Survey on the use of Machine Learning algorithms

Objective: The survey is aimed at mapping the use and diffusion of Machine Learning algorithms within the national insurance market, with particular reference to those used by companies in processes having a direct impact on customers. Non-Machine Learning algorithms are expressly excluded from the survey.

How to fill out the survey: The questionnaire should be filled in via the LimeSurvey platform.
The contact person of each company has received an email from the address "sondaggi-noreply@bancaditalia.it" with the subject "IVASS - Invitation to the survey on Machine Learning algorithms".
Inside the email there is a link of this type: "`https://sondaggi.bancaditalia.it/ivass/xxxxxx...`"
By clicking on this link (specific to each company) the user will be redirected to the start page of the questionnaire.
.
The questionnaire in text format will be attached to the invitation email to anticipate/facilitate its completion on the platform.
It is possible to stop filling in the questionnaire and then resume it at a later stage (the system will automatically save the answers provided).
Some of the answers are mandatory to continue filling in the questionnaire.
Once the questionnaire has been filled in, the contact person will receive an email confirming that the answers have been correctly received by IVASS.

Timing:

30/09/2022   Deadline for responses to the questionnaire

Some definitions in the questionnaire:

- **"artificial intelligence":** software developed using one or more of the following techniques and approaches and which can, for a given set of human-defined objectives, generate outputs such as content, estimates, recommendations or decisions that influence the environments with which it interacts:
   - a) machine learning approaches, including supervised learning, unsupervised learning and reinforcement learning, using a wide range of methods, including deep learning);
   - b) logic-based and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inferential and deductive engines, (symbolic) reasoning and expert systems;
   - c) statistical approaches, Bayesian estimation, search and optimisation methods.

- **"Machine Learning algorithms':** a data analysis method that aims to automate the creation of analytical models that learn iteratively from a database. These are complex algorithms capable of finding solutions by learning from input data and based on experience, without the use of mathematical models and predetermined equations.

- **"blockchain/distribuited ledger technology"**: these are decentralised electronic transaction records distributed in a network of participants (so-called 'nodes' that all have the same copy of the database) without a centralised control. Blockchains are a particular type of distributed ledger technology.

- **"cloud computing"**: the ability for data, applications and services to be delivered through an architecture (internal or external) that provides shared, highly scalable computing resources accessible via the Internet.

- **"internet of things"**: means a set of technologies that enable the interconnection and exchange of data between the network and devices (cars, home automation, sensors, wearable devices, cameras, etc.).

- **"churn models':** refers to models that enable the identification of customer intention to churn and timely intervention with appropriate actions, improving the company's ability to retain customers and stemming potential losses.

- **"chatbot"**: refers to software designed to simulate a conversation with a human being.

- **"robo advisor"**: algorithms that provide insurance advice in an automated manner, without any human intervention, consistent with the client's characteristics and objectives.

- **"training phase":** refers to the learning phase of supervised Machine Learning algorithms performed by administering a large amount of data for which the answers are known.

- **"fallback":** refers to the means of enabling users to benefit from products in the event of unavailability or failure of the application based on Machine Learning algorithms.

- **"data-poisoning"**: refers to the possibility that the data used by the Machine Learning algorithm during the training phase may be polluted by a deliberate attack.

- "**bias**": bias in the results produced by artificial intelligence models arising from the design or quality of the data, definition of the algorithms or processes for interpreting the results.

- **"human in the loop":** presence of human oversight that can help mitigate risks in Machine Learning algorithm decisions (e.g. by providing a check before Machine Learning algorithm decisions become operational).

- "**explainability**" of Machine Learning algorithms: refers to the possibility of explaining the algorithm's decisions from the input data to different stakeholders (customers, internal audit, supervisor - with different degrees of detail).

