June 29, 2021

By Electronic Submission

Office of the Comptroller of the Currency ("OCC")
[Docket ID OCC–2020–0049]

Board of Governors of the Federal Reserve System
[Docket No. OP–1743]

Federal Deposit Insurance Corporation
RIN 3064–ZA24

Bureau of Consumer Financial Protection
[Docket No. CFPB–2021–0004]

National Credit Union Administration
[Docket No. NCUA–2021–0023]

Re: *Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning*, Billing Code 4810-70-P

Ladies and Gentleman:

BAFT (Bankers Association for Finance and Trade)[1] respectfully submits this letter in response to the Request for Information and Comment ("RFI") on Financial Institutions' Use of Artificial Intelligence ("AI"), Including Machine Learning ("ML") issued jointly by the OCC, Federal Reserve Board, FDIC, CFPB, and NCUA ("federal regulators") on March 31, 2021 in the Federal Register.[2] BAFT greatly appreciates the federal regulators' solicitation of information from the financial services industry about the use cases for and risks and benefits of AI/ML to further support innovation in the sector.

Below, we respond to ten of the seventeen questions posed in the RFI: questions 1-8 and 16-17.

Question 1: How do financial institutions identify and manage risks relating to AI explainability? What barriers or challenges for explainability exist for developing, adopting, and managing AI?

[1] BAFT is a financial services trade association whose membership includes large global and regional banks, service providers, and fintech companies headquartered around the world. BAFT provides advocacy, thought leadership, education, and a global forum for its members in transaction banking, including international trade finance and payments. For over a century, BAFT has expanded markets, shaped policy, developed business solutions, and preserved the safety and soundness of the global financial system. Learn more at http://www.baft.org.

[2] 84 FR 16837 (Mar. 31, 2021).

Question 2: How do financial institutions use post-hoc methods to assist in evaluating conceptual soundness? How common are these methods? Are there limitations of these methods (whether to explain an AI approach's overall operation or to explain a specific prediction or categorization)? If so, please provide details on such limitations.

Question 3: For which uses of AI is lack of explainability more of a challenge? Please describe those challenges in detail. How do financial institutions account for and manage the varied challenges and risks posed by different uses?

Explainability and auditability are a cornerstone of any solution or approach used by financial institutions. Because the services they offer are heavily regulated, financial institutions are advised to use AI- based solutions that adhere to the "FEAT" principles: the solution should be fair, ethical, accountable and traceable.

There are various techniques used in AI solutions, such as ML, natural language processing ("NLP"), neural networks, knowledge base with rules engine, or a combination of multiple models and rules. In assessing any AI-based solution, a financial institution should understand the problem it is trying to solve and what AI techniques are used. For an AI system to be effective, trustworthy data from the past must exist that describes how various inputs result in a particular output. AI (and ML models) are algorithms that learn from the past and predict outcomes for a particular problem.

For example, if a financial institution wishes to review historical transactions for potentially fraudulent behavior, then a ML approach may be appropriate. Models built on historical transactions, as in this scenario, need to include an explanation of types of data used, types of data normalization performed, and acceptable deviation from the standard behavior. This information and any assumptions should be documented and available for later inspection. In addition, any review performed must provide model validation information to the end user in real time. There should be an audit trail of past decisions for use in providing explanations to regulators and internal auditors. As these ML models are further fine-tuned based on the additional information that is made available, a feedback mechanism should be implemented. The steps of train, validate, deploy and provide feedback in a loop to achieve better accuracy and explainability. Operators should validate models and results on a reasonable and representative sample set before deploying an AI solution in production.

In most non-financial industries (excluding medical), the primary use case of AI models has not been solving critical problems. Rather, the AI models make recommendations to customers browsing for products on an e-commerce site, make predictions on social media networking sites, etc. As such, model documentation and explainability have been generally less important. However, with repeated use as time goes by, the importance of such use cases increase (e.g., videos in social media that bring up similar videos repeatedly after reviewing one video, thereby amplifying the message). For use cases in financial services, auditor and regulators will expect rigorous documentation outlining model development as well as model monitoring.

