XTRACTIS®

The General Reasoning AI for Trusted Decisions

Automatic Discovery of Robust, Intelligible & Auditable Predictive Knowledge for High-Risk Applications

by Collective, Evolutionary, Corrective & Adaptive AI with Continuous Logics

[Augmented Fuzzy Cognitive AI]

ARISTOTE-X Conference

December 11, 2025

v1.0

Prof. Dr. Zyed ZALILA

President – Founder

www.xtractis.ai | zyed.zalila@xtractis.ai

Two approaches of AI – The original schism

Two paradigms [08/1956]

C-AI = **Connectionist AI** (Neurobiologists, Neurophysiologists, Neuropsychologists)

S-AI = **Cognitive** / Symbolic **AI** (Logicians, Cognitive Psychologists)

Perception

Language

Reasoning

slide 57

	UNINTELLIGIBLE (Opaque)	**INTELLIGIBLE** (Transparent)
Automatic Process **ROBUST**	**Con-AI** **Machine Learning** **Deep Learning** **Graphs** Neural Network Models	?
Manual Process **NON-ROBUST**	X	**BC-AI / FC-AI** **Maieutics** **Deduction** **Abduction** **Rules** Binary Expert Systems Fuzzy Expert Systems

Solution

Combine the advantages of these 2 approaches (without any Neural Network!), for reasoning and perception

➡ **Augmented Fuzzy Cognitive AI**
[Zalila & al. 1988-2025]

slide 28

Modeling Complex Processes – Data Driven Modeling Schools

slide 57

slide 35 slide 53

MACHINE LEARNING MODELING
(connectionism, graph theory, statistics, linear algebra)

Automatic Modeling
⊕
Binary Math

Opaque and **Robust** **Models**

COGNITIVE MODELING
(logics, relations, semantic, nonlinear math)
➡ Knowledge Discovery

Human Modeling
⊕ **Binary Logic**
⊕ **Continuous Logics**

Intelligible and **Non-Robust** **Models**

Automatic Modeling
XTRACTIS®
⊕
Fuzzy Relations of Order N
using Continuous Logics

Intelligible **and** **Robust** **Models**

Robust = high predictive performance
Intelligible = transparent model, revealing all its internal decision-logic

idagrai LABS

XTRACTIS®

Complex Modeling Techniques – Intelligibility × Performance Map

MACHINE LEARNING MODELING

G. Hinton, Y. LeCun, Y. Bengio
1990-2010
CONVOLUTIONAL NEURAL NETWORK

Ivakhnenko & Lapa | 1965
MULTILAYER PERCEPTRON
Machine Learning with Binary Math

J. Berkson | 1950
LOGISTIC REGRESSION
Automatic Induction with Binary Math

J. Friedman | 1999-2001
BOOSTED TREE
Machine Learning with Binary Math

L. Breiman | 2001
RANDOM FOREST
Machine Learning with Binary Math

F. Rosenblatt | 1957
PERCEPTRON

COGNITIVE MODELING

slide 27
slide 60

Z. Zalila & al. | 2002-2025
XTRACTIS
***Automatic** Inductive Reasoning*
***Automatic** Deductive Reasoning*
***Automatic** Abductive Reasoning*
with an infinity of Continuous Logics

L. Zadeh | 1965
J. Łukasiewicz | 1920
FUZZY EXPERT SYSTEMS
Human Inductive Reasoning
***Automatic** Deductive Reasoning*
with a Continuous Logic

Shortliffe & Buchanan | 1975
Aristotle | -4th c. BC
BINARY EXPERT SYSTEMS
Human Inductive Reasoning
*Human or **Automatic** Deductive Reasoning*
with Binary Logic

R. Bacon | 13th c.
al-H. ibn al-Haytham | 11th c.
EXPERIMENTAL SCIENTIFIC METHOD
Human Inductive & Deductive Reasoning
with Binary Math

Performance

Intelligibility

idagrai LABS

XTRACTIS®

12 AIs & Data-Driven Modeling Techniques

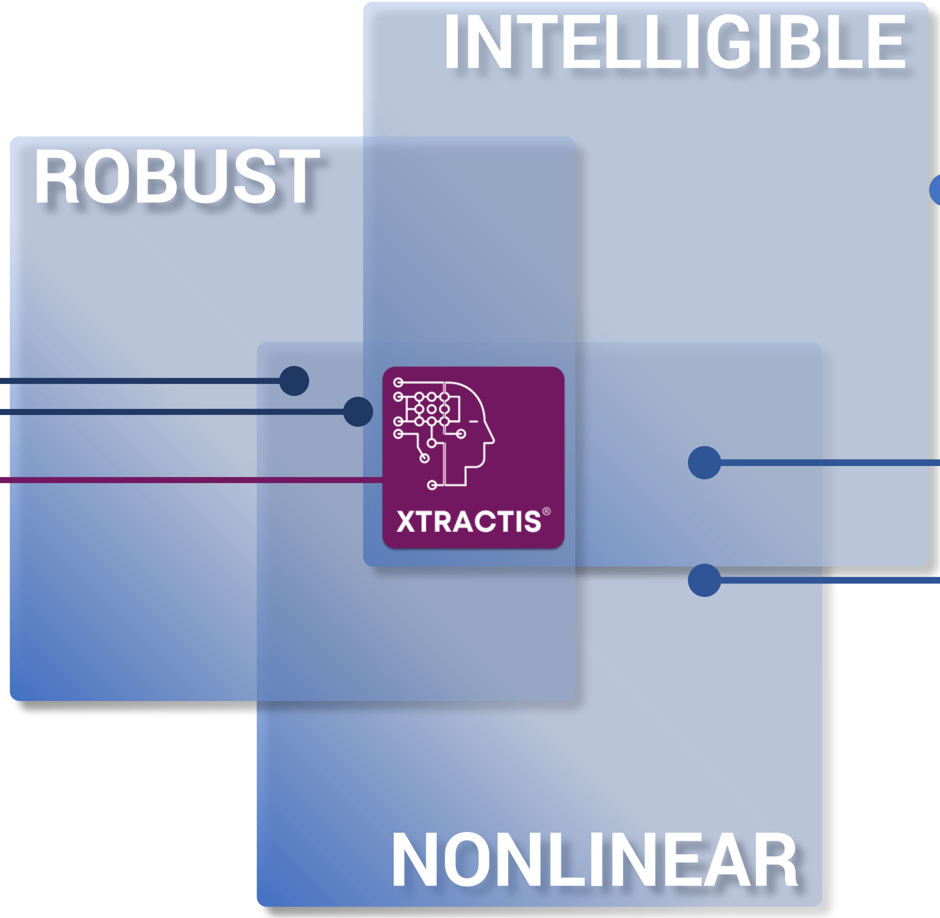
for Complex Processes and Phenomena

slide 34

Boosted Trees
Bayesian Networks
Kernel Support Vector Machine
Deep Learning / Neural Networks

slide 57

Random Forests

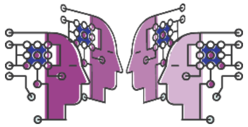


Support Vector Machine
Logistic Regression

CART Decision Trees
Binary Expert Systems
Fuzzy Expert Systems

Polynomial Regression

Partial Least Squares



Self improving Reasoning
for Knowledge Discovery

Augmented Fuzzy Cognitive AI

slide 18

slide 28

Theory of Fuzzy Relations of order N
+ Collective, Evolutionary, Corrective & Adaptive Inductive Reasoning

Robust = high predictive performance
Intelligible = transparent model, revealing all its internal decision-logic

General AI

- MULTI-PURPOSE for any business/**scientific** field
- Suitable for **HIGH-RISK CRITICAL APPLICATIONS** *when AI-driven decisions impact human lives or the environment or cause economic losses.*

Reasoning AI

Exobrain augmenting the **3 human reasoning** modes with an infinite plasticity (learns to reason better)

1. **INDUCTION**: automatically extracts knowledge-based models from data, as do scientists applying the Experimental Scientific Method
2. **DEDUCTION**: instantly predicts the outputs for a new case
3. **ABDUCTION**: discovers the most optimal solutions satisfying a fuzzy multi-objective request

Trusted Decisions

- Produces the most **ROBUST** and **INTELLIGIBLE** models = AI-driven decision systems having the highest predictive capacity AND understandable by humans
- **Auditable** by experts and can be **certified** by the regulator before deployment to end-users
- Decisions are instantly computed **RATIONALLY** & **DETERMINISTICALLY**

by design

All specificities are **natively** derived from the scientific foundation of our algorithms

Trusted AI for High-Risk Critical Applications when AI-driven Decisions Impact Human Lives or the Environment or Cause Economic Losses

HEALTH / PHARMA

Automated diagnoses for Personalized Medicine (from metabolic, epigenetic, physiological & anatomopathological data), Optimal Formulations & Drug Discovery, Monitoring, Virtual Screening, Protein Homology, Toxicity.

INDUSTRY / R&D

Product Design & Ergonomics, Sensory Marketing & Engineering, Smart Industry, Product Design, Quality Control, Maintenance & Diagnosis, Logistics, Risk Analysis, Environment, Geomatics, Optimization, ADAS, Autonomous Vehicles.

DEFENSE / CYBER / SECURITY

Command & Control, Autonomous Systems & Devices, Malicious Activities, Crime & Surveillance, Cybersecurity, Operational Research.

