
Response to Discussion Paper

Machine Learning for IRB Models

10 February 2022

The Association for Financial Markets in Europe (AFME) welcomes the opportunity to comment on the European Banking Authority's **DISCUSSION PAPER ON MACHINE LEARNING FOR IRB MODELS**. AFME represents a broad array of European and global participants in the wholesale financial markets. Its members comprise pan-EU and global banks as well as key regional banks, brokers, law firms, investors and other financial market participants. We advocate stable, competitive, sustainable European financial markets that support economic growth and benefit society.

AFME is the European member of the Global Financial Markets Association (GFMA) a global alliance with the Securities Industry and Financial Markets Association (SIFMA) in the US, and the Asia Securities Industry and Financial Markets Association (ASIFMA) in Asia.

AFME is registered on the EU Transparency Register, registration number 65110063986-76.

We have provided our response to the consultation questions, as well as an appendix which provides further details on our views.

Executive Summary

AFME supports the European Banking Authority's ('EBA') objective of understanding the challenges and opportunities coming from the world of machine learning (ML) should they be applied in the context of internal ratings-based (IRB) models to calculate regulatory capital for credit risk.

ML offers new opportunities to improve the accuracy of risk models, surpassing traditional models in terms of predictive power. The potential benefits for institutions would be significant¹. The financial sector should be able to take advantage of these opportunities. However, as the Discussion Paper reflects, one of the main challenges is the uncertainty about supervisory expectations. Therefore, we encourage the EBA to provide the necessary supervisory certainty for banks in this area. Within the IRB context, financial institutions do not have sufficient clarity on supervisory expectations about the interpretability of the results, which could pose a hurdle to the adoption of this technology. We believe that there is a need for the EBA to provide clarity when ML explainability techniques sufficiently meet Capital Requirements Regulation (CRR) requirements. There is also a need for the EBA to provide some clarity on the significance of business interpretability of the ML techniques. For example, would a data-driven model that is complex but explainable in quantitative terms qualify as a champion model (the winning model) despite lacking full business interpretability? ML models are not always straightforward and there could be a challenge with interpretability. ML techniques do not

¹ See Bank of Spain: UNDERSTANDING THE PERFORMANCE OF MACHINE LEARNING MODELS TO PREDICT CREDIT DEFAULT: A NOVEL APPROACH FOR SUPERVISORY EVALUATION [<https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/21/Files/dt2105e.pdf>]: According to the paper, ML models have better predictive power both in terms of classification and calibration than Logit or Lasso, regardless of the sample size and number of available features. Our benchmark results show that implementing XGBoost could yield savings from 12.4% to 17% in terms of regulatory capital requirements under the IRB approach. This leads us to conclude that the potential benefits in economic terms for the institutions would be significant

Association for Financial Markets in Europe

London Office: 39th Floor, 25 Canada Square, London E14 5LQ, United Kingdom T: +44 (0)20 3828 2700

Brussels Office: Rue de la Loi 82, 1040 Brussels, Belgium T: +32 (0)2 788 3971

Frankfurt Office: Bürohaus an der Alten Oper, Neue Mainzer Straße 75, 60311 Frankfurt am Main, Germany
T: +49 (0)69 153 258 967

www.afme.eu

allow for a direct and intuitive linkage between the model inputs and model outputs. ML models are often seen as a “black box” supported by statistical logic, but not necessarily intuitive business logic.

In addition to the principles-based approach proposed in this Discussion Paper, we wish to highlight that the Paper may not have sufficiently addressed the following:

- How the risk-based approach applies when using ML in certain steps of the development or when feeding ML into traditional methods;
- The importance on data selection which excludes biases within the model solution;
- When unfair bias has been properly identified, addressed and eliminated;
- The challenge of applying a "standard practice" for the model development. For example, there is no common approach to ML methodologies. High permutations of starting parameters can give similar results, where no consistent approach can give an optimal solution;
- How AI/ML incremental or continuous learning (model changes) can be viewed in the context of the regulations on prior notification or approval of model changes (EU 529/2014)
- Supervisory timing for issuing an opinion regarding these new approaches.