Further, some types of ML models are very difficult to explain at an algorithmic level. It is important to measure and confirm what the model predicted. This is akin to asking a human a question (to which he / she could give a right or a wrong answer) without determining what exactly happened in the brain that led to the answer (i.e., which neurons fired, which did not). Thus, model type (e.g. simple decisions trees are easier to explain than a deep learning model), the context around which the model is used, the robustness of model development, validation, and testing, and the robustness of model outcome monitoring may all contribute to the level of explainability at an 'outcome' level and at an algorithmic level.

Many of the challenges above can be resolved, or controlled, to a large extend by:

- Developing a well-defined body of knowledge and setting industry standards on AI definitions, model development techniques, model risk assessment techniques and explainability at two levels: outcome and algorithmic.
- Applying appropriate rigor in model development and model monitoring standards, depending on the criticality of the use case.

For some particular AI use cases, lack of explainability can be more of a challenge. Some of the challenges include quality of data, data categorization, data normalization, and understanding the meaning associated with each data element and interrelationships. Frequently, a combination of AI techniques are better suited to address some of these challenges. For example, NLP along with knowledge base systems are better suited for back office compliance at financial institutions. These AI techniques attempt to mimic human reasoning to assign meaning to various data elements and, if there is a conflict, resolve it by looking for additional information available in transactions. Once a conflict is resolved, these systems can easily provide an explanation for the decision. Any unresolved elements are left for the operator to resolve with assistance from an audit trail. To generate detailed explainability and avoid false analysis, firms should use a good knowledge based system and fine-tuned, targeted, and limited use NLP algorithms for a specific scenario.

Any financial institution that solely depends on ML based model solutions may encounter the challenge of valid results without explainability. For example in a loan approval process or credit card authorization process, it is important that models obtain correct, annotated, or tagged data along with closed loop feedback mechanisms to ensure constant learning and fine tuning.

ML models learn essentially from decisions made or behaviors / activities undertaken by humans in the past. Users should consider the implications of relying on outcomes that could change over time, but are influenced by the latest decisions and behaviors that are themselves derived from the use of ML models. As time passes, the risk that future ML model predictions will be based on inputs that its predecessor ML model versions have learned from behaviors influenced by its own outcomes. As an example, 20 years ago, humans acquired knowledge by reading books, attending classes by teachers who had spent a lifetime acquiring knowledge that had been tested and clarified over decades, considering different view-points and debates, pushing the boundaries of understanding of the subject by adding one's own views to it, all within a framework of value systems. Further, dissemination of knowledge for the most part was limited to the written word - newspapers, magazines, research papers and other documentation. The ML models of the past 20 years have had the advantage of learning from the 'shoulders of giants.' However, today, we see the proliferation of knowledge spread widely through ML-model-based social media, search engines, etc. that is not necessarily accurate or well researched and that results in poor outcomes for humanity. ML-models amplify the behavior of the user, and over time, may result in human thinking getting re-wired and fed by the ML-models. This could 'normalize' the perceived reality and reinforce or cement behaviors.

In the context of financial services, it is important that we prevent such polarization of outcomes over time. One solution could be maintenance of a 'reference' standard: a retention of knowledge that is not overly influenced by amplified views from ML models and is preserved within the framework of certain focused on the right outcomes for humanity. Ensuring that critical models continue to perform well against such 'reference' standards may help reduce the long-term shift in the knowledge space and ensure fair outcomes.

Question 4: How do financial institutions using AI manage risks related to data quality and data processing? How, if at all, have control processes or automated data quality routines changed to address the data quality needs of AI? How does risk management for alternative data compare to that of traditional data? Are there any barriers or challenges that data quality and data processing pose for developing, adopting, and managing AI? If so, please provide details on those barriers or challenges.

The risks associated with data quality and processing are inherent with the use of any technology and not unique to AI. AI models may use larger volumes of data, however, which may in turn increase the costs of effectively controlling the risks.

Using a risk-based approach to model validation and testing, financial institutions may enhance their continuous monitoring of model data and performance, depending on the nature of the algorithm or the use case. AI models themselves are particularly effective at performing these data quality checks to identify missing information or bad data. The output of the models and the related metrics will also provide vehicles to measure changes based on tolerance.

Question 5: Are there specific uses of AI for which alternative data are particularly effective?

Generally, there is value to having more data (including alternative data) available for use by AI when the data is relevant, available in an appropriate amount, and of reasonable quality. For instance, if only 200 of one million data sets have an IP address, that alternative data will not be helpful in modelling.