FINANCE / BUSINESS

Scoring & Risk Analysis, Econometrics, Malicious Activities, Venture Capital, Behavioral Finance, Wealth Management, Real Estate Finance, Strategy, Marketing & CRM, Legal, HR & Administration, Operational Research.

Ethics & Governance of AI for Health
(June 2021)
3rd guiding principle

EU AI Act
(February 2025)
New principles to rule the use of AI systems

Trusted AI – Intelligibility & Explainability, mandatory for critical decisions

The best Model we induced from your data is:

$(x_1$ is $\sim A_1)$ and $(x_2$ is $\sim A_2) \rightarrow$ y is $\sim B_1$

. . .

$(x_3$ is $\sim A_3)$ or $(x_4$ is $\sim A_4) \rightarrow$ y is $\sim B_3$

. . .

Example of an XTRACTIS Model

XTRACTIS® REVEAL

INTELLIGIBILITY of the internal logic of the decision system
(transparent models)

slide 60

For case #257, my decision is:
$\{B_8|0.9, \ldots, B_1|0.2\}$
and I explain my reasoning…

XTRACTIS® PREDICT

EXPLAINABILITY of each decision made
(using the model already induced)

Audit & Certification of AI systems
(for Trusted AI-driven Decisions)

Increase of expert knowledge

Quick generation of IP assets
(models belong to licensee)

Consolidation of decision-making processes
(automated, transparent and reliable decisions)

Reduction of ethical and moral biases
(Responsible AI by design)

Compliance with the new regulations for high-risk AI applications
(EU AI Act, WHO guiding principles)

XTRACTIS® Software Robots – Extending by design the 3 Human Reasoning Modes

INDUCTION

XTRACTIS® REVEAL

Inductive Robots automatically discover robust knowledge-based models from data

(Collective and Evolutionary)

DEDUCTION

XTRACTIS® PREDICT

Deductive Robots predict the outputs for a new case in real-time

(Collective and Deterministic)

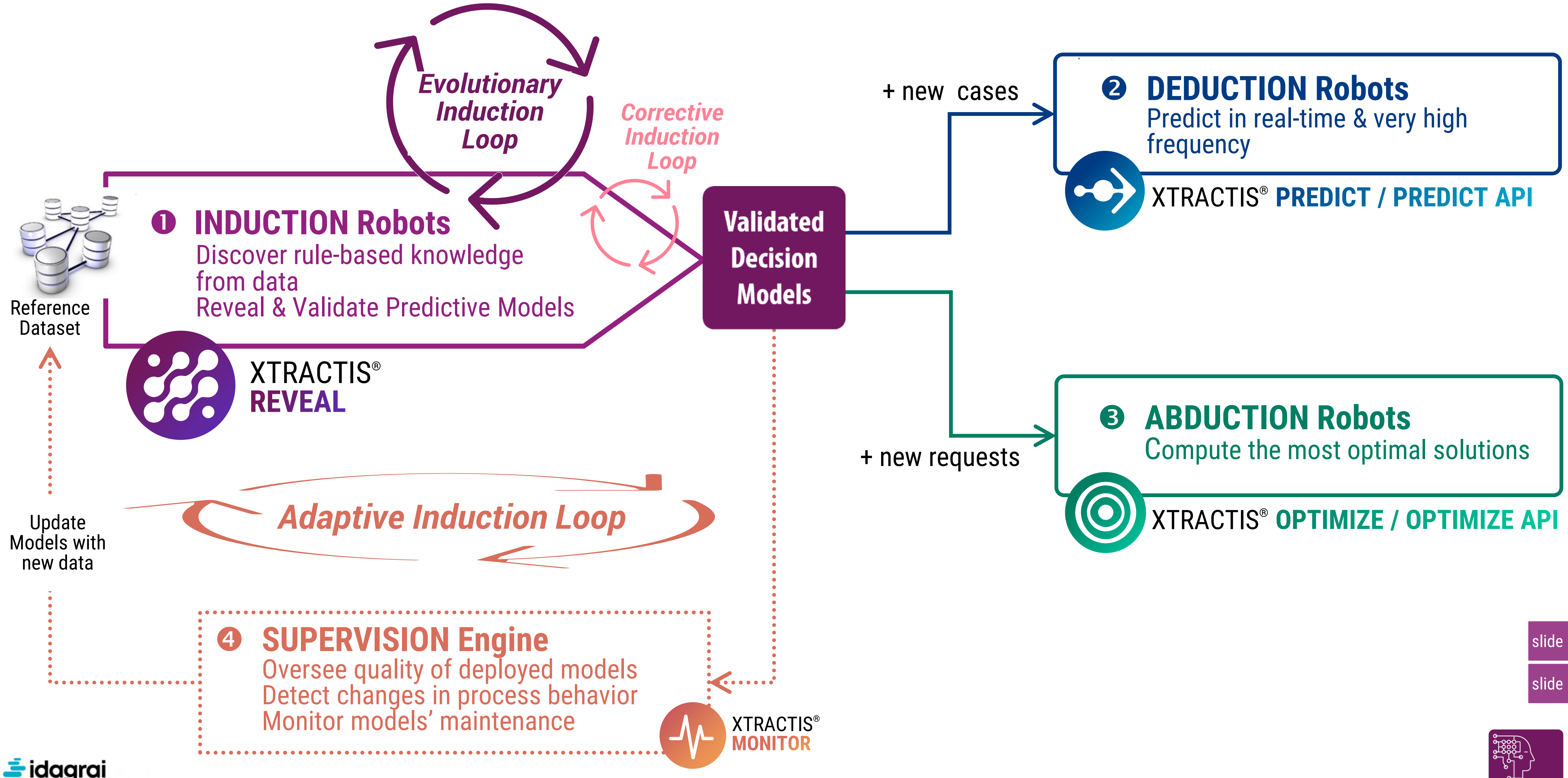
ABDUCTION

XTRACTIS® OPTIMIZE

Abductive Robots discover the most optimal solutions satisfying a fuzzy multi-objective request

(Collective and Evolutionary)

XTRACTIS® workflow — From data to validated models to trusted decisions & solutions



slide 58
slide 59

Data for XTRACTIS® – Structured, Quantitative/Qualitative

Variables to predict

Intelligible potential predictors

Regression | *Scoring / Binomial Classification* | *Multinomial Classification*

Label	Binary Variable	Nominal Variable	Numerical Variable	…	Var 1 to predict	Var 2 to predict	Var 3 to predict
Observation 1	0	Modality A	1.2558	…	2.35	1	D
Observation 2	1	Modality C	0.2356	…	1.256	0	A
Observation 3		Modality D	4.568	…	12.03	1	C
…	1		125.36	…	1.002	1	A
Observation n	0	Modality A		…	7.056	0	F

Observations

Missing values allowed

Production of Decision Models

Humans + their XTRACTIS Exobrain at Work!

Problem Setting

- Problem Definition
- Potential Predictors
- Variables to Predict

Reference Data Prep

- Compliant Data:
  - numeric and categorical variables
  - small & big data
  - missing values without imputation
  - structured in tables
- **Intelligible** preprocessing of unstructured data (image, signal, spectrum, text)

Evolutionary Induction Loop

XTRACTIS® REVEAL

- Collective Competitive **Evolutionary** Decremental **Holistic** Automatic Induction of Models
- Automatic Selection of Impacting Predictors
- Thousands of Inductive Strategies producing thousands of fuzzy rule-based-Models
- Intensive Cross-Validation
- Ranking of Models

XTRACTIS® Models **Intelligibility** (transparent model)

+ Built-in Structure Report:

- Reveals all the internal decision logic of the Model
- Ensures that humans understand the Model

Final Audit / Certification

- Selection of Top-Model by Expert/Modeler
- Audit by Expert
- Certification by Regulator

