 to make predictions about how the markets behave, then consumers might lose money since the market in 2021 might behave much differently than it did in 2007. The app would likely give financial advice only relevant to that particular year rather than taking into account overall market trends over many years. Overfitting could also lead to discriminatory outcomes in


\(^{16}\) Barocas and Selbst, "Big Data's Disparate Impact."

lending or credit scoring algorithms if only certain parts of the population are represented in training data, and could classify those underrepresented as "higher risk" which can prevent them from accessing credit when they need to.

B. Use of Alternative Data and Models

Some alternative credit score companies claim that using nontraditional data and modeling techniques can be more inclusive of those who have historically been "credit invisible." However, more research needs to be done in order to ensure that new ways of assessing and providing people with credit opportunities are both fair and inclusive. Groups like FinRegLab are looking at ways that alternative data such as using information from banks and small business software to expand access to those historically left out by traditional credit scoring models. But, it is overall unclear how alternative credit score companies are dealing with risk management of using alternative data and whether they are using data control processes. This data can include information that has not traditionally been included on credit reports such as social media activity, internet browser history, utility bill or telecom payments, and educational background. Not only does this raise privacy concerns that could lead to a chilling of free expression, but there is little evidence that these types of data are actually effective in calculating credit risk.

Alternative data companies making these evaluations may be pulling information from datasets that might be incomplete or non-inclusive of the populations for which they are making decisions. Under the Fair Credit Reporting Act (FCRA), the credit reporting agency and the information provider are responsible for correcting errors on a consumer report. However, some of these alternative bureaus try to dodge being classified as a credit reporting agency in attempts to avoid regulation under the FCRA, and there is some uncertainty among the courts in determining whether some of these bureaus are covered by the FCRA. Due to the legal ambiguity, it is unclear whether these companies are required to provide pathways for consumers to correct errors in their reports or even notify consumers what kinds of data is being collected and how their score is

---

calculated. It is possible that some companies are collecting thousands of data points and processing them in a manner that further complicates transparency to the consumer (for example, processing and reducing something as complicated as social media usage to arbitrary data points), even if they were given access to this information. Furthermore, the types of models these companies are using might be more complex than traditional scoring models, which could further preclude transparency on how the algorithm arrived at its final result.

C. Model Type

The type of model chosen for decision-making can often contribute to an increase or decrease in accuracy rates. For example, standard models used in machine learning like linear and logistic regressions have been used in statistical decision-making for decades and are also relatively easy to understand. The public is able to interpret how the factors under consideration led to the decision. The downside of these model types is that their simplicity can lower accuracy rates for certain kinds of data — although it is more straightforward to explain why the model doesn't work for certain individuals or groups. Neural networks, a newer and more complicated model type, are often used for more complex decisions due to its potential for higher accuracy rates. The downside here is that the model itself is relatively opaque — even to the engineers that use them. It is near impossible to identify exactly why a decision can be inaccurate or wrong when a neural network is used because there are often thousands if not millions of data points being fed into the system. For this reason, we recommend against using these systems for critical decision-making like credit scoring and lending decisions that must be transparent. Overall, there is sometimes a trade-off between accuracy and interpretability, something financial institutions need to keep in mind particularly for more sensitive applications that affect people's access to basic resources like credit. Regardless, any financial product or service that affects someone's ability to access financial opportunities needs to be interpretable, whether AI is used or not.

D. Dynamic Updating

Algorithmic decision-making is made even less transparent by the fact that the algorithms are often changing, a process referred to as dynamic updating. As an algorithm is fed more training data over time, it can adjust its behavior in response. For example, if engineers build a system that provides personalized advice on how individuals can adjust their spending habits to save money

---

26 Sharayu Rane, "The Balance: Accuracy vs. Interpretability."
and discover it is more accurate for people who do not spend much on transportation, they might look to ensure that their training data adequately represents people who DO spend significantly on transportation. As engineers adjust training data and other parameters, the algorithm's outputs can differ even with the same inputs. Similarly, if an algorithm is found to have discriminatory impacts, the company can adjust that algorithm to mitigate the issue.

Due to the lack of transparency requirements, companies do not need to inform consumers or others about the changes they are making. This can be concerning for victims of discrimination who cannot find out the details of the — since modified — algorithm that had caused them harm. For purposes of regulatory oversight, it will be essential to set strong requirements for public documentation which should include updates to training data, other adjusted parameters, and how the algorithm learned over time.

E. Model Accuracy and Testing

Poor accuracy rates of an algorithm can have significant harmful impacts on consumers. An example of this is automated savings apps that, when authorized to access a person's savings and checking accounts, claim to analyze income and spending trends and move money to a savings account to help a customer save money over time. However, there are risks associated with this business model — it is possible that this algorithm could cause an overdraft if a customer withdraws more than the model predicts they would. For individuals or households with a lower or unstable income, this could be particularly dangerous. According to data from 2015, 60 percent of households experienced a financial shock within the previous year, and the median household spent half a month's income on its most expensive financial shock. It is unclear to what extent automated savings companies are taking into account variability in spending in their algorithms, especially in the cases of a financial shock. Automated savings companies like Digit acknowledge that overdrafting due to their algorithms is a potential risk (although they claim the risk is "very unlikely"), and state that they will reimburse insufficient funds fees for the first two instances of overdrafting due to their algorithms. However, it is unclear how companies like Digit test their algorithms and put guardrails in place to take into account and mitigate the impacts of outliers.

