Governance of Artificial Intelligence in Finance

Accompanying questionnaire for consultation

This questionnaire has been designed for consulting about the discussion paper published by the ACPR's Fintech-Innovation Hub on the governance of AI algorithms in the financial sector.

The public consultation aims to submit to financial actors and other concerned parties (researchers, service and solution providers, control authorities, etc.) guidelines sketched in the discussion paper for feedback, and more broadly to gather any useful comment, including on supervisory authorities' best practices.

Please send your answers to fintech-innovation@acpr.banque-france.fr before **September 4, 2020**. They may be illustrated through AI (specifically ML) use cases, whether already deployed or in progress.

Context

This section of the questionnaire pertains to the respondents, who are also invited to provide information about their identity and organization.

QUESTION 1: EXPERIENCE WITH ML

- What kind of knowledge or experience do you possess regarding AI in general and ML in particular (R&D, Data Science, operational tradecraft, etc.)?
- If you are answering on behalf of a financial institution, what is the level of familiarity with AI within your personnel (both in technical and in business roles)?

QUESTION 2: IMPLEMENTATION OF ML *(QUESTION SOLELY FOR FINANCIAL CORPORATIONS)*

- What are the ML algorithms implemented in your organization?
- For each algorithm type, specify their use cases and the type of environment (development, pre-production, production)?
- For each use case, according to which criteria and evaluation methods has the algorithm been selected (raw performance, explainability/efficacy trade-off, etc.)?
- What are the respective roles of the teams involved in the design and implementation of ML algorithms in your organization (Data Scientist, software architects, project management, business experts, compliance officers, etc.)?

Explainability principle

This section of the questionnaire refers to pages 12-18 of the discussion paper, wherein a number of expectations pertaining to the explainability of AI algorithms were outlined on the basis of exploratory works on three topics.

The relevance of those guidelines needs to be confirmed on several points, which are the object of the following questions.

QUESTION 3: DEFINITION OF THE EXPLANATION LEVELS

Are the four explanation levels emerging from this analysis (1: observation, 2: justification, 3: approximation, 4: replication) clearly defined? If not, indicate the points of misunderstanding.

QUESTION 4: ADEQUACY OF THE EXPLANATION LEVELS

Do those explanation levels appear to represent an adequate scale in the following senses:

- Do they span the entire spectrum of current and future applications of AI in finance, from full transparency all the way to algorithms operating as "black boxes"?
- Does the choice of four levels seem appropriate (if not, should there be fewer or more levels)?

QUESTION 5: PRACTICAL EXAMPLES OF EXPLANATION LEVELS

The table presented in section 3.4.1 of the discussion paper suggests an appropriate explanation level for a few use cases of AI in the financial sector.

- How suitable are those suggested levels? If insufficiently, for what reason?
- Are those suggestions adapted to your own usage scenarios of AI (specify those scenarios)? If not, in what sense?

Performance principle

This section of the questionnaire refers to page 10 of the discussion paper.

QUESTION 6: TECHNICAL PERFORMANCE METRICS

How do you view the technical performance metrics commonly used for ML (AUC or F1 score, GINI score, etc.), specifically:

- Their adequacy with respect to the various ML algorithms?
- The availability of methods to choose between those metrics?
- How those metrics are used (model validation, selection of its operating point, model drift detection, etc.)?

QUESTION 7: FUNCTIONAL PERFORMANCE METRICS

- Which functional metrics (KPI) seem relevant when evaluating an AI component? Do those metrics account for compliance requirements specific to the processes considered?
- Who should be responsible for defining functional metrics (technical or domain experts, with or without input from risk management and compliance teams)?

Stability principle

This section of the questionnaire refers to pages 10-12 of the discussion paper.

QUESTION 9: MODEL GENERALISATION

- What limits to the generalization power of ML models have been identified, whether in relation to overfitting or to intrinsic limits of the model?
- How can those limits be handled (out-of-sample testing, etc.)?

QUESTION 10: RETRAINING AS A SOURCE OF INSTABILITY

- Based on your experience, are model retraining phases (whether on a periodic or continuous basis) a source of model instability?
- What techniques are or could be used to limit this source of instability (non-regression testing with appropriate datasets, etc.)?

Appropriate data management principle

This section of the questionnaire refers to pages 7-10 of the discussion paper.

QUESTION 11: REGULATORY COMPLIANCE OF DATA MANAGEMENT

In your experience, which methods or techniques appear advisable in order to ensure compliance with various regulatory requirements relative to data management:

- The GDPR?
- Other cross-cutting regulations?
- Sector-specific regulations, such as the European IDD (Insurance Distribution Directive)?

Specify what stage(s) of the AI development process (design / training / prediction) involve these methods and techniques.