Consultation Questions:

Q1: 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.)

Many of our Member firms do not use Machine Learning (ML) directly for IRB regulatory RWA calculations.

For those firms that use Machine Learning in an IRB context, examples of use cases provided include, applying a Machine Learning based component in an IRBA-approved rating system for corporate counterparties and the use of a Machine Learning-based sub-model to assess the textual parts of annual reports within the rating methodology applied to Large Corporate clients.

Furthermore ML models are often used for variable selection and for benchmarking purposes (both by development and validation functions).

A few firms also use ML-based algorithms to support or validate certain steps in the model development phase within the IRB context. As an example, an organization used a model based on a convolution neural network (CNN) as a challenger model to a recently-developed local retail unsecured probability of default (PD) IRB model.

The main reason cited by Member firms for not using ML-models for IRB purposes extensively, primarily lies in the difficulty to demonstrate adherence with strict, applicable regulatory requirements. For instance, where model re-development plans have already been agreed with competent authorities and the introduction of new ML-models may not be compatible with existing plans. Supervisors also often require humans in the loop when applying the models in production, which limits leveraging the full potential of ML e.g., with respect to automation.

As outlined in Sections 3.1 and 4.1 of the discussion paper, there is uncertainty on how models could comply with existing regulations. In the absence of certainty that ML would be accepted, it is a questionable option for resource commitment. Without clear criteria for acceptance or precedents, it would be unlikely to be a feasible option. For example, it is unclear if ML would satisfy all the criteria set out in specific EU regulatory requirements such as the EBA Guidelines etc

A more limited constraint in using ML in the context of IRB, is data availability. Within Wholesale credit, there are few default observations, making it challenging to apply ML in this context. Although noteworthy that data availability might be a problem in some context, this challenge is not specific to the use of ML and should not forestall the opportunities that exist in using ML within the IRB context.

We would like to emphasize that the reasons for not calculating parameters using ML are current realities which could change in the future.

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

ML-based algorithms are used in the estimation of the parameters for the probability of default (PD), the loss given default (LGD), the exposure at default (EAD) and the credit conversion factor CCF.

Specifically, some instances of use of ML algorithms include:

- Risk differentiation
- Decision on number of grades.

The use of ML is still very limited as explained in Q1. Member firms mainly use or would use ML for PD (ratings). However, ML offers opportunities to improve the estimation of any of these parameters. We do believe that there are opportunities in using ML for:

- (1) Incorporating alternative information sources like textual data
- (2) Risk driver selection;
- (3) Imputation of missing variables;
- (4) Simulation of different combinations to select the most suitable tree;
- (5) Data quality control;
- (6) Information processing;
- (7) Challenger models (performance challenge of the algorithms of traditional techniques).

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.

ML models are currently used for:

	Development	Validation	Monitoring
Data preparation	No, but see note in 1.3 below	-	-
Risk differentiation	Benchmarking Model component	Benchmarking, challenge of variables used and selected by experts	-
Risk Quantification	-	-	-
Validation	Benchmarking	Benchmarking, challenge of variables used and selected by experts	-

Ideally if there was some regulatory assurance, firms would consider the use of ML-based algorithms for the following purposes:

- dimensionality reduction;
- interaction detection – ML based-algorithms (e.g. decision trees) are used to understand features and the differentiation between data points to assist with explainability;
- feature generation –ML is used to create transformations of raw data;
- residual modelling – upon completion of model selection, ML is used to conclude whether any residual structured information remains in the data to further improve model performance;
- issuing data imputation;
- text classification - (re)classification of issues and findings raised on models and rating systems in Risk;

We welcome some clarity from the Regulators on the requirements or regulatory constraints(if any) that may apply in the afore-mentioned scenarios.

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?

The use of ML models in the IRB context whilst still developing, has been applied in the context of model benchmarking the last few years in order to challenge the expert-driven champion models.

Examples of ML models and algorithms used include: Gradient Boosting Machine; Random Forest; Decision Tree; One-layer Near Neighbour (NN); Unsupervised Clustering (KNN); Natural Language Processing with the Latent Dirichlet Allocation.