The definition alternative data as used in the RFI is very broad. Clearly, some data sources not typically found in a consumer reporting agency file can be very relevant in making a decision, such as bank account transaction history, rental payments, employment history, etc. The value of these data sources is dependent on the application / use case. Credit decisioning is an obvious example for which alternative data sources are very effective. Identification of fraud and the risk of fraud are other examples in which use of alternative data such as email, IP address, etc. can be extremely valuable and informative.

The use of alternative data by many potential applications needs further careful study to ensure absence of bias or other unfair outcomes. Additional regulatory safe harbours for the use of certain alternative data in proven and identified use cases may encourage further adoption by market participants.

Question 6: How do financial institutions manage AI risks relating to overfitting? What barriers or challenges, if any, does overfitting pose for developing, adopting, and managing AI? How do financial institutions develop their AI so that it will adapt to new and potentially different populations (outside of the test and training data)?

Over- and under-fitting is a well-known challenge for model development. Methods for recognizing and avoiding these pitfalls are replete in academic literature, including studies on cross validation, sampling techniques, etc.

Over and under-fitting is a key validation point for a model risk management framework in a bank. These pitfalls are reviewed during model validation and before a model is approved for use in the field. Model validation includes reviewing an on-going monitoring plan to be executed alongside the use of the model. After the model is in use, on-going monitoring of the model will involve periodic reporting coupled with manual checks that measure model performance against predefined thresholds for acceptable operation. On-going monitoring metrics are reviewed by the business control process for the model's use.

Solutions that are known to operate on populations not found in the training and testing data will introduce model risk. Depending on the context, additional controls and tooling are added to the control process that include human data check steps before using the output of the model. For example, many solvable problems using NLP techniques have an accuracy of around 80 to 90%. These less than perfect models can be incorporated in semi-automated workflows to provide higher accuracy faster than humans alone.

Question 7: Have financial institutions identified particular cybersecurity risks or experienced such incidents with respect to AI? If so, what practices are financial institutions using to manage cybersecurity risks related to AI? Please describe any barriers or challenges to the use of AI associated with cybersecurity risks. Are there specific information security or cybersecurity controls that can be applied to AI?

AI and AI models are subject to the same level of vulnerabilities and risk as other software assets within any business. In many instances, the core AI models have a higher risk as they access high volumes of complex data. As we increase our dependence on AI models, we increase the incentives for attackers to target those AI algorithms. A successful attack can have severe consequences, which can take a long time to detect. Therefore, businesses should apply the same or higher levels of core data security to the AI model as they apply to all sensitive data. This should include the testing data used to train a model.

Data scientists are not security experts and cannot be expected to assess the risk and potential vulnerabilities throughout the AI algorithm development process. The data scientists try to create new meaningful results, which often requires combinations of different sources of sensitive data within their models. Many will use a combination of open-source code (a risk that must be evaluated), plus new code when creating an AI model. A best practice is to keep a full inventory of all AI models / algorithms and conduct regular audits and testing and ensure a clear understanding of the level of sensitive data inputted to the model, how and what test data was used to train the model, and the parameters and weightings within the model. Regular model testing should be a best practice. Each component of the model has a set of security risks, which have to be taken into account and continually monitored.

Question 8: How do financial institutions manage AI risks relating to dynamic updating? Describe any barriers or challenges that may impede the use of AI that involve dynamic updating. How do financial institutions gain an understanding of whether AI approaches producing different outputs over time based on the same inputs are operating as intended?

As noted in the RFI, some risks regarding AI are known and not unique to AI. These risks include: "AI induced operational vulnerabilities, such as internal process or control breakdowns, cyber threats, information technology lapses, risks associated with the use of third parties, and model risk, all of which could affect a financial institution's safety and soundness." The risks are rooted in financial services operations and do not warrant unusual or heightened sensitive from the AI community, financial services industry, regulators, thought leaders, or policy makers.

Other risks are of significant concern: "Consumer protection risks, such as risks of unlawful discrimination, unfair, deceptive, or abusive acts or practices (UDAAP) under the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act), unfair or deceptive acts or practices (UDAP) under the Federal Trade Commission Act (FTC Act), or privacy concerns." These risks are the consequence of rules (transparent and hidden) and, in part or whole, subject to subjective interpretation and tolerance.

