

Explainable Artificial Intelligence and its impact in the Financial Sector



Management Solutions

Index

1 Who we are

2 What is XAI

3 Challenges and solutions

4 A use case

1

Who we are?

About Management Solutions

Management Solutions is an international consulting company with over 3,300 professionals, offices in 44 cities and operations in over 50 countries

A global and international company



50+ countries



1,200+ customers



+ 600 F.I. ⁽¹⁾

A confirmed success



ECB-accredited reference consultant



20+ years of experience



Strong growth in 20 years (x32)

Reference consultant accredited by supervisory authorities and supranational bodies (ECB, FCA, PRA, BoS, DNB, GNB, SNB, MNB, SBIF, SBS, BCCR, SSN, EIOPA, MEDE/ESM, WB, EIB, etc.)

One team



+ 3,300 consultants



A multidisciplinary team with 1/3 "quant" profiles

Presence in different industries

- Finance
- Energy
- Telecommunications
- Insurance
- Public administration
- Distribution and industry

Offices



Europe

Madrid, Barcelona, Bilbao, Coruña, Málaga, London, Frankfurt, Düsseldorf, Paris, Warsaw, Wrocław, Zurich, Milan, Rome, Bologna, Lisbon, Amsterdam, Oslo, Copenhagen



America

Toronto, New York, New Jersey, Houston, Boston, Birmingham, Pittsburgh, Atlanta, San Juan de Puerto Rico, Mexico City, Monterrey, Querétaro, San José, Medellín, Bogotá, Lima, Quito, Sao Paulo, Rio de Janeiro, Buenos Aires, Santiago de Chile



Asia, Africa, Oceania

Beijing, Istanbul, Johannesburg, Sydney



Through its **44 offices**, MS regularly serves clients operating in 50 countries in Europe, America, Asia and Africa

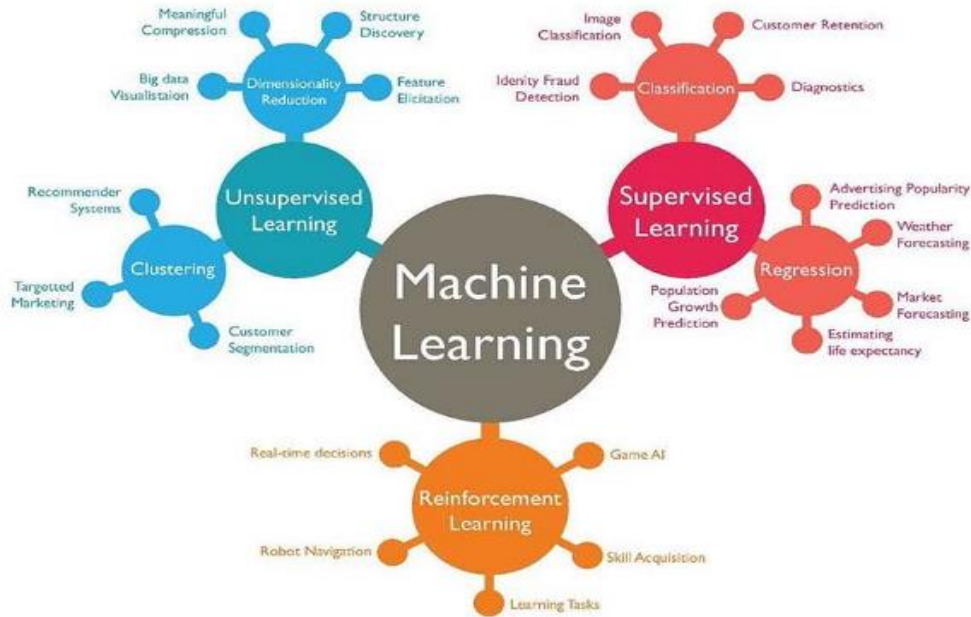
2

What is XAI

Challenges of advanced modelling

The need to understand AI decisions stems from the widespread use of advanced modelling, and its critical impact on human and societal decisions

Context



- **Digital transformation** has made available a massive amount of structured and unstructured data.
- **Increased appetite for modelling** and **technological infrastructures** that provide greater **computational capacity**.
- The proliferation of the use of **machine learning** has generated the **need to explain its results**.

Can we understand well enough the decisions made by AI algorithms...

