FÉDÉRATION
BANCAIRE
FRANÇAISE

10th February 2022

FBF RESPONSE
TO THE
EBA DISCUSSION PAPER
ON
MACHINE LEARNING FOR IRB MODELS

Ref.: EBA/DP/2021/04

The French Banking Federation (FBF) represents the interests of the banking industry in France. Its membership is composed of all credit institutions authorised as banks and doing business in France, i.e. more than 390 commercial, cooperative and mutual banks. FBF member banks have more than 38,000 permanent branches in France. They employ 370,000 people in France and around the world, and service 48 million customers.

* * *

I. GENERAL COMMENT

The FBF welcomes the opportunity to respond to the EBA Discussion Paper on the use of machine learning in the context of internal ratings-based (IRB) models to calculate regulatory capital for credit risk.

The key ideas developed in this document refer to:

- challenges about interpretability and implementation,
- uncertainty about the acceptability from the regulator (expected constraints associated with CRR3),
- more generally, a welcome clarification of the roles and responsibilities of the various industry and European regulatory bodies in the regulation of Machine Learning models (it is not always clear how the various initiatives, for instance the current one from EBA, some from ACPR and the AI Act for the European commission, are articulated together).

II. COMMENTS ON THE CONSULTATION

Question 1:
Do you currently use or plan to use ML models in the context of IRB in your institution? If yes, please specify and answer questions 1.1, 1.2, 1.3. 1.4; if no, are there specific reasons not to use ML models? Please specify (e.g. too costly, interpretability concerns, certain regulatory requirements, etc.)

NB: indications as to institutions practices here refer to the main/broad trends gathered from FBF members. Sometimes the findings are specific to some members or some business units, in which case this is specified.

Globally speaking, FBF members do not currently use sophisticated ML models in the context of IRB. Naturally, some use simpler models (e.g., logistic regression) for their IRB models. The reasons for not relying on more sophisticated ML are (i) challenges about interpretability and implementation and (ii) uncertainty about the acceptability from the regulator. In a specific case for example, a member is considering using a random forest technique for a cure score for which logistic regression and decision tree algorithm performances are poor.

Question 1.1:
For the estimation of which parameters does your institution currently use or plan to use ML models, i.e. PD, LGD, ELBE, EAD, CCF?

Some FBF members use simple models for both PD and LGD. The random forest model mentioned in the previous question would be aimed at LGD estimation. Note that in the remainder of this questionnaire we focus solely on sophisticated ML models as it seems to be the focus of the discussion, leaving logistic regression and linear regression aside.

Question 1.2:
Can you specify for which specific purposes these ML models are used or planned to be used? Please specify at which stage of the estimation process they are used, i.e. data preparation, risk differentiation, risk quantification, validation.

FBF members' ML models are mainly used for feature selection and risk differentiation. They also rely on ML models as challenger models.

Question 1.3:
Please also specify the type of ML models and algorithms (e.g. random forest, k-nearest neighbours, etc.) you currently use or plan to use in the IRB context?

Currently FBF members' main focus is on leveraging random forest algorithm.

Question 1.4:
Are you using or planning to use unstructured data for these ML models? If yes, please specify what kind of data or type of data sources you use or are planning to use. How do you ensure an adequate data quality?

Not at this stage.

Question 2:

Have you outsourced or are you planning to outsource the development and implementation of the ML models and, if yes, for which modelling phase? What are the main challenges you face in this regard?

The majority of our members / their business units have not outsourced and no plans to do so in the future.

Question 3:

Do you see or expect any challenges regarding the internal user acceptance of ML models (e.g. by credit officers responsible for credit approval)? What are the measures taken to ensure good knowledge of the ML models by their users (e.g. staff training, adapting required documentation to these new models)?

The answer to this question varies based on the maturity of the business unit in terms of ML usage outside of IRB models. Some business units indeed have concerns that business representatives and credit officers are comfortable with the understanding and interactions associated with the standard approaches and would not necessarily welcome ML models as they are less familiar with the way to interpret them. There can be also challenges associated with the complexity of input data oftentimes more complex and numerous than in models with standard approaches. Other business units mentioned having no particular concern mentioning that such techniques are already used in the granting process in which specific sessions for understanding model behaviour are organized when ML models are implemented.

Regarding the question about training of staff and framing of the development of AI Models, there are numerous initiatives within some of our members, including specific trainings on Artificial Intelligence, Machine Learning but also interpretability and bias in AI. Moreover, for example, in one case a governance around ML models has been defined at group level and is progressively being implemented to accompany the transition of all modelling teams interested in leveraging machine learning (not focusing specifically on IRB models). In particular, recently, the second version of AI Model Risk Management guidelines (targeting LOD1), stipulating a framework for documentation, has been distributed throughout this group. In addition, a Generic Control Plan for Model Risk has also been adapted to include the specificities of ML.

Question 4:

If you use or plan to use ML models in the context of IRB, can you please describe if and where (i.e. in which phase of the estimation process, e.g. development, application or both) human intervention is allowed and how it depends on the specific use of the ML model?

The question is unclear and does not seem to be specific to ML.

Question 5:

Do you see any issues in the interaction between data retention requirements of GDPR and the CRR requirements on the length of the historical observation period?