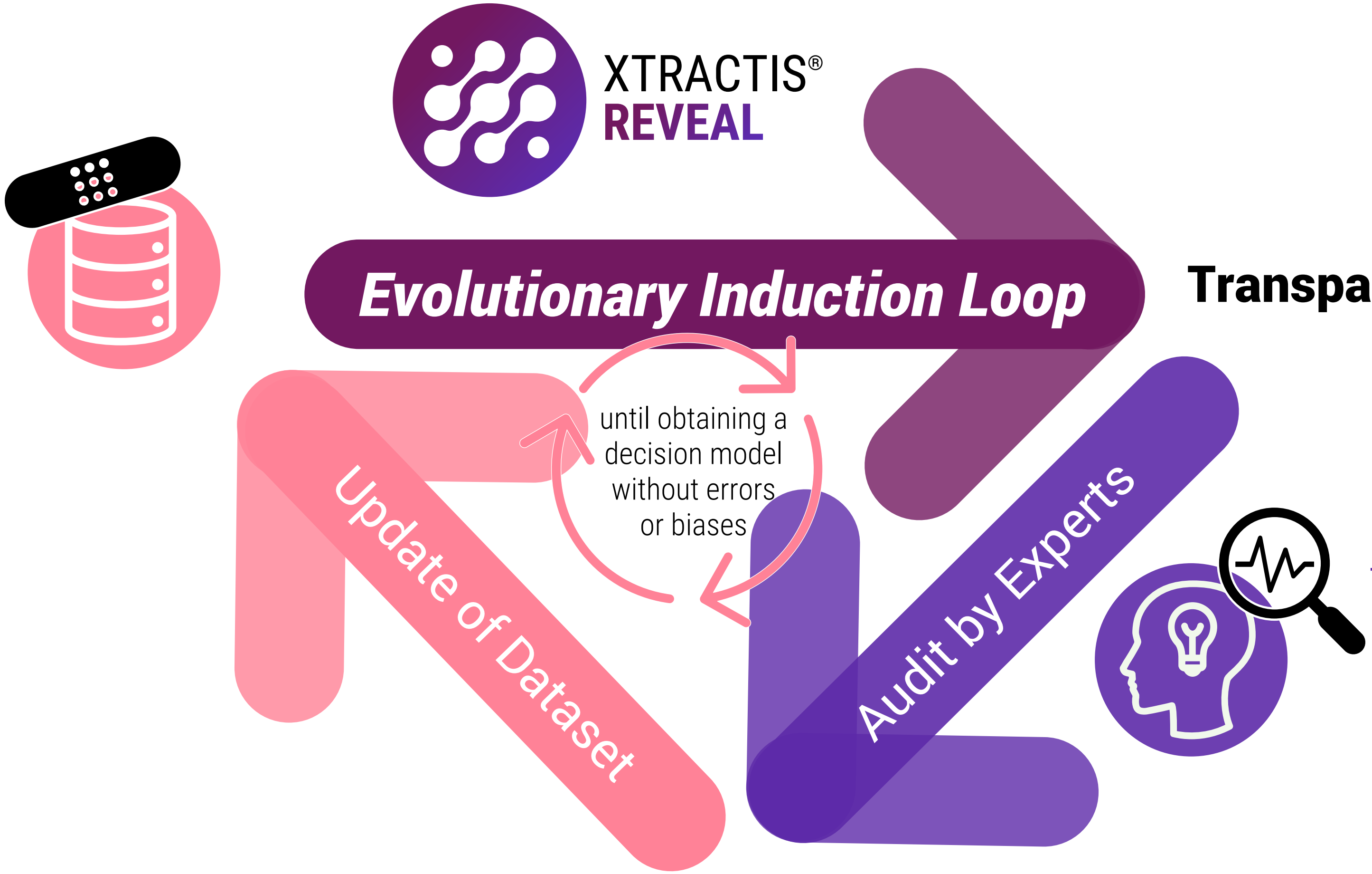
Deployment

Pushing Validated/Certified Models to end-users for real-time Predictive & Prescriptive Analysis (API, C code)

XTRACTIS®

Corrective Induction Loop – Detection of Errors and Biases

slide 10
slide 44



Transparent & **Robust** Model

Detection of **biased rules** or **“stupid” rules**

Detection of **errors** or **conflicting** cases in the training dataset

Deployment of AI Decision Models for Predictive Analysis

Predictions by Fuzzy Deduction

XTRACTIS®
PREDICT / PREDICT API

Validated Model

+

New Case/Situation

- Real-time predictions at 7.6+ million decisions per second, on 8-core CPU @2.5 GHz, offline or online (4 predictors, 8 rules)
- For each new case, instantaneous triggering of the involved rules, computation of each of their fuzzy decision and then their collegial final prediction
- Collective **Deterministic** Automatic Deduction from Models
- **Built-in Prediction Report** explaining how the decision is made from the Model's rules

Explainability & Traceability of each predicted decision

Deployment of AI Decision Models for Prescriptive Analysis

Optimization by Fuzzy Abduction

XTRACTIS®
OPTIMIZE / OPTIMIZE API

Validated Model

+

Multi-objective request on the predicted variables with binary or flexible constraints on predictors

→

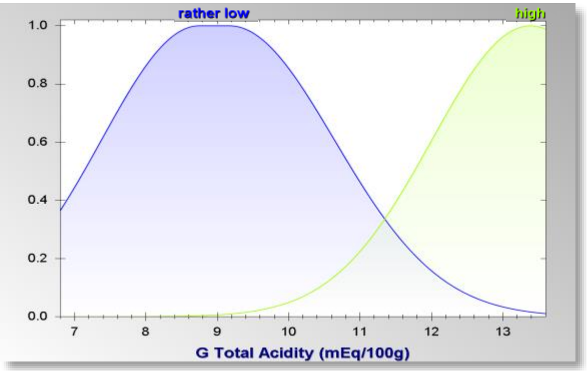
- Real-time search on CPU of the fuzzy-optimal solutions, i.e., the most satisfying ones, in the multidimensional operating space
- By computing the predictors values that best satisfy this request and its constraints
- Collective Competitive **Evolutionary** Automatic Abduction from Models
- **Built-in Request Satisfaction Report** confirming the relevance of each optimal solution and revealing the most limiting constraints

Explainability of each discovered solution

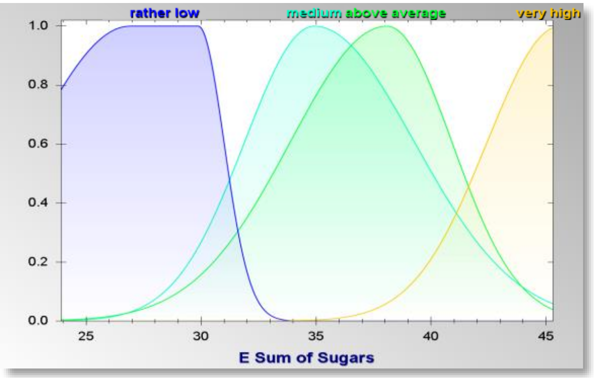
XTRACTIS® model – Making explicit an implicit unconscious decision

Sweet perception of a fresh tomato: 2 variables, 6 classes, 4 rules (complexity 33.0)

>> Classes



predictor #1: *Total Acidity*



predictor #2: *Sum of Sugars*

>> Rules

Rule ❶
If *Total Acidity* is *rather low*
And *Sum of Sugars* is *rather low*
Then *Sweet* equals *3.39*

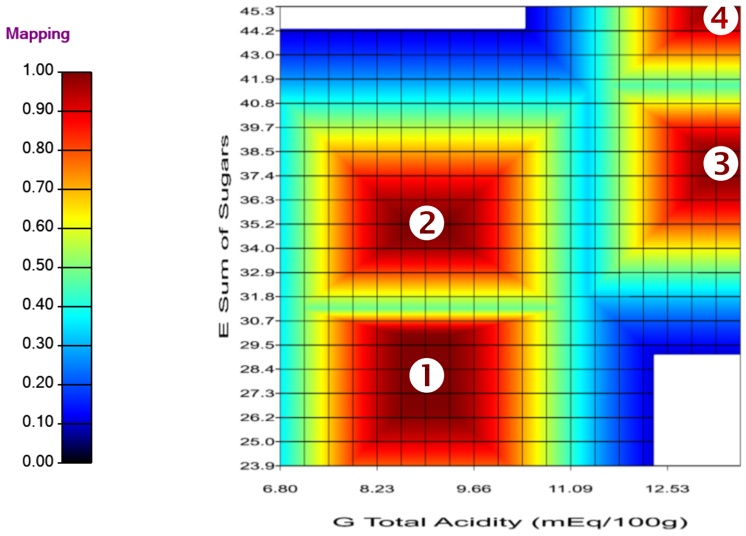
Rule ❷
If *Total Acidity* is *rather low*
And *Sum of Sugars* is *medium*
Then *Sweet* equals *7.19*

Rule ❸
If *Total Acidity* is *high*
And *Sum of Sugars* is *above average*
Then *Sweet* equals *3.30*

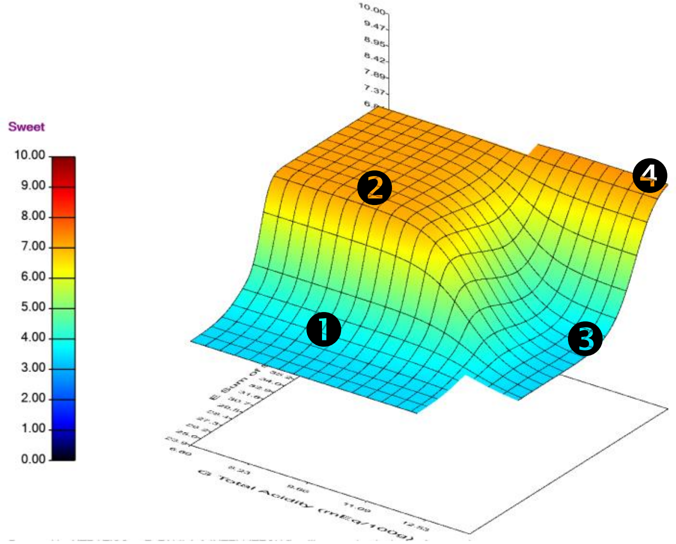
Rule ❹
If *Total Acidity* is *high*
And *Sum of Sugars* is *very high*
Then *Sweet* equals *7.49*

>> Inference

Mapping



Decision Surface



Results from XTRACTIS®
REVEAL 9.1.16325

Data Source:
INRA (Institut National de la Recherche Agronomique) et CTIFL (Centre Technique Interprofessionnel des Fruits et Légumes) - 7th Sensometrics Conf., July 2004, Davis, CA, USA