Many financial institutions are called on to pre-publish aspects of their processes including RPA, AI, scorecards, etc. The concept of simply changing AI models that fulfil a scoring function, as an example, will not be acceptable to all regulators.

Dynamic updating could potentially be catastrophic unless the change parameters: define the implications of a prediction / classification change in a process and are governed (rate of change) in terms of change consequence (cost, capacity, risk). These implications are potentially complex as they could change based on the nature of the inputs (e.g., using AI to score companies where the consequence of rejecting a large corporate is different from that of a small corporate).

Institutions should look to their data scientists to build varied approaches to modelling. These approaches should be assessed against effectiveness within the different stages of processes in which they are deployed. For example, the effectiveness of AI that initially scores a customer can be evaluated multiple times during the lifecycle of a customer in terms of transaction profile, product uptake, delinquency, etc.

The challenge with dynamically updating a model is that the need for a model change / evolution is based on a changing or changed operational environment. It becomes apparent that the historical data used for training is outdated. In theory, new data will have been captured which will represent this new operational change. In practice, though, this is not always the case. Fast changing risk environments will suffer from a loss of sufficient data for training to take place. New risk is often interpreted as noise to the AI algorithm. In these data-scarce risk environments, the data scientist might need to consider training data sources that fall outside of traditional historical data. Such data might come from human assessment of the potential risks of AI and the potential negative impacts of dynamic updated models.

In summary, dynamic updating:

- May lead to model drift. Risk may crystallize at any size. The model framework must be sound.
- Might need to evaluate a prediction / classification in terms of decision consequence (cost, capacity, and risk).
- Can have predefined acceptable parameters for dynamic change when the decision consequence is stated in advance.
- Can have a form of manual acceptance.
- Could adopt an ensemble approach where the new model comprises an ensemble of a number of previous models.

Question 16: To the extent not already discussed, please identify any additional uses of AI by financial institutions and any risk management challenges or other factors that may impede adoption and use of AI.

In general, banks are considering the application of AI in the following areas:

- Customer service (using conversational AI to improve customer service and relationship management).
- Marketing (using AI to fine-tune potential product matches with customers).
- Underwriting (using AI to make loan processing, risk assessment, and on-going monitoring more efficient and accurate).
- Risk and regulatory compliance (using AI to monitor rules and regulations by country to improve back office compliance management processes, such as automation of sanctions screening and monitoring transactions for indicia of money laundering).

Risk management challenges or other factors impeding adoption of AI, include:

- Leadership's fear of regulatory action and penalties if the AML results from using the AI application differ meaningfully from using the traditional rules-based transaction monitoring systems. There is a concern that the AI application will not flag transactions similar to the ones flagged by traditional systems. A result of equal concern is the AI's ability to identify potentially suspicious transactions in the past data used to train the algorithm that the bank originally missed. A bank's risk governance models for new technology solutions also prevent the speedy rollout of AI/ML applications.
- Complex, highly variable data sets, such as are present in trade finance documentation, where there is no standard format for commercial invoices and other documentation presented to banks.
- Poor data quality, such as is present in trade finance documentation, where scanned images of paper documents may have low image resolution.

Question 17: To the extent not already discussed, please identify any benefits or risks to financial institutions' customers or prospective customers from the use of AI by those financial institutions. Please provide any suggestions on how to maximize benefits or address any identified risks.

In addition to the above, the use of AI may benefit customers in the following ways:

- Improved customer service through more customized and efficient processing of requests.
- Lower overall fees resulting from lower operating costs and improved efficiency.
- Faster transaction processing resulting from more accurate compliance screening and automation of back office functions.

- Improved communication of the benefits of products that may fit customers' needs.

The use of deficient AI in the areas noted above, of course, could introduce risks or costs to customers, such as decreased efficiency and accuracy of customer services, increased costs of internal operations and ultimately fees to customers, increased processing times due to the introduction of AI, and recommendation of products to customers that are not relevant.

To minimize the above risks, banks must invest in and deploy AI solutions cautiously, ensuring that there is clear customer value compared to the current state before implementation and that controls are in place to continuously monitor risks and costs.

*****

Thank you for your consideration of these comments. If you should have any questions regarding this letter, please do not hesitate to contact me at spelosi@baft.org or 202-317-2173.

Sincerely,

Samantha J. Pelosi
Senior Vice President of Payments & Innovation