Prot. 2022/21

POSITION PAPER
"EBA DISCUSSION PAPER ON MACHINE LEARNING FOR IRB MODELS"
(EBA/DP/2021/04, 11 November 2021)

Assilea, the Italian Leasing Association, represents the Italian leasing industry. Our members are leasing companies classed into generalist banks, specialist banks, non-banking financial intermediaries, brokers and dealers, long-term rental companies, outsourcers specialized in the leasing market. The Association's key task is to carry out institutional activities with a view to providing information and assistance to its members and contributing towards the solution of leasing-related issues at different levels, in different domestic and international venues. Assilea is member of Leaseurope (the European Leasing Federation) and of ABI (the Italian Banking Association).

We thank you for the opportunity offered with this consultation. In this paper we reply to some of the main questions of the EBA/DP/2021/04, 11 November 2021. As we are an Association, our reply is based on the experience we have in creating tools from our databases that could be used by our members in the process of asset value assessment and in the credit risk evaluation and monitoring.

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.).

Our Association (the Italian Leasing Association), in partnership with academic experts and data manager advisory companies, has recently developed machine learning models in the field of real estate evaluation assessment and monitoring and in the field of leasing credit worthiness assessment and monitoring (i.e. *early warning mechanism*). They could be integrated in IRB models.

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?

The first of the above-mentioned tools could be used in the calculation of LGD parameters and for credit risk mitigation purposes in the collateral's evaluation. The second one could be part of the PD estimation process (one-year time horizon PD): different rating classes of the client in the specific transaction are provided according to the credit worthiness of the client and the characteristics of the transaction (based on private and public data bases). Of course, both the tools should be integrated into the specific leasing companies' evaluation systems, in terms of company's data set and models, according to the leasing company risk appetite and strategy.

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.

According to the structure of the specific leasing company, these tools could be used in different stages of the estimation process.

assilea

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?

In the real estate evaluation model, the machine learning technique which has been used is the neural network one.

In the credit worthiness model, the machine learning model which has been used is that of an ensemble of gradient boosting trees.

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?

In the real estate evaluation machine learning model, some destructed data are used for the location of the asset and the prediction of price evolution in the area (e.g., geolocation data, economic information from various sources). In the credit worthiness machine learning model, there was a first attempt to use open data of a regional chamber of commerce (data about the evolution in companies' demography by sector), but with no additional value to the model. We do not exclude the use of unstructured data for the future implementation of leasing machine learning models.

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 development of ML model could be outsourced in most of the modelling and implementation phase, the choice is based on the data availability and expertise of the lease company and, in our opinion, is proportional to the organization complexity, in terms of dimension and activity diversification.

As risk differentiation is concerned, regarding the following paragraph on page 11 "Risk differentiation [...] through text mining", we think that text mining is not properly a technique used/suitable for revising PD of other models. We suggest removing "through text mining" or to replace e.g., with "through calibration algorithms".

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)?

As an association, we try to expand the culture about the evolution in credit risk and in credit risk management. We also offer statistics, the access to common data bases, and information services to improve the use of the available data sets. Machine Learning models are the natural evolution of all of that. We organize specific events, meetings, training courses for the leasing companies and specific sessions with the credit officers responsible for credit approval.

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?

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?
The recent Opinion 11/2021 of the European Data Protection Supervisor[1] may be related to this question. It gives specific recommendation about the use of consumer credit data in the new context of digitalization and in the interaction with Artificial Intelligence Act. Therefore, in our opinion, to this question a sufficiently general answer should be provided after the consultation phase.
Concerning other specific issues raised in your discussion paper, we add the following remarks.
Page 13 "Finally, the possible [...] requirements only", you might want to stress, however, that technical solutions for different requirements are not orthogonal, but interdependent, e.g., for the fairness requirement there are specific approaches as reported in Nikita Kozodoi, Johannes Jacob, Stefan Lessmann: *Fairness in credit scoring: Assessment, implementation, and profit implications* [Eur. J.Oper. Res. 297(3): 1083-1094 (2022)] and these approaches may conflict with other requirements, e.g., as in the accuracy-fairness trade-off.
Page 14 "The definition and assignment [...] complex ML models". Finding the required clear economic theory behind the models might be impossible. It appears more realistic that economic theories/constraints are integrated in the logic of the model, e.g., as constraints or domain knowledge.

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)?**

The validation activities include a first dataset training and a following dataset test. In the initial phase of the project, through questionnaires and interviews, the validation metrics have been set. They are manifold and cover different issues (ranking of the client credit worthiness based on the score, prediction errors, errors calibration as a probability, misclassification errors in the rating notch).

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

The overfitting problems of the model have been tackled with regularization techniques L1 and L2 and with ensembling methods.

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

[1] https://edps.europa.eu/system/files/2021-08/opinion_consumercredit-final_en.pdf

functioning of the models or when integrating control models for identifying possible incidences?

The performance of the ML model for leasing credit worthiness has been evaluated on the whole database of the leasing credit bureau and on the individual leasing company's data; it has been evaluated on clusters of the leasing credit bureau data in which the model performs particularly well, on clusters in which the model performs in line with the average level, on clusters where the model is underperforming. In this way it is possible to outline those features (or missing values) that impact the models results.

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).

ML parameters are recalculated on a quarterly basis. Hyperparameters are re estimated in relation to the monitoring assessment, e.g. they are re estimated either when ML performances degrade or to implement forward-looking policies based on the macroeconomic conditions.

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).

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

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?

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

As mentioned in the reply to question 1, our association developed a ML model for real estate assets evaluation and price monitoring.

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

ML models may provide uncertainty estimates of their predictions, e.g., as in conformal prediction (https://en.wikipedia.org/wiki/Conformal_prediction).
These can be exploited for improved risk differentiation/estimation, and for model validation.

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?

Projection reproducibility is guaranteed by keeping a copy of the models and keeping track of the software libraries that are used. Projection explainability is guaranteed through feature relevance technics (Shapley values) and counterfactual explanation that have been thought to support human decision in final the credit worthiness assessment.

As far as “The technical box on interpretability techniques” on page 22 of the consultation document is concerned, you might find more general and complete categorizations of these techniques in survey papers such as:
Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, Dino Pedreschi: A Survey of Methods for ExplainingBlack Box Models. ACM Comput. Surv. 51(5): 93:1-93:42 (2019)
Michael Bücker, Gero Szepannek, Alicja Gosiewska, Przemyslaw Biecek: Transparency, Auditability and eXplainability of Machine LearningModels in Credit Scoring. CoRR abs/2009.13384 (2020)

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?

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