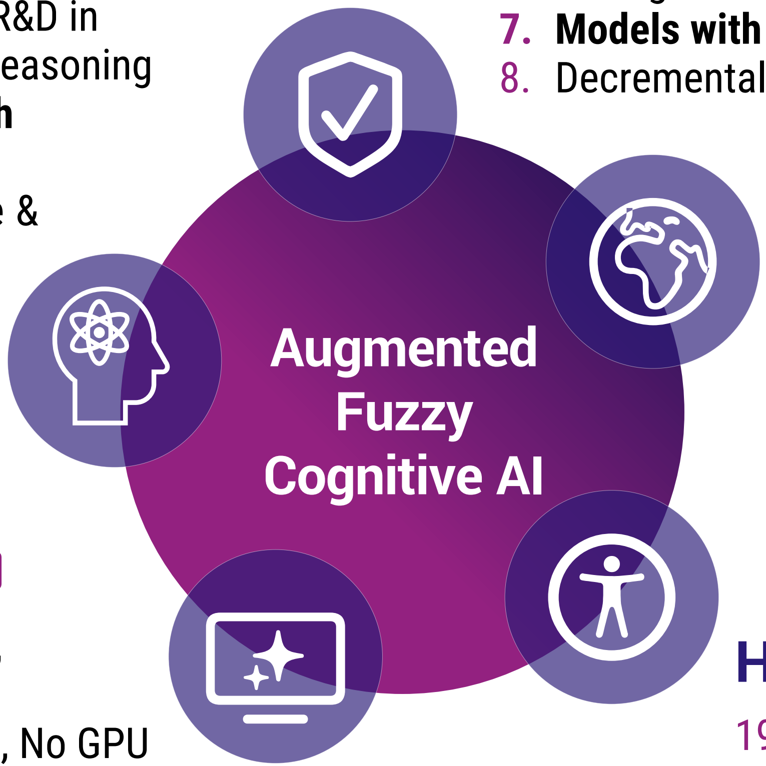
Scoring
Binomial Classification
Multinomial Classification
Regression

TRUSTED

5. Robust Models: High Predictive Performance, Intensive Validation
6. Intelligible & Explainable Models
7. **Models with Proven Stability**
8. Decremental Holistic Modeling, Detecting and Handling Weak Signals

REASONING AI

1. 260+ scientist-years of high-level R&D in Fuzzy Mathematics & Automatic Reasoning
2. **Nonlinear Rule-Based Models with Continuous Logics**
3. Collective, Evolutionary, Corrective & Adaptive AI
4. Frugal AI, using evenly small or big data



STRATEGIC

9. General Purpose Behavioral AI
10. Sovereign, 100% Proprietary Fuzzy Algorithms
11. Compliant with International Regulations
12. **Transparent Reverse-Engineering of Opaque/Old Models**
13. Globally Unique

READY-TO-USE

14. **No-Code, No Framework,** All-in-One Solution
15. Scalable Induction Power, No GPU
16. Private PaaS & Secure Access or Client Premises
17. Prediction in real-time at a very high frequency
18. Easy Deployment

HUMAN-CENTRIC

19. Reproduces the 3 Human Reasoning Modes
20. **Highlighting Biases & Producing Ethical Decision Systems**
21. Enhancing Knowledge & Augmenting Experts
22. Audit & Certification before Deployment of Decision Systems

XTRACTIS® Augmented Fuzzy Cognitive AI – 7 bases [Zalila 1993-2025]

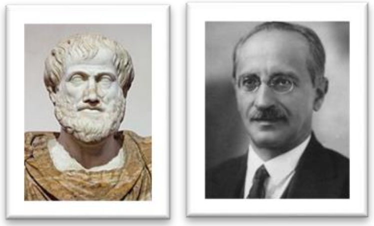
slide 5

1. Continuous/Fuzzy Logics *vs* Binary Logic

slide 35 slide 53

It must capture imprecision, epistemic uncertainty and subjectivity in an infinity way.

➡ Theory of Fuzzy Relations of order N [Zalila 1993]



[Aristotle (-384, -322)]
[Lukasiewicz 1920]

2. Generalized Possibility/Necessity Measures *vs* Probability Measure

slide 48 slide 54

It must handle low quality data, multi-variable dependencies without any a priori definition of probability laws.

➡ Generalized measures of possibility and necessity which are frugal and handle ignorance [Zalila 1993]

[Pascal & de Fermat 1654]

3. Robotic Induction *vs* Human Induction

It must extend human understanding and reveal the behavior of unknown complex processes/phenomena.

➡ Humans pose the complex problem that they will never be able to solve and their exobrain solves the complex problem that it will never be able to pose. [Zalila 2003, 2017]

[al-Hassan ibn al-Hassan ibn al-Haytham al-Basri 1015-1021]
[R. Bacon 1267]

4. Complexity *vs* Complication

slide 47

It must preserve medium and weak signals, and interactions between variables.
It must discover the true level of complexity of PPC.

➡ Fuzzy Decremental Holistic Automatic Induction = Polymorphic modeling
[Zalila & Idagrai Labs/Intellitech 2002-2025]

[Descartes 1637]
[von Bertalanffy 1950, 1951]

idagrai LABS

XTRACTIS®

XTRACTIS® Augmented Fuzzy Cognitive AI – 7 bases [Zalila 1993-2025]

5. Robust Predictive Modeling *vs* Descriptive Modeling

It must ensure high robustness (generalization capacity) and detect errors and biases in the data.

➡ Systematic multi-cycle Cross-Validation

➡ Corrective Induction [Zalila 2024-2025]

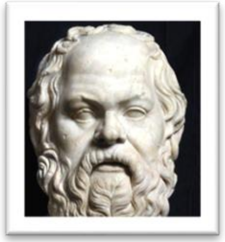
slide 38 slide 39
slide 44
slides 40-43

6. Wisdom *vs* Foolishness

It must know that it does not know = Wise ignorance

➡ Right of refusal [Zalila & Idagrai Labs/Intellitech 2002-2025]

➡ Adaptive Induction



[Socrates (-470, -399)]

slide 56

7. Robot Intelligence Test on Reasoning *vs* R. I. Test on Conversation

An intelligent robot must be able to automatically design and perfect its most efficient **inductive** strategies to discover the most robust & intelligible decision systems, then use them to predict in real time by explaining its deterministic **deductive** reasoning, and discover the most optimal multi-objective prescriptions by automatically perfecting its **abductive** reasoning. [Zalila 2017]

[Turing 1950]

Examples of XTRACTIS® Use Cases – benchmarks vs. LoR, RFo, BT & NN

HEALTH / PHARMA

Anatomopathological Diagnosis of Breast Cancer (2022)

Genetic Diagnosis of Prostate Cancer (2022)

Cardiotocographic Identification of Fetal Heart Conditions (2022)

Spectrometric Diagnosis of Ovarian Cancer (2022)

Serological Diagnosis of Chronic Kidney Disease (2023)

Genetic Identification of Lung Cancer (2022)

BUSINESS / FINANCE

Chemical Identification of Wines' Type (soon)

Discovery of Discriminatory Biases in the Professional Evaluation of Employees (2024)

Nonlinear Multi-Objective Optimization of a Supply Chain Under Flexible Constraints (2023)

Detection of Fraudulent Credit Card Transactions (2023)

Prediction of Telecom Customer Churning (2023)

INDUSTRY / R&D

Prediction of the Degradation of a Naval Propulsion Unit (2022)

Emergency Detection for an Automatic Braking System (2021)

Prediction of the Toxicity of Chemical Molecule Residues & Discovery of New Nontoxic Herbicides (2023)

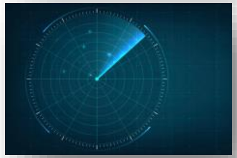
Prediction of the Compressive Strength of Concrete (2023)

Identification of the Longitudinal Action Required when Approaching Traffic Lights (2023)

Prediction of the Rupture of Flexible Underwater Pipe (2023)

DEFENSE / CYBER / SECURITY

Log-based Detection & Identification of Cyber Intrusions (2022)



Acoustic Detection of Underwater Mines (2022)

Identification of an UAV Intrusion Based on Wi-Fi Analysis (2023)

Temporal Identification of Criminal Profiles and Action Phases from Communications Metadata (2023)

Passive Magnetic Identification of Land Mines (2023)

Prediction of Robbery Crimes in American Cities (2023)

xtractis.ai/use-cases

UC#08: Genetic Diagnosis of Prostate Cancer

Design an AI-based decision system which **accurately** and instantly makes a **rational** medical diagnosis of prostate cancer from genetic sequencing of prostate tissue.