Other types of ML algorithms under consideration for use include:

- random forest;
- stochastic gradient boosted Trees or Decision Trees from;
- (linear) support Vector Machine;
- probabilistic neural networks;

This is not a definitive list of ML algorithms that would be useful in the future, and the uses will evolve over time.

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?

Some of our Member firms do not currently use unstructured/semi-structured data for development in IRB ML models. However, one institution reported the use of textual data from annual reports in a ML-based component of an IRBA-approved rating system for corporate counterparties.

Some of our Members acknowledge the potential of a variety of data sources that can be used in ML to explain customers' behaviours. Examples include customer comments in collections, email addresses, text files (e.g. comments, email messages), audio, video, images, financial reporting and log files. For this reason, unstructured data should be considered as an option as it may provide value to the model.

ML can also be used to classify comments into categories in Validation.

With regards to ensuring adequate data quality, unstructured data can be cleaned to adhere to controls in such a way that the resulting data meets the quality and completeness requirements of existing regulatory requirements in IRB. These controls are part of the monitoring process of the models and they are performed periodically to mitigate any possible risks.

To ensure data quality, some firms apply minimum standards that constitute a set of data quality tests covering various dimensions irrespective of the model technique/approach and can be applied across various model categories (e.g. provision or for capital modelling purposes).

In order to mitigate the risk that ML-based approaches may introduce or recommend any decision in violation of ethical principles, some firms introduced a Model Ethics framework.

Q2: 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?

More generally, firms tend to use externally developed ML-based algorithms, particularly in cases when detailed data are accessible only via third party or there are no specific appetite or resources available to develop the models internally.

The validation of ML-based models/algorithms is a core function predominantly performed internally and not often outsourced, but it may be expedient to outsource in limited circumstances, for example where technical niche expertise is required or a shortage of resources becomes an impediment to timely outcomes.

Q3: Do you see or expect any challenges regarding the internal user acceptance of ML models (e.g. by knowledge of the ML models by their users (e.g. staff training, adapting required documentation to these new models)?

In general terms, development and validation processes do not change substantially due to the use of ML. Some of the techniques and tests might be different but they will be based on the same fundamental principles as for other traditional techniques.

Firms acknowledge that the internal users' understanding and acceptance of the model and its functioning are key elements to ensure the model delivers on its purpose of supporting business decisions. This requires significant efforts for example in explaining ML models when compared to traditional models. ML-based approaches could pose an additional challenge for interpretation from a business user perspective, particularly when it relates to gaining comfort over the reasonableness of the model outputs, particularly for wholesale banking. It is noteworthy that this challenge exists in every use case of ML and is not peculiar to ML models within IRB. However, partial explainability of the relationship or outcomes of the ML models is generally acceptable in other user cases of ML, but unclear if this would also be acceptable in the IRB context.

One key issue for user acceptance would be the explainability, which is not a clearly defined term; especially in cases where the model view diverges from the expert view, it would be crucial for the experts to understand the methodology of the model and, in some cases, the specific factors/reasons for the model outputs.

In spite of the challenges highlighted above, there are techniques that can be used to explain and describe the results in a similar way as we do with traditional models (e.g. by describing the relevant factors involved in the decision). To address the challenge of internal user acceptance, some firms establish model risk management framework with controls to ensure that internal user acceptance is built-in the selection of the approach. Typically a simple more intuitive approach may be preferred over a more complex model to retain a deep understanding over the model usage (e.g. use of loss rates per LTV buckets over a model).

Generally, the understanding of the model output and its functioning is achieved by segregation of duties across multiple control functions, depending on the role and the model use. As with the usage of traditional models, it is key to ensure that 1st and 3rd lines of defence have the skills to understand the results of the models. Minimum standards for model use require user training, clear and current instructions to model usage and pre-authorization for additional uses. In essence, the effectiveness of the controls from end-to-end is achieved via reliance on integrated controls and by compounding the effectiveness of specific reviews along the model lifecycle steps.