- **'black box' algorithms:** these are Machine Learning algorithms that appear from the outside as a black box and for which it is not possible to provide a detailed analysis of the internal mechanisms (e.g. Neural Networks, Deep Learning)

QUESTIONNAIRE FOR INSURANCE UNDERTAKINGS

Survey on the use of Machine Learning algorithms

Question no.	Sections	Question	Answers
1	**Identifying data**	*Corporate name of the insurance undertaking*	`Alphanumeric field`
2		*Number of registration in the Register of Undertakings kept by IVASS*	`Alphanumeric field`
3		*Email of the person responsible for the models or innovation (or in his /her absence, of the person in charge of the Compliance function) if any clarification is required*	`Alphanumeric field`

Use of Artificial Intelligence and Machine Learning

Question no.	Sections	Question	Answers
4	**New Technologies**	*Does the company use artificial intelligence models?*	*Yes/No*
5		*If so, are these models accompanied by other innovative technologies?*	`Blockchain/Distributed Ledger Technology, Cloud Computing, Internet of Things (IoT), Big Data Analytics, Other (specify). Multiple choice answer`
6		*Are Machine Learning algorithms used for AI models?*	`yes / no (if the answer is no, the survey stops here)`
7		*If so, in what areas of operation?*	`Only in internal operational processes, only in processes having a direct impact on customers, in both areas, other (specify)`

The answers to the following questions refer ONLY to Machine Learning models having a DIRECT IMPACT ON CUSTOMERS

Governance of Machine Learning algorithms

8	**Governance**	Is there a specific policy for applications/services using Machine Learning algorithms?	Yes/No/in the process of being defined
9		Has the company changed its risk management, compliance, internal auditing and IT policies in light of the use of Machine Learning algorithms?	Yes/No/In progress (Please provide a description of the main changes)
10		Indicate the main risks identified for the use of Machine Learning algorithms and the measures taken to mitigate these risks	Open-ended answer
11		Is the coordination of the development and management of applications using Machine Learning algorithms centralised or distributed across different areas/departments?	Centralized/Distributed (Please provide a description)
12		How is the governance of the data used in Machine Learning algorithms ensured?	Open-ended answer
13		Have the Machine Learning algorithms undergone internal or external validation/auditing?	Yes/No/Under validation (if yes, provide a description)
14		Have specific KPIs/KRIs been developed to assess the performance/risks of Machine Learning algorithms? If yes, which ones?	KPIs/KRIs/under development (multiple answer) (Please provide a description)
15		Is there any human supervision (human in the loop) of decisions made by Machine Learning algorithms?	Possible answers: Total, Partial, None (Please provide a description)

16		*Are there any training plans for the company's internal staff using Machine Learning models in day-to-day operations to ensure their correct use?*	`Yes/No/under implementation (Please provide a description)`

Customer communications and consent

17		*Is information provided to customers on the use of Machine Learning algorithms in products and services designed for them?*	`Yes/No (Please provide a description)`
18		*If yes, what information is provided on the use of Machine Learning models, the data processed and the main variables that impact the decisions made by the algorithms?*	`Open-ended answer`

Security of Machine Learning algorithms

19	**Security of the Algorithm**	*Does the company perform regular audits to verify the integrity and robustness of applications using Machine Learning algorithms against possible cyber attacks?*	`Yes/No/Under development (Please provide a description)`
20		*Are backup and disaster recovery mechanisms in place for applications using Machine Learning algorithms?*	`Yes/No/Under development (Please provide a description)`
21		*What fall-back mechanisms are in place in the event of unavailability of the applications using Machine Learning algorithms?*	`Open-ended answer`
22		*What procedures are in place to protect algorithms from data-poisoning attacks?*	`Open-ended answer`

Explainability of Machine Learning algorithms

23	**Explainability**	Does the company use "black box" Machine Learning algorithms?	Yes/No (If yes, please provide a description of any minimum explainability criteria defined)
24		With regard to "black box" or particularly complex Machine Learning algorithms, does the company use tools to ensure explainability? If yes, please specify (e.g. challenging methods or specific explainability reports).	Yes/No/under development (Please provide a description and, if helpful, attach relevant documentation)