There is undeniably some lack of clarity in the interplay between constraints from GDPR and CRR requirements. While some members/business units report that CRR has been accepted as a valid reason to keep record necessary for the validation of IRB models, others mentioned that it can sometimes be difficult to justify the need to store 20 years of observation without full anonymization (the person ID is required for the audit trail).

Question 6:
Do you have any experience in ML models used for estimating credit risk (if possible, please differentiate between models where ML is used only for risk differentiation, only for risk quantification or used for both)? If so, what are the main challenges you face especially in the areas of:
a) Methodology (e.g. which tests to use/validation activities to perform).
b) Traceability (e.g. how to identify the root cause for an identified issue).
c) Knowledge needed by the validation function (e.g. specialised training sessions on ML techniques by an independent party).
d) Resources needed to perform the validation (e.g. more time needed for validation)?

There is some variability in the experience of members/business units. Some are quite experienced at using ML models for risk granting, others have used them as feature selectors, i.e., to select risk drivers, others have looked at advanced methods for acquisition scores.

In terms of challenges, one member reported that ML is not used in an “automated re-calibration mode”, in which the model would be constantly retrained as more data becomes available. Therefore, the difference between ML & standard approaches lies only in the statistical learning methodology (based on random forest or gradient boosting) applied. The main challenge resulting from this change lies in explainability – see question 15. On another perimeter, it was highlighted that one model that leverages pseudo social network to build features for scoring has been both an implementation and validation challenge.

Question 7:
Can you please elaborate on your strategy to overcome the overfitting issues related to ML models (e.g. cross-validation, regularisation)?

Overfitting is not perceived as a prevalent issue in the current context of limited usage of ML models and the focus on simpler models. However, whenever banks are faced with the issue, the common reaction is to look at feature selection, cross-validation and regularisation.

Question 8:
What are the specific challenges you see regarding the development, maintenance and control of ML models in the IRB context, e.g., when verifying the correct implementation of internal rating and risk parameters in IT systems, when monitoring the correct functioning of the models or when integrating control models for identifying possible incidences?

One of the challenges that FBF members feel is key is the reproducibility of the model between the development platform and the production platform. In addition, they note that the verification that a model is properly implemented in production is not straightforward.

Question 9:

How often do you plan to update your ML models (e.g., by re estimating parameters of the model and/or its hyperparameters) Please explain any related challenges with particular reference to those related to ensuring compliance with Regulation (EU) No 529/2014 (i.e. materiality assessment of IRB model changes).

Model update strategies will depend on the use case and the reason triggering the update.

Question 10:

Are you using or planning to use ML for credit risk apart from regulatory capital purposes? Please specify (i.e. loan origination, loan acquisition, provisioning, ICAAP).

ML is being used or being planned for credit scoring and operational risk monitoring at loan origination and also for credit monitoring / early warning systems.

Question 11:

Do you see any challenges in using ML in the context of IRB models stemming from the AI act?

The question is very broad and the AI Act being at the same time very prescriptive and ambiguous, we cannot at this point answer this question.

Question 12:

Do you see any additional challenge or issue that is relevant for discussion related to the use of ML models in the IRB context?

On capital, expected constraints associated with CRR3 will limit the ability of ML models to offer a gain in performance compared to simpler approaches.

Question 13:

Are you using or planning to use ML for collateral valuation? Please specify.

FBF members report that nothing is planned at this stage, but that ML could be a useful tool to estimate the value of a real estate asset.

Question 14:

Do you see any other area where the use of ML models might be beneficial?

The question being fairly open ended, we focus more specifically on credit risk aspects. As mentioned before, ML can be a powerful tool in both early warning systems to accompany credit monitoring and also for operational risk monitoring, enabling a filtering of the alerts to reduce false positive and escalate important problems.

Question 15:

What does your institution do to ensure explainability of the ML models, i.e. the use of ex post tools to describe the contribution of individual variables or the introduction of constraints in the algorithm to reduce complexity?

Explainability is essential within ML modelers community as it is required to ensure business understanding and validation of the models being developed. Banks focus on both global explainability aspects (typically for Model Owners) and local explainability aspects (typically for Model Users). The most commonly used techniques include Shap values, feature importance, partial dependence plots (PDP) and accumulated local effects (ALE). Some have also experimented with sophisticated algorithms that are explainable by design such as AGBoost.

To be noted however, the ability to explain is more acute when the results are not aligned with the expectations of the user, i.e., for exceptions, problems rather than for all possible cases.

Question 16:
Are you concerned about how to share the information gathered on the interpretability with the different stakeholders (e.g. senior management)? What approaches do you think could be useful to address these issues?

As a general consideration, banks' senior managements are proactive in ensuring proper model risk management of ML models. In that context, they are also interested that relevant aspects of AI model development deployment and consumption (including interpretability) are addressed.
Having said this, training (CRCU, Senior management,..) can represent a challenge given the complexity of some ML models, leading to the need to set up specific training sessions.

Question 17:
Do you have any concern related to the principle-based recommendations?

FBF members welcome principle-based recommendations for risk management that take into account the materiality of the risk and that foster pragmatic discussion with the supervisors.