Typically, a detailed understanding of the technical aspects is predominantly expected from the validation functions which would perform dedicated tests for specific uses. In addition, the model output and its reasonableness and responsiveness to inputs is evaluated by panels of experts consulted during the development phase and prior to completing the internal approval process. The implementation design also progresses starting as early as the development phase and it is approved by the model owner (or delegates) prior to receiving internal approval.

Overall, we do not consider that the introduction of ML models will increase any risk associated to models' results interpretation. With regards to the development and validation processes, these processes do not change substantially.

ML is used in many areas apart from IRB models such as in AML/CFT and fraud prevention. When ML is used in those contexts, the explainability techniques that are already available are proving to be effective to understand the outcomes. However, within the IRB context, financial institutions do not have sufficient clarity on supervisory expectations about the interpretability of the results, which becomes a barrier to the adoption of this technology. We believe that there is a need for the EBA to develop guidance on this issue and clarifies what level of explainability meets CRR requirements.

Q4: 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?

ML models in the context of IRB are often used in the development process. Some firms use ML-based algorithms only to support or validate certain steps in the model development phase within the IRB context, human intervention is allowed and required during the model development phase to evaluate any proposed model output or instances where experts may simply disagree with the model output. In application, output of ML models is reviewed and approved by credit officers for further consideration in IRB models.

ML has also been utilised in the context for IRB for low value-added tasks such as the preparation of information sources or the selection of variables to make the process more efficient and eliminate the subjectivity of these processes, as every decision is statistically supported.

Specific examples include:

- challenging variable selection based on deep knowledge of the business, professional experience and/or economic intuition
- visual inspection of the proposed model results, e.g. definition of number of grades obtained from the clustering algorithm or verifying possible overtime intersections of the target variable across grades
- bucketing of grades/pools
- evaluating performance and fitness-for-purpose of challenger model

The use of ML often allows for human intervention in the development and implementation of the algorithms. Tasks such as the algorithm selection, the establishment of the boundary conditions, and parameterization of the algorithms (number of nodes, clustering...) are performed by human intervention. Although there are also opportunities to automate the end-to-end process and this offers the advantages of producing simpler solutions, faster creation of those solutions, and models that often outperform hand-designed models.

Q5: 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?

We do not foresee the use of ML posing any additional challenge to comply with data protection regulation. However, some firms face challenges not to compromise any of the following requirements based on legal uncertainties:

CRR and GDPR

- CRR requires firms to collect and store all relevant data to provide effective support to its credit risk measurement and management process. ([CRR Article 144\(1\) \(d\)](#)). This in practice also includes personal data.
- CRR specifies that this dataset need to be retained for at least for five years. If the available observation period spans a longer period for any source, and this data is relevant, the longer period shall be used. ([CRR Article 180 1. \(h\)](#)) . Whereas, GDPR does not specify retention periods for personal data but stipulates that a firm must ensure that personal data is stored for no longer than necessary for the purposes for which it was collected. That period should take into account the reasons why the firm needs to process the data as well as any legal obligations to keep the data for a fixed period of time. CRR also requires traceability of the dataset: The history; processing and location of the data under consideration can be easily traced ([CRR Article 174\(b\)](#)). With GDPR, If data is anonymized, the GDPR allows you to keep it for as long as you want. By way of an exception, personal data may be kept for a longer period for archiving purposes in the

public interest or for reasons of scientific or historical research, provided that appropriate technical and organization measures are put in place (such as anonymization, encryption etc.)

Q6: 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:

ML-based models have been used by some firms for the estimation of both risk differentiation and quantification. There are no specific challenges identified in the aforementioned areas.

a) Methodology (e.g. which tests to use/validation activities to perform)

- Although many firms are not currently using ML in credit risk models, we do not think that the use of ML requires a different methodology or specific steps in the development process. The steps are the same (selection of LR, selection of variables, selection of drivers, construction of decision trees etc.).

b) Traceability (e.g. how to identify the root cause for an identified issue)

- As noted in 3 above, explainability of models is key. Explainability is in turn linked to Traceability, as models would need to be transparent to internal model users and developers. Traceability (e.g. how to identify the root cause for an identified issue) is a part of the validation activities. Gaining a deep understanding of how hyperparameter settings impact a model and detect potential issues (e.g. for risk drivers selection, assessing statistical dependencies, testing assumptions, evaluating human judgement, etc.) is an iterative process and typically entails replicating model output by re-running scripts (or part thereof) and challenging the model design.