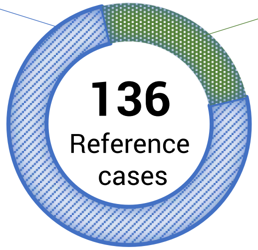
Variable to Predict among 2 classes: Normal | Tumor

12,600 Potential Predictors, which are the level of expression of genes characterizing each patient, normalized to the median: gene_39939_at, gene_33792_at, gene_35178_at…

Source:
D. Singh & al., Department of Adult Oncology, Brigham and Women's Hospital, Harvard Medical School

12,600 numeric variables

75.00%
Training 80%
Validation 20%



25.00%
External Test

102 genetic sequencing of prostate tissue from patients	
Normal	**Tumor**
52 | 50.98%	50 | 49.02%

34 genetic sequencing of prostate tissue from patients	
Normal	**Tumor**
9 | 26.47%	25 | 73.53%

XTRACTIS top-model: Intelligible Decision System

100 + 2,000 inductive reasoning strategies explored, Induction optimization on F_1-Score,
Top-model selection on validation F_1-Score

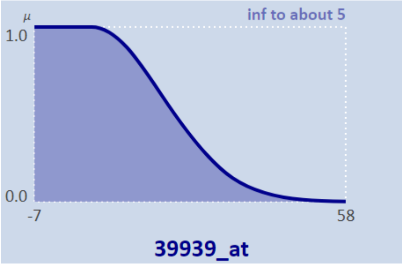
IS = 4.98 **F_1-Score** (on ETD) **= 100.00%**

PREDICTORS

- **7 predictors** selected by XTRACTIS out of 12,600
  (12,600 numeric variables)
- Ranked by impact significance:
  (2 strong signals, 3 medium signals & 2 weak signals)
- Labeled by fuzzy classes

Example:

fuzzy interval “**inferior to about 5**”



RULES

- **4 connective fuzzy rules without chaining**
  (aggregated into 2 disjunctive rules)
- **2 to 4 predictors per rule** (on average, 3.0 per rule)

R1

R2

R3

R4

IF	gene 39939_at	IS	inf to ~5
AND	gene 35178_at	IS	inf to ~-2
AND	gene 36883_at	IS	inf to ~87
AND	gene 40282_s_at	IS	inf to ~77
THEN	Diagnosis	IS	TUMOR

idagrai LABS

XTRACTIS®

Explained Prediction for a Case from the External Test Set

Patient #1

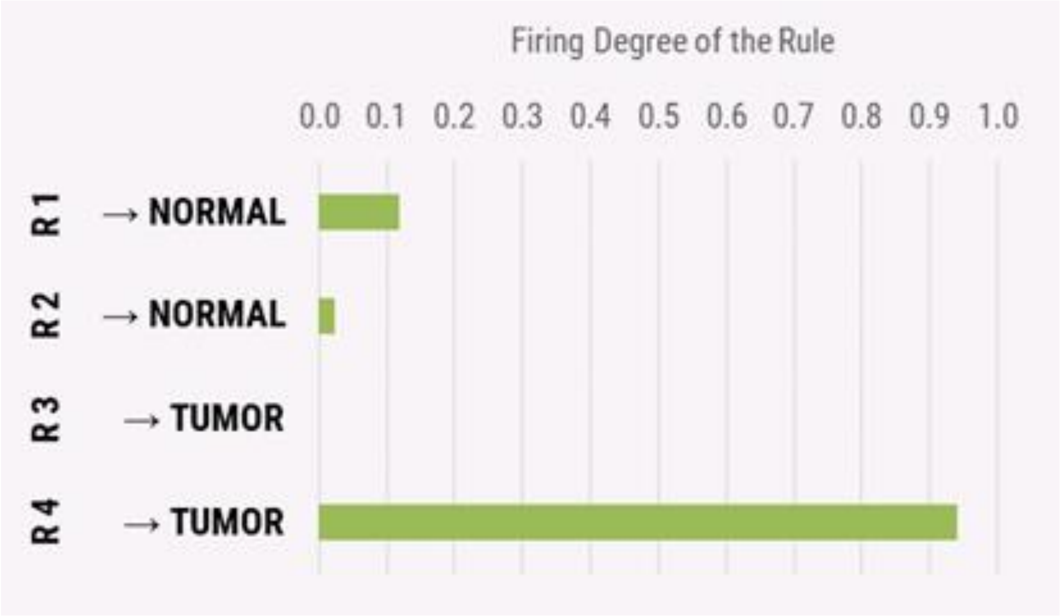
gene 39939_at	5
gene 33792_at *	1.7
gene 35178_at	2
gene 36883_at	39
…	…
gene 40282_s_at	26

actual value = TUMOR

**Predictor value outside the variation range of the model but inside the allowed extrapolation range. XTRACTIS will refuse to give a result for an extrapolation far from the allowed extrapolation range.*

Real Time Deductive Inference of the model's rules

Number of triggered rules: **3 / 4**



FUZZY PREDICTION

{ TUMOR | 0.940,
NORMAL | 0.117 }

FINAL PREDICTION

{ TUMOR }

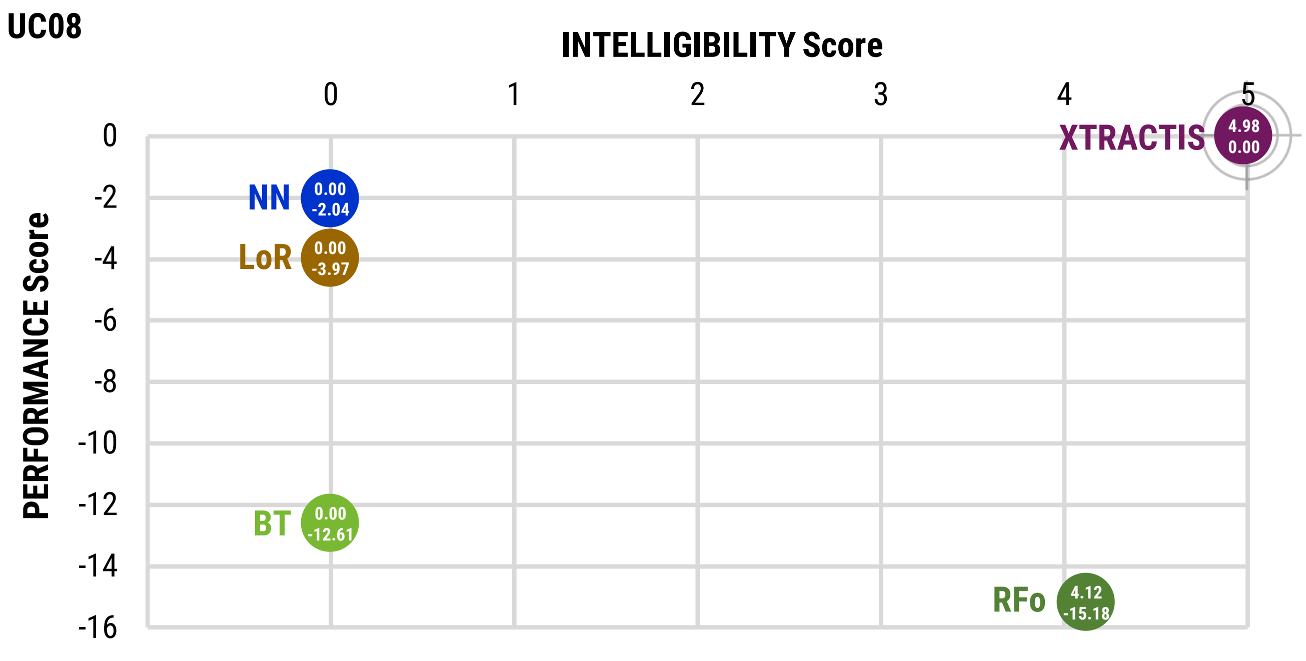
The system delivers a correct diagnosis of cancer compared to that given by the genetic oncologist:

TUMOR

UC#08: Benchmark XTRACTIS vs. its Challengers

XTRACTIS[1]	7 predictors | 4 gradual rules without chaining aggregated into 2 disjunctive rules | Each unitary rule uses only 3.0 predictors on average | A few rules triggered at a time
Logistic Regression[2]	120 predictors | 1 linear equation
Random Forest[2]	19 predictors | 15 trees without chaining | 50 binary rules | Each rule uses only 1.8 predictors on average
Boosted Tree[2]	24 predictors | 1 chain of 14 trees | 48 binary rules | Each rule uses only 1.9 predictors on average | Tree #N corrects the error of the N-1 previous trees
Neural Network[2]	12,600 predictors | 1 hidden layer | 14 equations | 13 hidden nodes | Each equation uses 11,700.9 predictors on average | Unintelligible synthetic variables

F_1-Score calculated on all unknown cases (CVE & IVE).
Leader performance is 100.00%



[1] Results from XTRACTIS® REVEAL v11.2.38531 (2021/06)

[2] Results from XTRACTIS® BENCHMARK module embedding Python 3.6 package (2021/04): Scikit-Learn 1.0.2 | LightGBM 2.2.2 | TensorFlow 2.6.2 | Keras 2.6.0