Fairness of Machine Learning algorithms

25	**Fairness**	How does the company identify bias in the data used for Machine Learning algorithms (even indirectly, through algorithm inference) and prevent the exclusion or discrimination of customer groups?	Open-ended answer
26		Are there any internal mechanisms that signal unwanted exclusion or discrimination of customers caused by Machine Learning algorithms? If yes, please describe these mechanisms and the actions to be implemented in case of bias in the results.	Yes/No/Under development (Please provide a description)
27		In the case of Machine Learning algorithms used for pricing, are the results checked to see whether customers with the same risk profile are assigned different prices? If this is possible, is there a maximum price range set by the company?	Yes/No (Please provide a description)

Outsourcing of Machine Learning algorithms

28	Outsourcing	By what process does the company determine the critical applications using Machine Learning algorithms, which require special caution in case of outsourcing (Art. 62 IVASS Reg. 38 / 2018)?	Open-ended answer
29		In particular, has the company assessed the possible risks (e.g. lock-in or concentration) arising from the outsourcing of Machine Learning algorithms?	Yes/No/Under evaluation (Please provide a description)
30		Have the possible operational, reputational or legal repercussions of operational disruptions, malfunctions, loss or theft of data for outsourced applications using Machine Learning algorithms been considered? If yes, please report the main evaluations.	Open-ended answer
31		What control processes are in place in the case of machine learning algorithms managed by external outsourcers, with particular reference to explainability and fairness?	Open-ended answer

Specific Machine Learning algorithms used by the company
Please list the main Machine Learning algorithms used (up to a maximum of 3). Non-Machine Learning algorithms and those that, although using this technology, are dedicated to the companies' internal processes are excluded from the survey.

32	Main Machine Learning Algorithms Used	Specify whether the company uses an in-house developed Machine Learning model or a solution developed externally. In this case, please specify the commercial name of the solution/platform.	Internally developed / externally acquired model (in the second case, indicate the name of the acquired solution).
33		Specify in which field the Machine Learning algorithm/application is used.	Possible answers: Customer profiling, Policy pricing, Cross-selling/Up-selling of products, Customer

			behaviour analysis (Churn models), Claims management, Chatbot communications, Robo Advice, Text Analysis, Sentiment Analysis, Fraud prevention, Other (specify). Multiple choice answer
34		*Type of Machine Learning algorithm.*	*Possible answers: Support Vector Machine, Decision Trees, Random Forest, Association Rules, Neural Networks and Deep Learning, Clustering, Nearest Neighbor, Other (specify)*
35		*Describe the purposes and methods of using the Machine Learning algorithm.*	*Open-ended answer*
36		*Specify the type of products for which the Machine Learning algorithm is used.*	*Life: unit-linked, hybrid, with-profit, term life insurance, LTC, Non-Life: Home, Health, Accident, Motor, General liability, Credit and Suretyship, Travel, Other.*
37		*Specify the development status of the Machine Learning algorithm.*	*Under production, Under testing, Under design*
38		*Is the service using the Machine Learning algorithm accessible via the Web, via an App on a smartphone or both?*	*Possible answers: Web, App, Both*
39		*Is the Machine Learning algorithm managed by external outsourcers? If yes, please indicate the outsourcers.*	*Yes/No (if yes, please indicate the name of the outsourcer)*
40		*What are the main data required from customers and which are taken into account for the training / validation / testing phase of the Machine Learning algorithm?*	*Open-ended answer*

41	How does the company ensure that the data used for the training / validation / testing of the Machine Learning algorithm is accurate, appropriate and complete?	Open-ended answer
42	How does the company ensure that the data used for the training / validation / testing of the Machine Learning algorithm is non-discriminatory towards customers, in particular that there is no bias - even indirect - on factors such as gender, ethnicity, religion, nationality, sexual orientation?	Open-ended answer
43	Are bias monitoring, detection and correction activities in place?	Open-ended answer
44	How often is the calibration of the Machine Learning algorithm reviewed?	Open-ended answer
45	How are the operation, decisions and results of the Machine Learning algorithm documented?	Open-ended answer
46	Does the company use other Machine Learning algorithms?	Yes/No (if Yes, go back to question 32)