There are available techniques that can be used to explain and describe the results in a similar way as is done with traditional models, but legal certainty about the sufficiency of these explanations to meet supervisory expectations will be helpful.

c) Knowledge needed by the validation function (e.g. specialised training sessions on ML techniques by an independent party)

- We do not consider knowledge required by the validation function constitutes a challenge in that firms can address the continuous need for professional development by offering in-house training on ML where required. In addition, validation teams are structured to allow mobility across projects and countries, thereby offering the opportunity to improve validators' experience with ML approaches and develop an appreciation for pros/cons of various methods.

d) Resources needed to perform the validation (e.g. more time needed for validation)

- Subject to the adopted approach, validation of candidate models (or live models) may require significantly more tests and methodologies than current approaches. (E.g. linear regression tests are simple, standardised and well understood. A Neural network needs additional and more complex validation)
- As highlighted above, resources needed to perform the validation do not necessarily represent a challenge any more than non-ML models. For non-ML models factors that are key to the successful completion of a validation activity are the planning and pre-agreed validation procedures. Additional time would be allocated if required to build a model or test the justification to use a ML algorithm. Nonetheless, ML-based models do require resources with a specialist knowledge of ML model statistical properties and programming skills.

Q7: Can you please elaborate on your strategy to overcome the overfitting issues related to ML models (e.g. cross-validation, regularisation)?

Some Member firms have indicated that models would be subject to all known techniques available to avoid overfitting, such as out-of-sample performance, cross-validation, k-fold testing, regularisation, hyper parameter tuning etc. Other Members have explored common approaches such as splitting data into training/testing, cross-validation, limiting the number of features (by their importance), regularization and increasing the size of training data.

Cross validation and regularisation remain viable tests, in addition to limiting the number of features by their importance and testing model outputs by increasing the size of training data.

There are also techniques available to verify that the relationship of “individual attributes to the target” remains stable over time. These continue to evolve as further work is done in ML.

Additionally, with regards to transparency and explainability, the strategy could be for Credit professionals to check and validate the risk factors, impacts, economic links, etc to ensure that the model follows common business sense and is not overfitted.

We wish to reiterate that monitoring and control processes for ML-based models are similar to those for models based on other techniques. The building steps do not change.

Q8: 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?

Some of the challenges regarding the development, maintenance and control of ML models in the IRB context may include:

- IT implementation, controls and governance . ML processes, cleans and monitors large amount of data so it requires different controls ,requires more capability and IT expertise than would be required for traditional methods e.g. logistical regression In this regard, legacy IT infrastructure is very important and could be a challenge. Some existing IT infrastructure may not fully support the implementation of ML models in credit risk (e.g. number of parameters, non-linear relationships between risk drivers and target variable, unstructured data used in ML models, computation capability). However this challenge is not peculiar to ML models, technological problems are intrinsic to any new model implementation.
- There may also be a challenge with interpretability, when poor performance is identified (e.g. high override rates); it may be more difficult in such an instance, to attribute this or to trace the key reason for the poor performance if the model is essentially a black box.

- It is noteworthy that the challenges highlighted above, are not insurmountable.

Q9: 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)?

The common use of ML-based algorithms in the IRB context is to support or validate certain steps in the model development phase. Model-related changes would occur when a material and consistent deterioration of the model performance is observed. Examples include deterioration in discriminatory power or loss of accuracy.

Similarly, changes to the non-model components of a rating system (e.g. implementation platform) are driven by business requirements when efficiency gains or risk mitigation opportunities justify the change.

Depending on the use of the IRB models, some firms undertake a full review every 3 years, which includes a reassessment of the model design and changes in parameters.