XTRACTIS® vs. its 4 challengers in HEALTH/PHARMA

Scores from 9 Public Use Cases

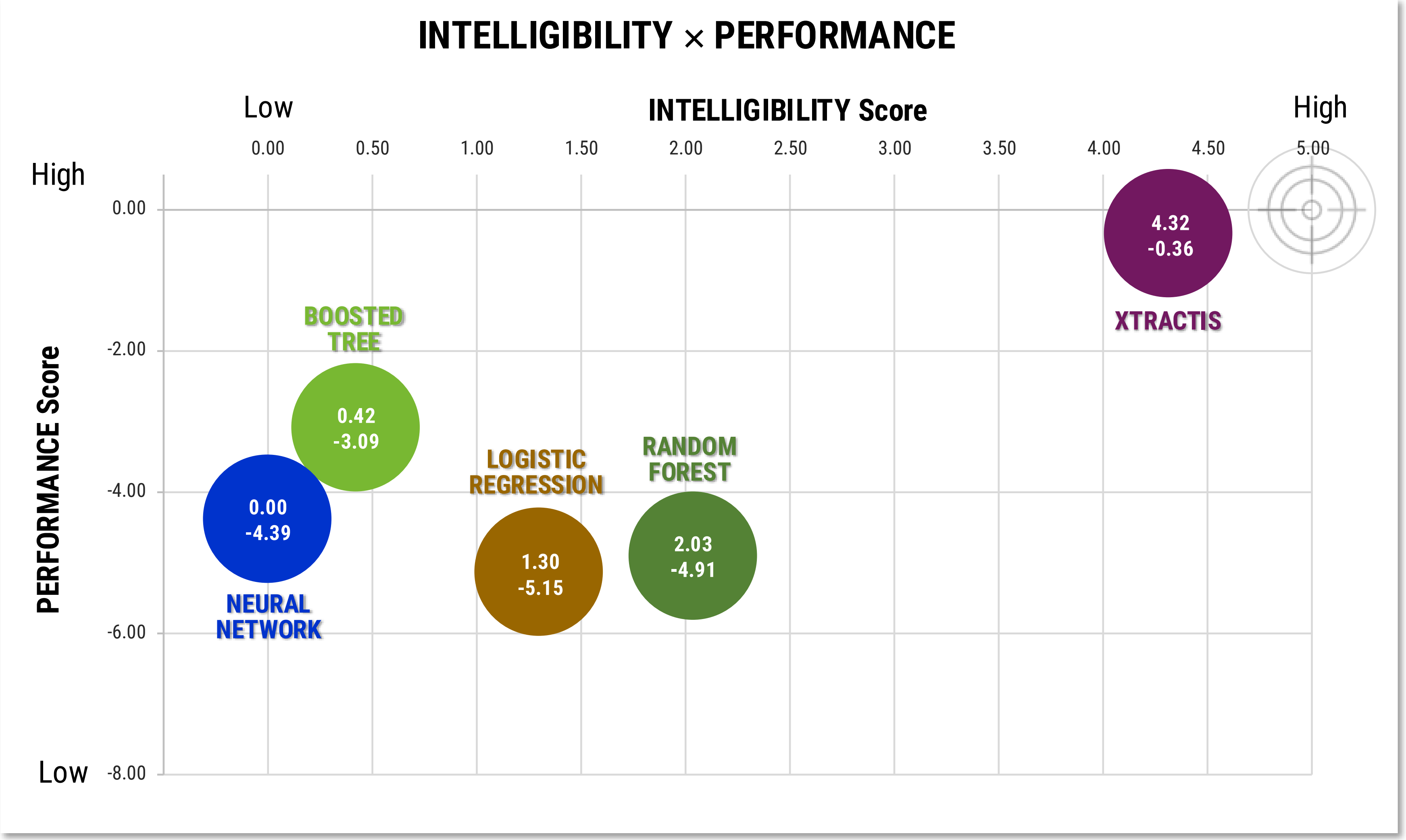
A bubble on the **top-right** is the Holy Grail for critical AI-based decision systems:

an AI Technique which produces predictive models with

the highest Performance

and

the highest Intelligibility



idagrai LABS

XTRACTIS®

XTRACTIS® vs. its 4 challengers in all sectors

Scores from 27 Public Use Cases

Logistic Regression excluded: inapplicable for 3 UC, unavailable for 1 UC

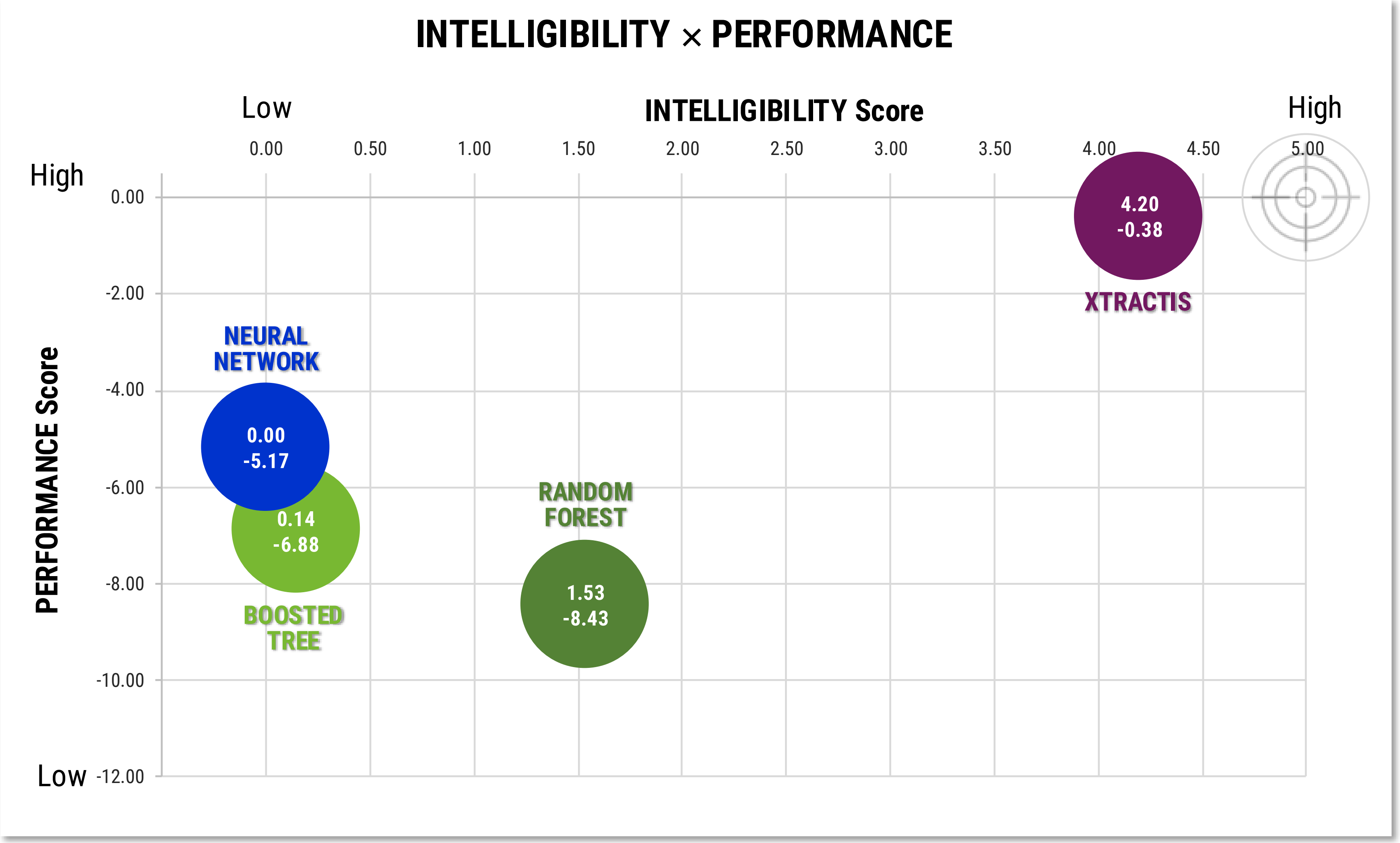
A bubble on the **top-right** is the Holy Grail for critical AI-based decision systems:

an AI Technique which produces predictive models with

the highest Performance

<u>and</u>

the highest Intelligibility



XTRACTIS® vs. its 4 challengers in all sectors

Scores from 23 Public Use Cases* out of 27

**for which Logistic Regression is applicable*

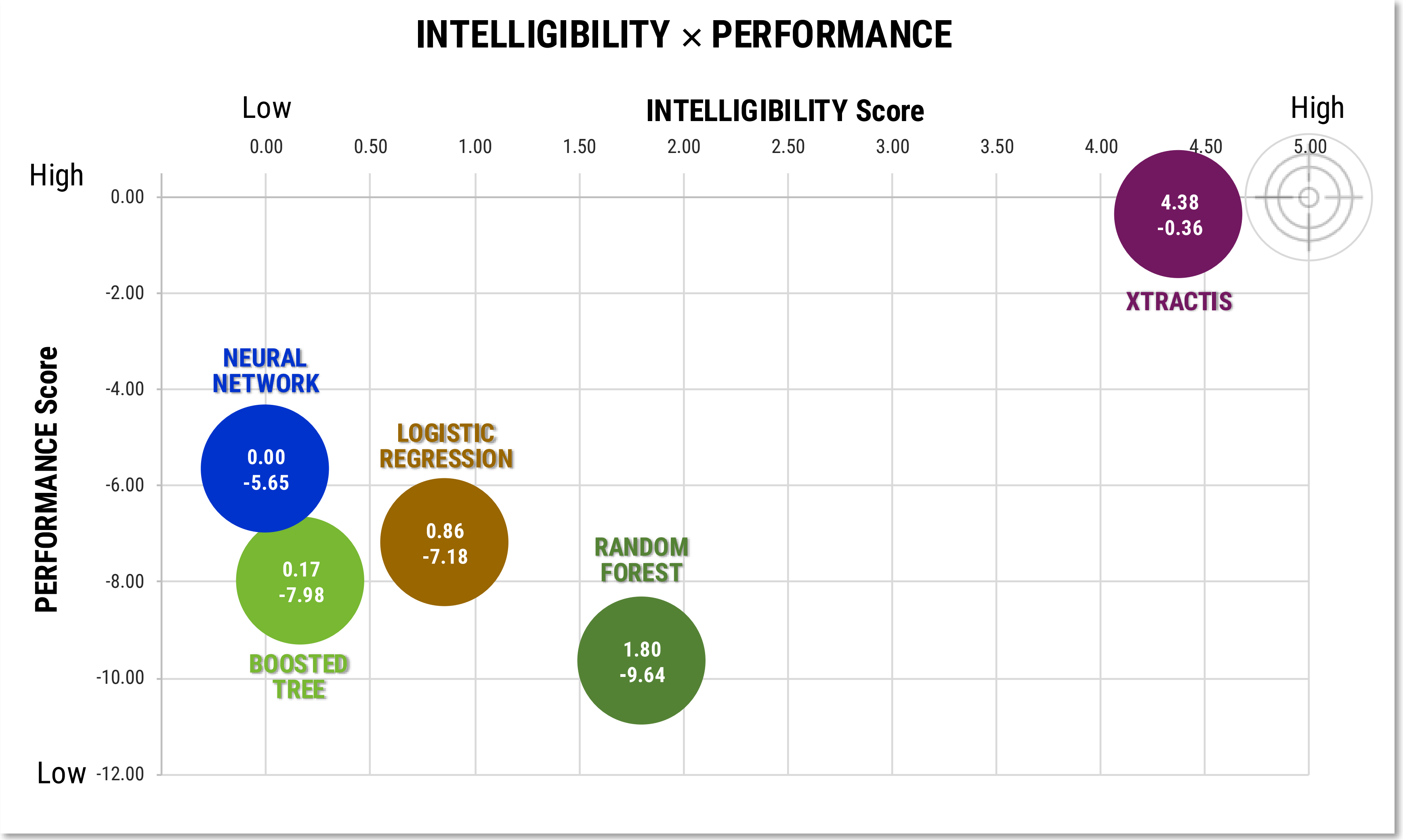
A bubble on the **top-right** is the Holy Grail for critical AI-based decision systems:

an AI Technique which produces predictive models with

the highest Performance

<u>and</u>

the highest Intelligibility



Complex Processes – 40 properties for a Trusted General Reasoning AI [Zalila 2013-2025]

+ **Multidimensional, Handling dependences, General Purpose**
+ **Frugal, Handling data with low quality (Fuzziness), heterogeneous data**
+ **Handling Regression, Multiclassification, Scoring, Clustering issues**
+ **Automatic, Without a priori, Rational, Collective & Evolutionary Inductive Reasoning**
+ **Nonlinear, Non-monotonic, Deterministic, Polymorphic, Deductive Reasoning**
+ **Intelligible/Transparent, Traceable, Explainable, Local, Granular** slide 52
+ **Decremental Holistic, Handling low signals, Handling ignorance, Right of refusal**
+ **Reflective, Robust/Accurate, Formally Stable, Ethical**
+ **Handling ill-posed problems, Group decision**
+ **Collective & Evolutionary Abductive Reasoning, Multi-objective Optimization under flexible constraints**
+ **Corrective Induction to detect and correct errors and biases in the reference data**
+ **Adaptive Induction to update the discovered knowledge** slide 10

XTRACTIS® Augmented Fuzzy Cognitive AI – the best of 3 worlds

	UNINTELLIGIBLE (Opaque)	INTELLIGIBLE (Transparent)
Automatic Process **ROBUST**	**Con-AI** **Machine Learning** **Deep Learning** **Graphs** Neural Network Models	**AFC-AI** **Induction (=KDD)** **Deduction** **Abduction** **FR-*N*** Fuzzy Predictive Models
Manual Process **NON-ROBUST**	X	**BC-AI / FC-AI** **Maieutics** **Deduction** **Abduction** **Rules** Binary Expert Systems Fuzzy Expert Systems
		ESM **Induction** **Deduction** **Equations** Analytical and Statistical Models

***Con-AI**: Connectionnist AI*
***BC-AI**: Binary Cognitive AI*
***FC-AI**: Fuzzy Cognitive AI*
***ESM**: Experimental Scientific Method*
***AFC-AI**: Augmented Fuzzy Cognitive AI*

Augmented Fuzzy Cognitive AI

∝-valent, ∝- relational, ∝-operable, ∝-measurable, ordinal, non-additive, knowledge-based AI
with Decremental Holistic, Inductive, Abductive, Deductive, Reflexive, Collective, Competitive, Cooperative, Evolutionary, Corrective and Adaptive Reasoning.

➡ **Cumulative Social Learning** trough emulation (innovation) and imitation (conformity)
= intergenerational transmission and improvement of meta-knowledge and of knowledge.

➡ **Robust Universal Cognitive Approximator**
of non-linear, non-monotonic multidimensional functions, and of non-convex, disconnected, non-decomposable sets.

idagrai LABS

XTRACTIS®

BENEFITS OF XTRACTIS AI

Since 2003, the only one robust **and** intelligible decision-making AI

- capable of detecting **biases** and **errors** in data and decisions
- capable of **auditing** human or virtual experts to detect **fraud** and errors (cf. Aristotle's Organon)
- capable of revealing, in an intelligible manner, the **unconscious decision-making strategies** of an animal, human or artificial brain
- capable of producing **trusted critical high-risk decision-making systems**
- capable of discovering **new robust scientific knowledge**
- **sovereign** & ITAR-free that does not use any open-source AI code
- using only **full CPU, at 100% of its FP64 speed, 100% of the time**, without any GPU
- that consumes **10^3 to 10^4 times less** energy in model training and **10^5 to 10^6 times less** in deductive inference than LLMs
- that successfully passes new **robot intelligence test**, proposed by [Zalila 2017]

IMPACTS OF USE OF TRUSTED PREDICTIVE MODELS

- Autonomous development of trusted models, for personalized decisions, from authorized and controlled data
- **Industry**. Greater production of IP assets, discovery of new products adapted to each (group of) consumer(s)
- **Healthcare**. More accurate and informed medical decisions: ↘ deaths, ↘ medical costs, ↘ medical deserts
- **Education**. 24/7/365 accelerated and informed training of students thanks to virtual coaches, justifying their decision in a deterministic and rational manner
- **Governance**. More fair and informed political decisions: optimization of state budgets,↘ social movements, ↗ social welfare

General Reasoning AI – Synthesis

- **Defense**. Deterrent weapon (swarms of intelligent drones): ↘ conflicts, ↘ collateral damage, no destruction of infrastructure. AI vs. AI battle.

- **Security**. Accelerated and informed fight against malicious acts: ↘ fraud, ↘ money laundering, ↘ terrorist actions, ↘ cyberattacks, ↘ repeat offenses, ↘ femicides

- **Social Justice**. Accelerated and informed fight against discrimination (hiring, promotion, education admissions algorithms, credit, insurance): ↗ inclusion, ↘ unemployment, ↘ unnecessary training expenses

- **Sustainability**. ↘ energy and hydraulic consumption (frugal AI without GPU, small CPUs for very high frequency deduction)

↗ Intelligence **of Humanity**

The AI of **explicit knowledge** *vs.* the AI of **belief**

XTRACTIS®

The General Reasoning AI for Trusted Decisions

idagrai LABS

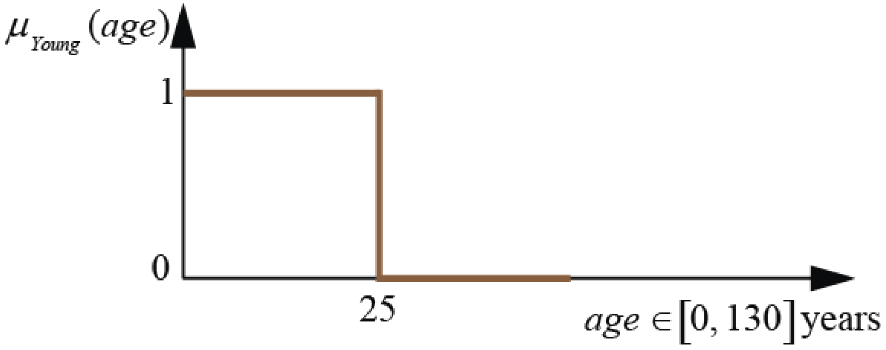
Prof. Dr. Zyed ZALILA
President — Founder
@zyedzalila
+33 6 21 02 03 37
zyed.zalila@xtractis.ai

Complex Processes & Phenomena

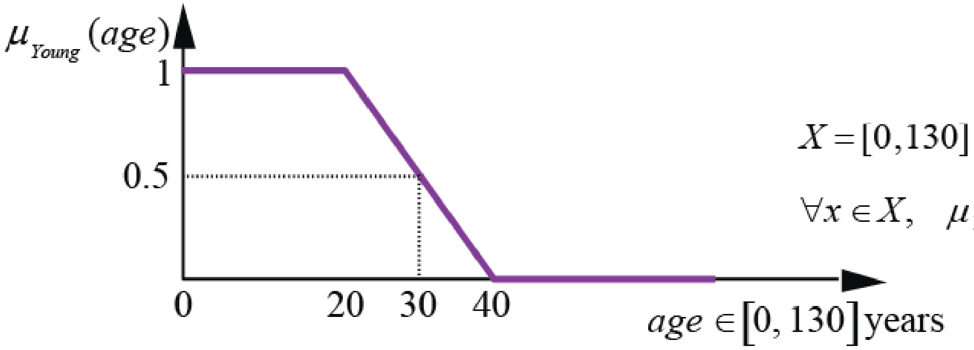
- **High dimensionality** (hundreds / thousands / tens of thousands of variables)
- Strong interactions between variables, Weak signals
- Non-linear, Non-monotonic, Non-convex, Disconnected, Non-decomposable
- Tacit / implicit knowledge (intuition, instinct, perception)
- Heterogeneous reference data
- Low quality of reference data
  - noisy
  - incomplete (missing values)
  - medium & weak signals
  - few quantity
  - fuzzy (imprecise, uncertain, subjective)