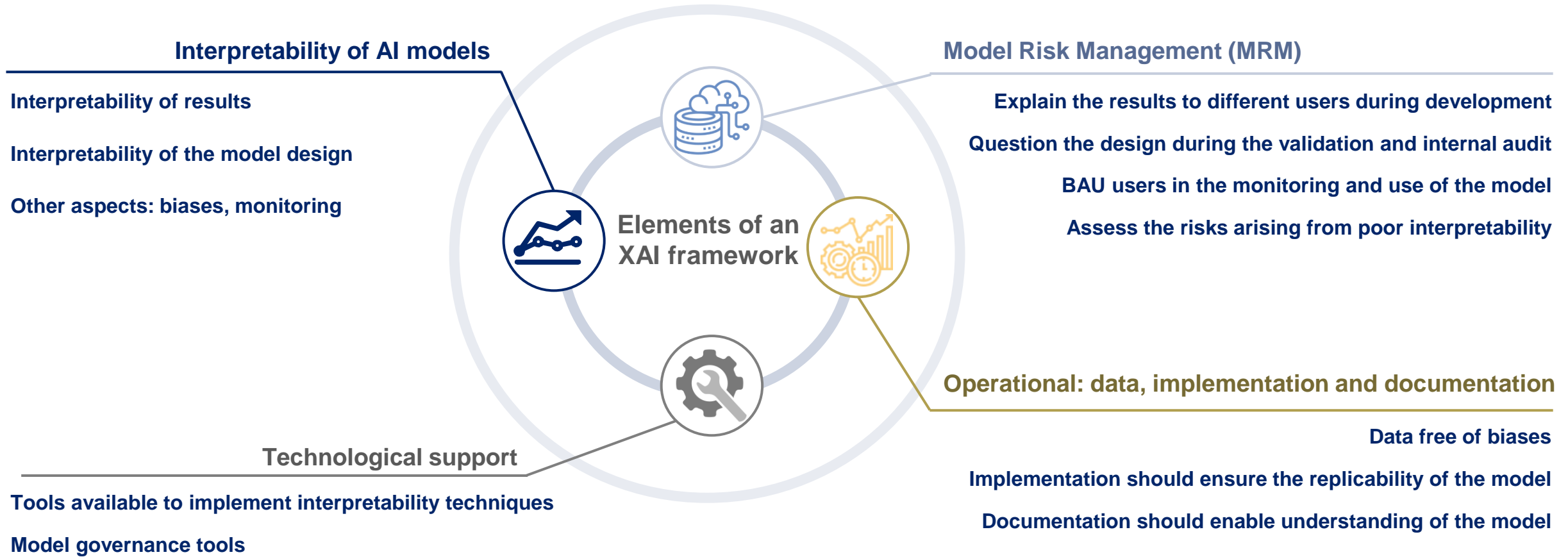
... especially when they have a critical impact on human, moral, or social decisions?

- 1 Since 2012, Deep Learning methods have dominated benchmarks of model predictive power and yield better results in business cases ¹.
- 2 There is a trade-off between the predictive power of Machine Learning models, and their ability to produce explainable and interpretable predictions.
- 3 This has led to the emergence of Explainable Artificial Intelligence (XAI), a field of AI that focuses on understanding and interpreting it.

Explainable Artificial Intelligence (XAI) can prevent and mitigate risks that affect all sectors, and is increasingly found in all proposed regulatory frameworks for AI adoption



XAI Framework: What elements make up an XAI Framework?



3

Challenges and solutions

Challenges in XAI

Challenges in XAI: Main challenges and proposed solutions

Challenges

Solutions

Interpretability of the model

Model risk

Data, documentation and implementation

The modelling process must be reproducible

Interpretability explanations may not be unique

The sequence of most likely model predictions may not be consistent

There is no mathematical formalism to describe the properties of many ML models

There are no globally accepted measures of evaluation for ML models, nor of interpretability.

Model users must be able to access the interpretability modules to understand the predictions

There is implementation variance in certain AI algorithms.

The level of documentation should reflect the complexity of the algorithm.



Randomisation control

Comparison of interpretability techniques, verifying consistency in similar data. Creation of interpretability reports

Include specific sensitivity analyses when revising the model.

Comprehensive documentation on models and their mathematical basis

Application of a diverse and comprehensive testing framework throughout the model lifecycle.

Implementation of XAI modules

Adequate documentation and implementation controls.

Include in detail the results of the interpretability tests, guaranteeing their replication by a third party.

4 A use case

Interpretability of an intake model with ModelCraft® (1/2)

Problem: Intake

Objective: apply the ModelCraft® interpretability module to explain and detail the predictions made by an advanced Machine Learning model

- Find the best Machine Learning model that fits the data in a credit admission portfolio using ModelCraft®.
- Application of the ModelCraft® interpretability module for the Machine Learning model to be developed.
- Generate ModelCraft® automatic reporting of model development and interpretability information explaining the average weights of the model over the entire portfolio and the specific weights of the variables on each prediction.

Approach

Two phases:

- Application of the ModelCraft® interpretability module and analysis of the results
- Customisation and automatic report generation of the development and interpretability process.