Changes to IRB models are evaluated against regulatory requirements and are governed by internal policies. Given the current regulations, it does not seem to be feasible to have models that are updated more frequently than every few years. Even at an optimistic interpretation, re-estimating parameters in a model may qualify as “changes to the rating criteria and/or their weights or hierarchy” which would in turn trigger a pre-notification or pre-approval).

Challenger models for instance, should be developed at a certain interval and then will be compared with a champion model; whenever the challenger model is significantly better than champion. Ideally, the challenger should ideally become the champion. However, this is challenging under existing guidelines.

It is necessary to differentiate between scenarios where ML models are used exclusively to design the model, but then the model remains static afterwards from scenarios where the training is allowed for ML models. We welcome a clarification from the Regulators, if either of the afore-mentioned scenario is considered more acceptable.

Q10 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)?

The usage of ML models are totally effective for calculating provisions. Estimations of expected losses and regulatory capital should be equivalent due to shortfall, and should be calculated using the same methodology. A practical example will be pricing models that use both capital and provisions costs what means that they have to be calculated by using the same methodology.

Apart from regulatory capital purposes, some firms use ML to develop credit acceptance models, affordability models (income and expenses estimation to assess customers’ repayment capacity), detect application frauds, generate early warning signals and create new variables for risk models (e.g. based on transactional data). In addition. Some firms are also investigating opportunities to develop challenger models of credit decision making models using ML. Other firms use ML for acquisition and collection scorecards. We see many opportunities in the use of ML within the credit risk area and beyond capital estimation, e.g. in the admission process.

Q11. Do you see any challenges in using ML in the context of IRB models stemming from the AI act?

We foresee certain challenges in using ML in the context of the IRB models stemming from the proposed AI Act. For example, Paragraph 80 of the draft AI Act states that “limited derogations should also be envisaged in relation to the quality management system of providers and the monitoring obligation placed on users of high-

risk AI systems to the extent that these apply to credit institutions regulated by Directive 2013/36/EU (i.e. defined in Article 3(1)(1) of that directive and “credit institution as defined in point (1) of Article 4(1) of Regulation (EU) No 575/2013” [i.e. the CRR]). This means that most likely, many institutions will have to comply with most/all of the High Risk section.

These poses the following few issues:

- The Definition of “AI” in Annex 1 includes “Statistical approaches, Bayesian estimation” which could be far reaching and cover almost any existing method used by Member Firms. We should welcome some clarity from the EBA on where to draw the line between Statistical non-AI / non-ML models, and new AI/ML models for which these directives and requirements would apply.
- Annex 3 of the proposed AI Act lists as high risk, “AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems put into service by small scale providers for their own use” – which therefore classifies Retail Rating/PD models as High Risk items.
- If IRB ML models are classified as High Risk AI systems i.e. “AI systems intended to be used for making individual risk assessments”; then they will need to be compliant with another list of requirements resulting from AI EU Act, such as:
 - Conformity assessment
 - Reporting obligations
 - Technical documentation
 - Human oversight (human-machine interface tools)
 - Bias and discrimination prevention

We wish to highlight this as a possible unintended consequence. We propose that when credit scoring is used as an input to IRB models, the requirements of the AI act should not apply to IRB models for the following reasons:

- The use of AI in IRB models does not adversely affect the decision-making process of the lender, and therefore has no impact on the rights of individuals;
- Any potential risk of harm to consumers will already be mitigated at this early stage, given that institutions will have to comply with the requirements set out in the AI act to calculate the credit scoring.

We believe there is a need to differentiate AI applications which are used in the wider credit process (e.g. AI applications used in the valuation of collateral, are used as background tools in the process and do not affect a person’s access to essential services; or applications used in any phases following the initial disbursement of the loan which are used for monitoring and internal process efficiency as well as other subsequent uses of the credit scoring in other applications from those applications that can cause harm to consumers (e.g. capital consumption models, or marketing campaigns) . AI applications which do not cause consumer harm, including IRB models should be clearly excluded from the scope of the AI Act.

Q12: Do you see any additional challenge or issue that is relevant for discussion related to the use of ML models in the IRB context?