---

Furthermore, there is also the issue of "snake oil" that is unfortunately common in several companies in the AI space, including some in the financial sector. Some companies claim that their technology is capable of doing certain things that are unsubstantiated by science, or claim that the correlations between different phenomena are actually due to causation. For example, companies like Lemonade claim that they can detect insurance fraud from videos of individuals discussing insurance claims based on how they speak and/or look. While certain insurance issues may not be under the jurisdiction of the CFPB, without more regulation there could be similar types of unsubstantiated or pseudoscientific claims made by companies in the broader financial space.

Due to the lack of transparency and dynamic nature of AI, it is difficult to hold companies accountable for the harms that inaccurate models can cause. As there are real financial consequences for individuals who rely on services that use AI, stronger testing and transparency standards are needed to ensure robust accuracy rates and prevent harm.

F. Use of Third Party AI

The main concern when using AI developed by third parties is that the algorithm can essentially be a black box. All the problems with algorithms being opaque and potentially biased are exacerbated when those algorithms are used by or licensed to third parties that have even less visibility into how the algorithms work and less ownership over the accuracy of the results. If an institution is concerned about bias arising in its processes when using a third party's AI technology, it needs information from the third party about how the model was designed including information about the training data, model type, how they tested for bias, etc. It is often not the case that third parties provide this information to the groups they sell to, as companies treat their technology and algorithms as trade secrets and there are few requirements mandating this disclosure.

Uncertainty in Application of Fair Lending Laws

Due to the lack of algorithmic transparency and testing standards, it is unclear how financial institutions' use of algorithms complies with existing fair lending laws. The potential for discrimination has been previously discussed, but it is difficult to say to what extent these companies are testing for and mitigating disparate impacts.

The Equal Credit Opportunity Act (ECOA), which is implemented by Regulation B, requires creditors to notify an applicant of the principal reasons for taking adverse action for credit or to provide an applicant a disclosure of the right to request those reasons. Currently, the official

---

32 https://www.lemonade.com/faq
interpretation of §1002.9(b)(2) states that the creditor need not describe how or why a factor adversely affected an applicant (for example, the notice may say “length of residence” rather than “too short a period of residence.”). This allowed flexibility could potentially be taken advantage of by creditors as they adopt more AI tools that further lead to vagueness for credit decisions. Consumers deserve concrete reasons for why they are denied credit, regardless of if AI is used to make that decision. Unfortunately, AI can often make decisions with little insight as to how the algorithm arrived at its final result. The use of AI in determining credit underwriting should not be used as an excuse for creditors to further avoid clarity when taking adverse action. If creditors discover that integrating AI into their decision-making hinders the ability to identify reasons for taking adverse action, they should simply avoid using AI.

Conclusion

Two engineers designing algorithms with the same goal in mind could potentially approach algorithm design drastically differently — including but not limited to dataset and data processing choices, model type, and feature selection — which could therefore affect bias and accuracy rates differently. All of these choices are made by engineers and/or their teams, people who have human biases. Some companies in the financial space claim that their algorithms "can reduce human biases and errors caused by psychological and emotional factors" but this is rarely true. Algorithms involve human input at all stages of the design process, therefore imbuing human bias into automated decision systems. With these concerns, we recommend the following:

- The use of algorithms should be transparent to the end users.
- Algorithmic decision-making should be testable for errors and bias.
- Algorithms should be designed with fairness and accuracy in mind.
- The data set used for algorithmic decision-making should avoid the use of proxies for prohibited factors.
- Algorithmic decision-making processes that could have significant consumer consequences should be explainable.

35 "Bias in AI: What it is, Types & Examples of Bias & Tools to fix it," AI Multiple, April 17, 2021, https://research.aimultiple.com/ai-bias/.
• If the use of AI or algorithms in financial services leads to less fair or inclusive outcomes, it should not be used.

The purpose of artificial intelligence is to attempt to mimic human thinking — ideally, the more the algorithm learns from its environment or by interacting with a user, the better it can perform its task. However, current AI systems operate nowhere near this level and this can have significant discriminatory and otherwise harmful impacts on consumers using financial products. It is important that innovation using AI in the financial space is done carefully. While the use of alternative data and AI/ML has the potential to be more inclusive of populations previously excluded from financial services, it also has the potential to further discriminate against certain groups, which is especially worrisome considering some financial products and services are already out of reach for many. AI brings along new risks that must be mitigated in order for these products to be deployed in a fair and inclusive manner. We urge the Bureau to establish clear rules regarding AI transparency, data usage, and algorithm design requirements in financial services to ensure consumer safety.

Sincerely,

Nandita Sampath
Policy Analyst