



National Institute of Standards and Technology (NIST)
100 Bureau Drive
Gaithersburg, MD 20899

Upstart Network Inc. Comment Re: Four Principles of Explainable Artificial Intelligence (Draft NISTIR 8312)

Upstart operates an online lending platform that enables lenders to offer loans to consumers. The platform includes access to a pricing model built by Upstart that uses both traditional and non-traditional credit variables and advanced modeling techniques, such as the use of artificial intelligence (AI) and machine learning (ML) algorithms, to more accurately price consumer credit. Our partnerships with banks supervised by the FDIC, OCC, Federal Reserve, CFPB and state regulators enable us to observe opportunities for regulators to encourage banks and credit unions' use of technology to expand and enhance the availability of credit in the United States. We believe that the work NIST is doing to examine the future of artificial intelligence is especially important -- and especially urgent -- now, as our country confronts and manages the economic consequences of the COVID-19 pandemic.

Benefits of AI in Credit Underwriting

Upstart believes that AI/ML credit underwriting models can contribute substantially to the recovery of our country's economy by enhancing credit availability in a manner that is consistent with fair lending and other consumer protection requirements, and without undermining bank safety and soundness. Our view is that the emergence of artificial intelligence and machine learning in lending will entirely reshape the banking and broader credit industry in the next 10-15 years. Almost all lending will have sophisticated machine learning (ML) and broad use of (what is today called) alternative data at its core – in customer acquisition, credit decisioning, anti-fraud, verification processes, servicing, and cross-selling. Upstart has found, first in a lab setting and then in commercial practice, that it is possible to have dramatic gains in reduction of loan loss rates without impacting loan approvals. In fact, Upstart has proven that you can reduce loss rates by about 75% while approving the same number of people, compared to traditional institutions. The use of an AI / ML model also offers a path where lenders partnering with Upstart can dramatically expand the pool of people that they are approving for prime credit products without significantly increasing loss rates.

It's important to note that the full benefits of AI in credit underwriting are still untapped. Many institutions powered by Upstart are implementing the Upstart model to expand access to credit only gradually. In 2019, the Consumer Financial Protection Bureau (CFPB) reported that the Upstart model approves 27% more applicants than the traditional model, and yields 16% lower average APRs for approved loans. The data showed expansion of credit access reflected in the results provided occurs across all tested race, ethnicity, and sex segments resulting in the tested model increasing acceptance rates by 23-29% and decreasing average APRs by 15-17%. In many consumer segments, the results provided show that the tested model significantly expands access to credit compared to the traditional model. In particular, under the tested model, the results provided reflect that: "Near prime" consumers with FICO scores from 620 to 660 are approved approximately twice as frequently. Applicants under 25 years of age are 32% more likely to be approved. Consumers with incomes under \$50,000 are 13% more likely to be approved.¹

Upstart also believes that the use of more data, automation, and AI can produce outcomes that can remove human bias. Human biases, even unintentional ones, have long led to unfair disparity in lending that are not easily explained. As the NIST framework points out, human decisions simply don't hold up under the scrutiny of the report's four principles. For instance, in testing in-person small business lending practices at physical bank branch locations in 2018, the National Community Reinvestment Coalition found that Black and Hispanic testers were requested to provide more information than their white counterparts, particularly personal income tax statements. Hispanic testers

¹ <https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/>

were asked to provide them nearly 32% and black testers 28% more frequently than their white counterparts. They also found that White testers were given significantly better information about business loan products, particularly information regarding loan fees where white testers were told about what to expect 44% more frequently than Hispanic testers and 35% more frequently than black testers.²

Comments on NIST Framework

NIST has put forward a proposed four-part framework for standards governing “explainable AI”. Upstart offers brief comments on each, based on our experiences using AI and ML technology to make better lending decisions.