QUESTION 12: BIAS DETECTION AND MITIGATION

Which methods appear advisable in order to analyze biases in ML systems, for each of the following types of bias:

- Pre-existing biases in the input data fed to the ML models?
- Biases present in the algorithms themselves?
- Biases within the models produced by the algorithms – and in their decisions and predictions?

More precisely, which fairness metrics enable the identification of biases, for example those with a discriminatory nature?

Which methods can be used to mitigate the undesired biases thusly identified?

Integration in business processes

This section of the questionnaire refers to pages 21-24 of the discussion paper.

QUESTION 13: ROLE OF AI

- Which are or should be, according to you, the outlines of a method to assess the integration of AI components in business processes?
- What should such a method enable to evaluate: how critical the function of those components is, how disruptive they are with respect to the traditional process, what human-machine interactions are possible, etc.?
- What are your thoughts on maintaining "parallel" processes assigned to human operators so as to continuously evaluate and/or correct an algorithm's results?

QUESTION 14: AI ENGINEERING METHODOLOGY

- Should the engineering methodology used for AI differ from that used for traditional models, and more generally from standard software engineering practices? If so, in what way?
- How should, according to you, the ML model-building process take into account the integration of those models in business processes?

Internal control system

This section of the questionnaire refers to pages 24-28 of the discussion paper.

QUESTION 15: RISK MANAGEMENT

- How does the introduction of AI into business processes impact risk management: does it generate new risks or magnify pre-existing risks (specify the nature of those risks: operational or financial, legal, etc.)?
- Are new, AI-specific risk management methods called for (for example, calibration of ML models in order to limit the exposure to a given type of risk)?

QUESTION 16: FUNCTIONAL VALIDATION

- What should be the initial functional validation process of an ML model (i.e. prior to deployment in production)?
- Should functional validation be re-iterated when deploying a new version? Specify if the answer depends on the type of update (patch, improvement, etc.).
- How should ML components be continuously monitored for business risks?

QUESTION 17: INTERNAL MODEL UPDATE POLICY *(INTERNAL RISK MODELS)*

- On what conditions may, according to you, ML algorithms be used within "Basel models" in the banking sector, and within internal models in the insurance sector?
- How should an organization's internal model update policy take into account the use of ML in its internal models?

QUESTION 18: TECHNICAL VALIDATION

- What should be the initial technical validation process of an ML model (i.e. prior to deployment in production)?
- What technical indicators and methods should be used to continuously monitor ML components deployed in production?

Security and outsourcing

This section of the questionnaire refers to pages 28-31 of the discussion paper.

QUESTION 19: OUTSOURCING

Does the use of AI generate specific challenges or risks when its development, hosting or administration are outsourced? If so, which ones?

QUESTION 20 : SECURITY

- What is the impact of using ML on IT security?
- Which types of attack against ML models (causative attacks, surrogate model attacks, adversarial attacks, etc.) appear the most important to you, both in terms of occurrence likelihood and in terms of damage inflicted in case of success? Specify according to the type of ML model, the use case, and the environment (dedicated hosting servers or cloud services, etc.).

Multi-pronged approach to evaluation

This section of the questionnaire refers to pages 7-10 of the discussion paper, which suggest implementing a multidimensional approach for auditing processes using AI. The following questions aim to further define this approach.

QUESTION 21: ANALYTICAL EVALUATION

- Which of the following elements are available for evaluating an AI algorithm in the relevant organizations: the source code? Its documentation? The resulting models? The training and validation data? Specify if the answer depends on the algorithm type, the use case involved, or the context of the evaluation (internal validation, external audit, etc.).
- Do you use standardized documentation frameworks such as information sheets describing the algorithm, the model, or the data used?

QUESTION 22: EMPIRICAL EVALUATION

- Which of the empirical evaluation methods suggested in section 5.5.1 of the discussion paper (benchmark datasets or challenger models) seems more appropriate in your opinion?
- Is the architecture of data processing workflows and AI systems within the relevant organizations sufficiently modular and robust to enable this kind of functional testing at the data or model level?
- Are the data format and schema sufficiently standardized (or flexible) to support a data benchmarking method without incurring data integration costs by the supervisor?
- Analogously, are they sufficiently documented and transparent to support the integration of challenger models developed by the supervisor, without this approach being rendered unrealistic by an information asymmetry?

QUESTION 23 : EXPLANATORY METHODS

- Which explanatory methods (cf. appendix 11 of the discussion paper) are currently implemented among the use cases of AI to your knowledge?
- Do you know of any explanatory method for AI other than those described in this document? If so, which ones? Have they already been implemented and deployed?
- Does the most appropriate explanatory method to use depend on the algorithm type?
- Does it depend on the intended recipients of the explanation, and if so, in what way?
- Does it depend on the level of risk associated with the business process, and if so, in what way?