



Barrett Burns President & CEO barrettburns@vantagescore.com

June 30, 2021

Comment Intake
Bureau of Consumer Financial Protection
1700 G Street NW,
Washington, DC 20552.

Delivery via email to: 2021-RFI-AI@cfpb.gov.

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

Docket No. CFPB-2021-0004

To Whom it May Concern:

The use of Artificial Intelligence and Machine Learning across a wide range of industries and the related benefits and potential concerns have been receiving significant attention recently. VantageScore Solutions, LLC ("VantageScore") applauds the federal financial regulatory agencies for seeking information to better understand the use of these methodologies by financial institutions and the related governance, risk and control considerations.

VantageScore develops generic consumer credit scoring models that deploy a consistent algorithm across each of the three national credit reporting companies: Equifax, Experian, and TransUnion (the "CRCs"). In choosing to work exclusively with credit file data, we benefit from the stringent regulations (i.e., ECOA and FCRA) and the data standards (e.g., quality, accuracy, standardization, and universality) that such data are subject to.

VantageScore Solutions was founded with the mission of scoring more people responsibly and predictively while at the same time increasing improving performance and inclusion. VantageScore models thus provide a greater opportunity to access sustainable mainstream credit by a broader group of consumers. Our models are used in every part of the credit process.

Our latest two models, VantageScore 3.0 and VantageScore 4.0, are able to generate a score for approximately 96% of those consumers who are 18 or older. As a result, there is a significant reduction in the number of conventionally unscorable consumers (who have been identified as



the “credit invisibles” in the CFPB’s May 2015 research report¹). That group includes approximately 37 million consumers who are otherwise unable to obtain a credit score if a lender is using a legacy credit scoring model. Of particular note is that within these 37 million otherwise unscorable, approximately 13 million achieve a score a 620 or more, and within this group of 13 million, approximately 3 million are African Americans or Hispanics.

To set the appropriate context, we view Artificial Intelligence (AI) as a broad field, generally defined as machines mimicking some form of human behavior. Machine Learning (ML) is a sub-discipline within the field of AI which is characterized by the use of algorithms that are able to mimic how humans learn, thereby improving with experience and new data. Many of the standard tools used for classification, regression, clustering and detection analyses are ML algorithms and have been around for a long time.

More recently, a more complex subset of ML algorithms has been put into broader use, fueled by the increased availability of computing power. These more complex methods, sometimes referred to as “Deep Learning,” include neural networks, natural language processing and image processing. The recent discussions related to AI and Machine Learning tend to center around these more complex methodologies, focusing on issues related to transparency, explainability and the potential “black box” nature of some of these methodologies.

Modern AI and ML techniques have the potential to provide significant benefits in consumer lending by providing increased accuracy in projections and estimates, deeper insights to solve complex problems, as well as significant efficiency gains and transferability of applications. At the same time, consumer lending is a highly-regulated industry with strict requirements regarding fair lending and fair credit reporting related consumer protections, as well as the need for safety and soundness controls related to risk management systems and the models that are used in decisioning. Therefore, any potential application of AI and ML tools must appropriately take into account these expectations and requirements, and weigh the potential the benefits of these tools against potential risks and the principles of transparency, explainability, consistency and fairness. Further, the use of advanced methods cannot be a substitute for domain expertise and sound business judgment. The resulting algorithms, underlying assumptions and results must be carefully reviewed and challenged by subject matter experts to ensure that business, policy, legal and control objectives are met.

At VantageScore, we have been able to successfully incorporate the benefits of using advanced machine learning techniques to enhance credit scoring models. This use brings significant value to both consumers and lending institutions by expanding the population of consumers who are able to receive a credit score while also increasing the predictive accuracy of the scoring models. The development of our latest model, VantageScore 4.0, demonstrates how ML methodologies can be harnessed in a way that supports innovation while maintaining the

¹ See CFPB Office of Research Data Point: *Credit Invisibles* (May 2015), https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf.



transparency, explainability and fairness of the algorithms and remaining compliant with existing regulatory requirements.

In VantageScore 4.0, our data scientists have incorporated ML-based approaches as part of the data exploration and model development processes to augment the development of credit data attributes in consumer segments for which there is limited credit history information available. These consumer segments are particularly difficult to assess using traditionally developed attributes. Traditional credit attributes typically incorporate one or two dimensions from the behavioral credit data that is available and are not sensitive enough to predict accurately the risk of default in these consumer segments. ML methodologies enabled our data scientists to examine the full spectrum of data available in the credit files efficiently and comprehensively in order to gain additional, predictive risk insights. Through ML techniques, tens of thousands of multidimensional combinations of data elements were analyzed as part of the data exploration and model development process to identify effective signals and significantly enhance the accuracy of risk predictions.

Our development approach did not, however, end with simply running ML algorithms to achieve these benefits. Our data scientists identified the highest performing multi-dimensional relationships that came to light using ML, as well as their common elements. They then translated these elements into traditional, structured attributes which could be incorporated into the model development process. This translation was done to ensure that consumers could understand the behavioral causes of the credit score they received and what behavioral changes would assist in improving their credit score. While this attribute conversion process resulted in some loss of predictive performance as compared to an unaltered ML-based model, it ensured full compliance with regulatory requirements and an intuitive, transparent and explainable model.

The benefits of a balanced approach which leverages ML combined with strong domain expertise has become very apparent in the case of VantageScore. We observe that VantageScore 4.0 provides a significant lift in predictive performance, particularly for consumer who have been recently inactive with credit but who have accounts on their credit files. There are approximately 24 million such consumers (out of the 37 million consumers scored by our models but are otherwise unscorable by a legacy scoring model) who have not actively used credit within the last six months. The ML algorithms allow deeper insights about these consumers for incorporation into the model. In return, these deeper insights help these consumers to receive a fair and accurate credit score that is more reflective of their creditworthiness and facilitates their access to mainstream credit products best suited to their needs.

Innovation, either through advancements in modeling methodologies such as ML or via improvements in the data available for the development of credit scoring models and other decisioning tools deployed by financial services, can greatly benefit both consumers and the financial system in general. Those advancements, however, must be used transparently, effectively and with appropriate controls and considerations to avoid potential pitfalls. The



benefits are particularly pronounced in the case of consumers who have been historically underserved by the legacy processes currently in place and who are at a disadvantage due to the limited information available in their credit files. At VantageScore, innovation is a core objective in our overarching mission to develop highly accurate credit scoring models while reliably scoring more consumers so that they can gain access to mainstream credit.

Sincerely,

A handwritten signature in black ink, reading 'Barrett Burns', with a long horizontal flourish extending to the right.