Until the further clarity is provided, it is very challenging to develop ML/AI models as the goalpost or acceptance-threshold is unknown. For example, the AI Act is not clear on who will be the Competent Authority. Member firms also seek some understanding on projected timelines as this could help with strategy discussions related to the use of ML models in the IRB context.

The regulatory framework for the approval of regulatory models is demanding. We would recommend that Supervisory teams should be sized and skilled properly to face the review of these models. These types of models will require an effective supervisory process.

Q13: Are you using or planning to use ML for collateral valuation? Please specify.?

A limited use of an ML model for this purpose is already underway. Some initiatives have been undertaken to explore the use of ML based models for collateral valuation in selected countries.

Q14: Do you see any other area where the use of ML models might be beneficial?

Potential beneficial uses of ML models span several applications in the Financial Services industry. Examples include:

- Willingness versus ability to pay in collections and recoveries
- Collateral valuations
- Know-your-customer (KYC)
- Application / transactional fraud
- Cyber security
- Operations
- Marketing
- Anomalies detection

Q15: 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/interpretability techniques are being developed in the industry. ML benchmarks are generally basic and self-explanatory (e.g. decision trees with very few nodes). If ML was expanded, we would have to introduce standard explainability techniques that would vary based on the context and types of explanations needed and would be generally accepted also by regulators.

From a control perspective, some firms have introduced a Model Ethics framework to govern the risk that models may infringe ethical principles.

Q16: 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?

We are not concerned about how to share the information gathered with different stakeholders. Understandable there are differing levels of technical knowledge and expertise within a firm.

In that regard, sharing information on models' interpretability with different stakeholders does represent a challenge. However, this is not a challenge peculiar to ML models and non-technical stakeholders do not need to get into technicalities, similar to the approach adopted when using other mathematical techniques to build IRB models. We believe that with the right involvement and training of different internal stakeholders this knowledge gap and that the senior management may have of these models can be bridged.

Some firms have developed various approaches and tools supporting such activities.

Some practical tips that some firms have adopted to bridge the knowledge gap in other areas, which have proved include:

- providing senior management with short summaries of key model characteristics. To this extent, MV reports already include model specifications and validation outcomes;
- relying on reader-friendly graphics to help committee members visualize the key elements of the model structure in presentations for committee meetings;
- providing dedicated training on the basic concepts and approaches used in ML;
- explaining the model score with segments (high, medium, low scoring) using a simple decision tree and top attributes as branches;
- shortlisting the most important variables;
- providing graphical representations of functions such as PDP/SHAP.

We would appreciate some regulatory assurance from the Regulators that having explainability and interpretability of results covered in a layered approach in the various stages of the life cycle would this be compatible with the CRR principle. For example, Senior Management will have sight of the final conclusions, whilst the technical people are involved with the granular details.

Q17:: Do you have any concern related to the principle-based recommendations?

We welcome the introduction of principle-based recommendations. Furthermore, it would be extremely beneficial if such recommendations explicitly stated the regulators' objectives and context, such that the interpretation and logic that institutions are expected to adopt is clear.

Additionally, the principle of regular updates does not appear to be compatible with CRR or EU 529/2014 in their current form. We would need clarification from the regulators on the conflict.

As raised previously in Q9, we believe it would be necessary to distinguish between scenarios where ML techniques are used exclusively to design the model and cases where the training of the model is allowed.

Also, we would encourage the regulator to :

- clearly differentiate what is expected to be minimum standards from recommended practices;
- Set out supervisory timing for issuing an opinion regarding these new approaches;
- provide usage of interpretability techniques that are well accepted and considered as robust enough for explaining the predictions made by the models and global behaviour consistently;
- provide guidance on supervisory expectations in terms of sufficient identification and elimination of possible unfair biases.

Appendix – Additional AFME Comments on the Discussion Paper

We have drafted this appendix to provide further details to our consultation response and to address general issues not addressed in the questions above.

General Comments:

The key points that we would like to raise are listed below.