AI systems should deliver accompanying evidence or reasons for all their outputs. As the use of AI expands, Upstart believes that market participants must rigorously test model outputs to ensure accuracy, fairness, and model integrity. We also believe that AI model outputs can be explained, and evidence provided to users. It’s important to note that AI systems may be making large numbers of decisions in moving towards a final output. In NIST’s framework of asking for “evidence or reasons,” Upstart would submit that the relevance of different types of “evidence” or “reasons” depends in part on the targeted audience. A borrower, for instance, is interested broadly in the reason or reasons that they may have received a negative or positive outcome from an AI model, and whether there is anything they can do to change their circumstances to receive a better outcome. While the model developer will be interested in those issues, they will also be interested in the more granular operation of the AI system and the technical evidence that it remains statistically sound. A regulator may be most interested in evidence that the model is performing with accuracy and that it is being operated by model-developers responsibly.

Systems should provide explanations that are meaningful or understandable to individual users. In lending, federal and state regulations require that models produce explanations for adverse actions (both declining a customer application for credit) and pricing decisions. Upstart has found that with its AI model, it is possible to use AI / ML to search for and surface meaningful explanations to users of outputs, accompanied by the evidence that is referenced in standard format. We use AI systems to issue consumers accurate Adverse Action Notices (AANs), so it is important that users impacted by AI model outputs understand how they may be able to change their outcome in the future. In general, Upstart agrees that it is important that the explanations are provided to users in a digestible form that is rank-ordered in a way that is relevant and actionable. For instance, it is not useful for a system to explain itself so comprehensively - all the relevant explanations provided in no particular order - that the impacted user cannot understand what is actionable or not actionable on their part.

The explanation correctly reflects the system’s process for generating the output. We agree with NIST that the explanation for a model output should first and foremost be truthful. As referenced above, different stakeholders may be interested in different types or levels of explanations. Upstart believes that because an AI system’s process for generating an output will often be extremely complex, it may be desirable in certain circumstances for an explanation of an output to be truthful, but focused at a high / general level. It is also possible, however, that such an explanation may not be deemed *complete* by all users or stakeholders who may have different interests in operation of the model. For instance, an explanation given to a borrower who is declined for credit could be technically “correct” but also so comprehensive and so technical that it is not meaningful to that particular end-user, even while being helpful to the model developer. In this instance, the goal is not to accurately explain the entirety of the model, but rather the decision and what can or cannot be changed to yield a different decision.

The system only operates under conditions for which it was designed or when the system reaches a sufficient confidence in its output. (The idea is that if a system has insufficient confidence in its decision, it should not supply a decision to the user.) Upstart believes it is crucial to test models rigorously before putting them into production. Model developers should understand the limitations of their model, and generally aim to use models in proportion to their confidence in the model in any domain. However, in certain practical circumstances, there may be no alternative to using a system to make a decision, e.g. when an applicant about whom no model is confident applies for a loan, a decision still must be made one way or the other to the applicant. It is up to the

² <https://ncrc.org/wp-content/uploads/2019/09/NCRC-Small-Business-Research-FINAL.pdf>

system developer to decide when the model can graduate from a training model and be put into production with real end-users. Through testing, human users gain confidence in an AI system and the system “learns” and renders outputs with greater confidence based on much larger volumes of data throughput. Upstart agrees with NIST that a system should not be put into production until its human operators have sufficient confidence in the system and know that the system is reliably operating and that they can vigilantly monitor and explain its decisions. In the testing phase, developers should monitor the amount of testing done and record the reasons they believe the model is ready to be put into practice.

Key Conclusion

Stepping back, it is important that we continue to seek to responsibly harness the potential of AI. To do so, Upstart believes that human users should not necessarily aspire for AI to solve problems in the same way, or using the same logic, that a human would solve them. A large part of the value of AI to people is that AI systems may approach problems in ways that are different than the way humans might think to approach them. In the context of harnessing, but also explaining AI systems, that means that society should want to ensure that AI developers are asking the questions that are important, and seeking the explainability within the framework of how an AI system “thinks” rather than relying most heavily on the frameworks of how a human would think about solving a similar problem.

Upstart is grateful to NIST for taking up this important topic.