Benefits

- Advanced XgBoost model with 94% hit rate
- Interpretability module that allows to know the influences of the variables in the prediction of a specific client.

Applied strategy

Using ModelCraft® as an accelerator

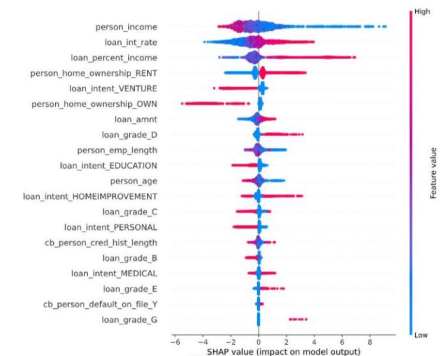
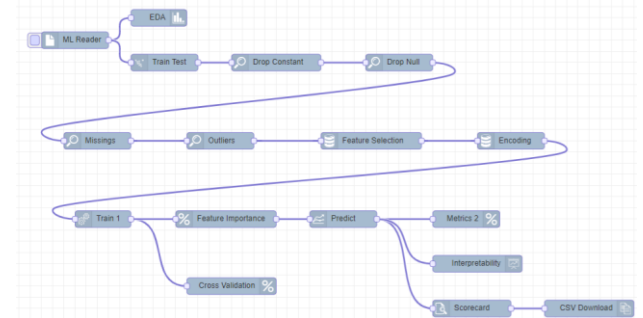
Implementation of various development algorithms to test the best accuracy:

- XgBoost (94% Accuracy - 87% AUC - 94% F1 Score)
- LightGBM (94% Accuracy - 86% AUC - 94% F1 Score)
- R. Forest (94% Accuracy - 86% AUC - 93% F1 Score)
- Logistics (87% Accuracy - 76% AUC - 86% F1 Score)

Interpretability module:

- Obtain, on average, the weights that the variables have for the model's decision on any prediction.
- To be able to analyse, for each prediction, the influence of each variable on the admission decision.

In order to cover both concerns, the SHAP algorithm is applied.



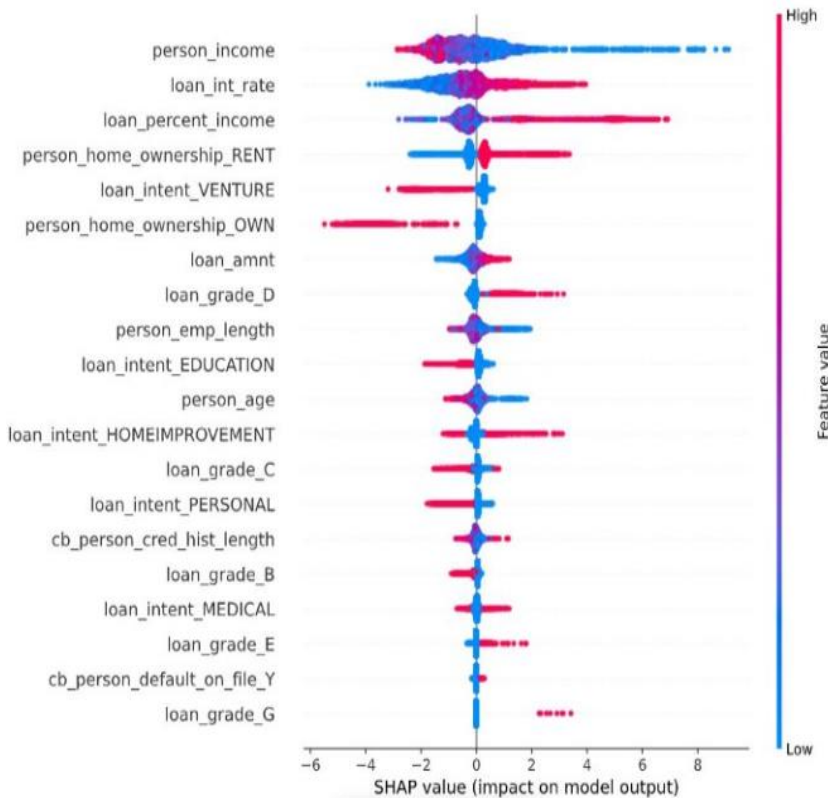
4 A use case

Interpretability of an intake model with ModelCraft® (2/2)

Admission

Objective: apply the ModelCraft® interpretability module to explain and detail the predictions made by an advanced Machine Learning model

Results



- Income turns out to be the most important variable in the model.
- People with low incomes (blue observations) show a positive correlation with default.
- The second characteristic identified as most important: interest rate of the loan.



- The influences of each variable in the model are also obtained.
- Owning a real estate property has a great impact on the admission,

Explainable artificial intelligence (XAI)
Challenges of model interpretability