ML-models to co-exist with traditional IRB approaches:

In the Executive summary (page 6) it is stated that- *“the EBA is considering providing a set of principle-based recommendations which should ensure an appropriate use of such techniques by institutions in the context of IRB models. This should ultimately ensure: i) a consistent and clear understanding of the prudential provisions, ii) how new sophisticated ML models can coexist and adhere with these requirements and so iii) that the outcome – in terms of setting capital requirements in a prudent manner – continues to be harmonised across Europe.”*

The EBA discussion paper references an IIF paper on Machine learning in credit risk which concludes that, the potential of ML models remains partially unused due to the FI's sentiment that the additional investments and (project) risks associated with the adoption of such approaches outweigh the potential benefits. The Paper acknowledges that better performing models would support the objectives of both Financial Institutions and Regulators.

Consequently, when describing how ML models can co-exist with traditional IRB approaches, it would be beneficial if the EBA recommendations clearly articulate the following:

- EBA's actual appetite for recognizing and supporting Financial Institutions' initiatives to explore the trade-off between the benefit and challenges introduced by ML-models as an alternative to traditional IRB approaches; and
- the interpretation of prudential provisions and any additional necessary conditions that competent authorities would require Financial Institutions to compensate for challenges and shortfalls experienced in the use of full ML-based models (e.g. increased complexity, lesser “explainability” and tie this to economic rationale).

Cost-benefit evaluation

The above referenced IIF Report cited one of the main challenges mentioned by institutions for the incorporation of ML models in credit risk modelling was the ‘lack of understanding’ from supervisors and the uncertainty with respect to the approval of such new types of models....For some types of models it could at least be challenging to demonstrate to the supervisors that the underlying economic rationale is conveyed.”

We wish to expand on the challenges already expressed in the discussion paper. The adoption of ML models commands significant additional investments and risks compared to the use of traditional IRB approaches. However, in the absence of guidelines that articulate how ML model outputs would complement the output of (existing) traditional IRB models and how competent authorities would use them to inform the setting of capital requirements, the benefits are uncertain.

Examples of current challenges include:

- Project risk - the redevelopment of IRB rating system typically follows a multiyear roll-out plan or an agreed return-to-compliance plan regularly monitored by Competent Authorities. The introduction of complex ML approaches increases significantly the delivery risk on such commitments;

- Internal justification of, the economic rationale for the additional investments required to scale resources and controls to IRB standards for data, infrastructures and experienced modelling resources for development and validation;
- Highly-regulated environments or industries experience challenges in resourcing highly qualified personnel with niche skills and effectively compete with specialized boutiques/start-ups. The ability to attract and retain qualified resources is fueled by the opportunity to innovate and the exploration of the full potential of ML models and cutting-edge applications;
- Ethical risks is an increasing concern for ML models compared to traditional IRB approaches.

We would welcome clarifications on:

- How the introduction of ML models as challenger models would inform the competent authorities' decision process and discretionary assessments in setting capital requirements;
- the expectations with respect to the use ML-based model output as input into traditional IRB approaches, or model (e.g. proportionality principle);
- the expectations of the internal controls applicable to challenger models.

Stability of model performance and velocity of model changes

The discussion paper (pages . 25 - Exec summary p.6 and others) states that: *"Where ML models are frequently updated, the reason for such regular updates needs to be analysed in detail and monitored by the institution. Generally, a break in the economic conditions or in the institutions' processes or in the underlying data might justify a model update. As credit risk is, however, not supposed to change frequently (in contrast to e.g. the market risk), such updates should, in general, not occur frequently. Therefore, the parameters of the model should generally be stable. Moreover, institutions should always compare the changes to the last approved model in order to make sure that many insignificant changes at a high frequency do not lead to an unnoticed material change within a certain time period. Therefore, institutions need to weigh the benefits against the potential risks of automatic and frequent updates of its IRB model."*

We wish to request that the EBA recommendations should articulate the rationale Financial Institutions should adopt (especially if different to the "EBA FINAL draft EBA/RTS/2013/06 RTS") and to identify those changes to ML models that would trigger pre/post notification and pre-approval by competent authorities.

AFME Contacts

Tola Gbadebo

Tola.Gbadebo@afme.eu

Constance Usherwood

Constance.Usherwood@afme.eu