Examples of CPP:
- Predictive maintenance of frigate turbines
- Modeling of engine tests
- Prediction of the ecotoxicity of chemical molecules
- Epigenetic & metabolic predictive medicine
- Behavioral finance…

Fuzzy Mathematics – suitable for real world

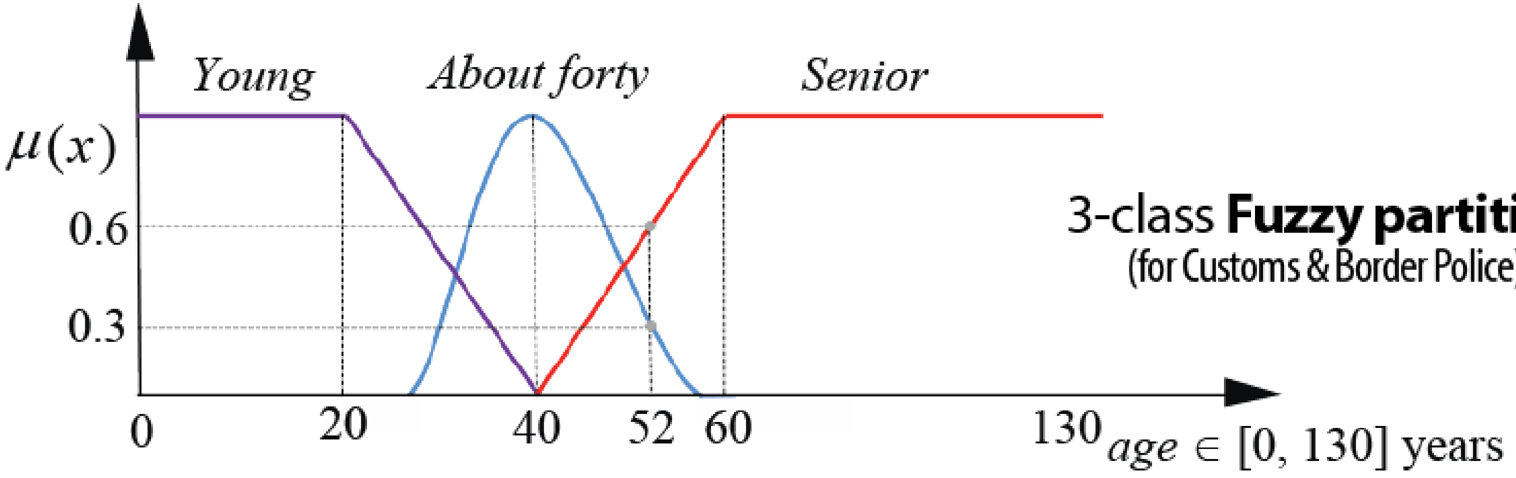


Binary set (BR-1): **Young traveler** (for airlines)



$$X = [0,130] \text{ years};\quad \forall x \in X,\quad \mu_{Young}(x) = \begin{cases} 1, & \text{if } x \in [0,20] \\ \dfrac{40-x}{20}, & \text{if } x \in [20,40] \\ 0, & \text{otherwise} \end{cases}$$

Fuzzy set (FR-1): **Young traveler** (for Customs & Border Police)



3-class **Fuzzy partition**
(for Customs & Border Police)

Lemma [Zalila 1993]

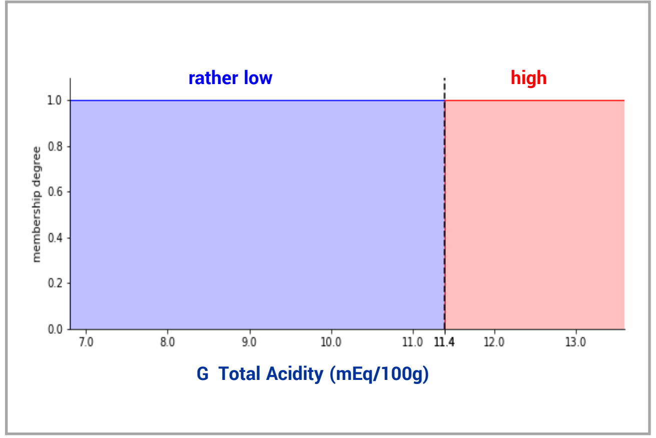
Let Q be an FR-n of X of height h. Let $\top$ be a T-norm.

$$Q = \bigcup_{\substack{\vee \\ \alpha \in [0\,,\,h]}} \alpha \langle \top \rangle Q_\alpha$$

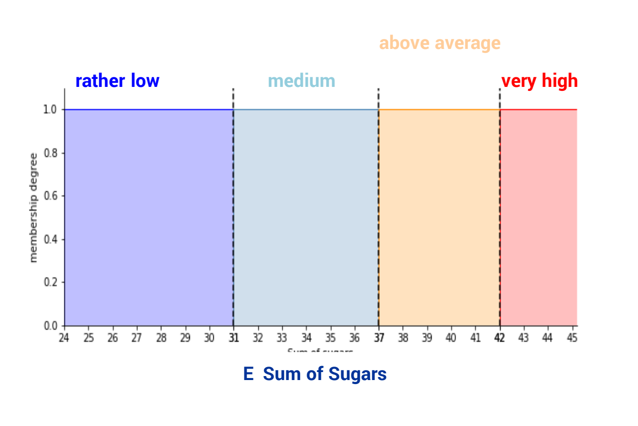
Binary rule-based model – Intelligible model, but not efficient

Sweet perception of a fresh tomato: 2 variables , 6 classes, 4 rules

>> Classes



predictor #1: *Total Acidity*



predictor #2: *Sum of Sugars*

>> Rules

Rule ❶

If *Total Acidity* is *rather low*
And *Sum of Sugars* is *rather low*
Then *Sweet* equals *3.39*

Rule ❷

If *Total Acidity* is *rather low*
And *Sum of Sugars* is *medium*
Then *Sweet* equals *7.19*

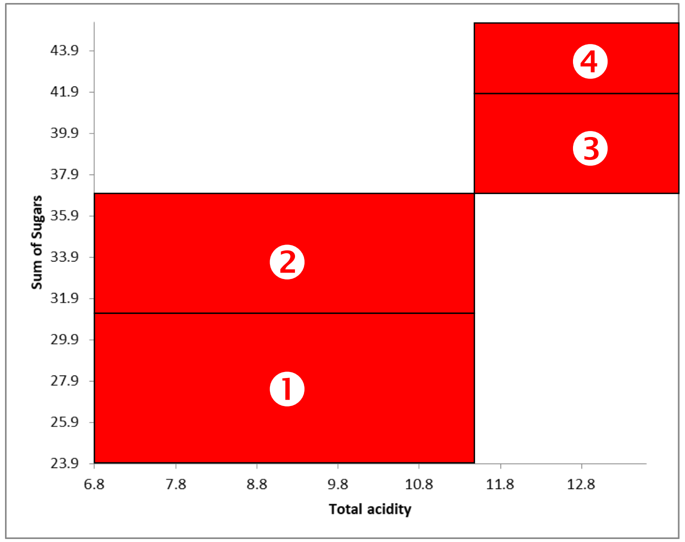
Rule ❸

If *Total Acidity* is *high*
And *Sum of Sugars* is *above average*
Then *Sweet* equals *3.30*

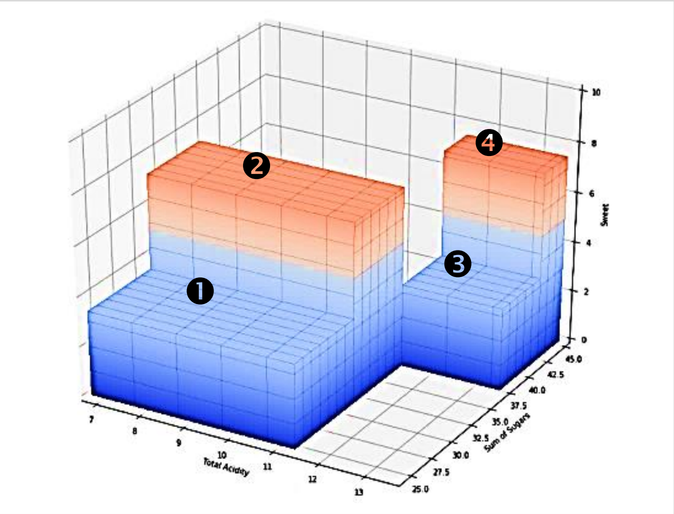
Rule ❹

If *Total Acidity* is *high*
And *Sum of Sugars* is *very high*
Then *Sweet* equals *7.49*

>> Inference



Mapping



Decision Surface

Data Source:
INRA (Institut National de la Recherche Agronomique) et CTIFL (Centre Technique Interprofessionnel des Fruits et Légumes) - 7th Sensometrics Conf., July 2004, Davis, CA, USA

Scoring
Binomial Classification
Multinomial Classification
Regression

idagrai LABS

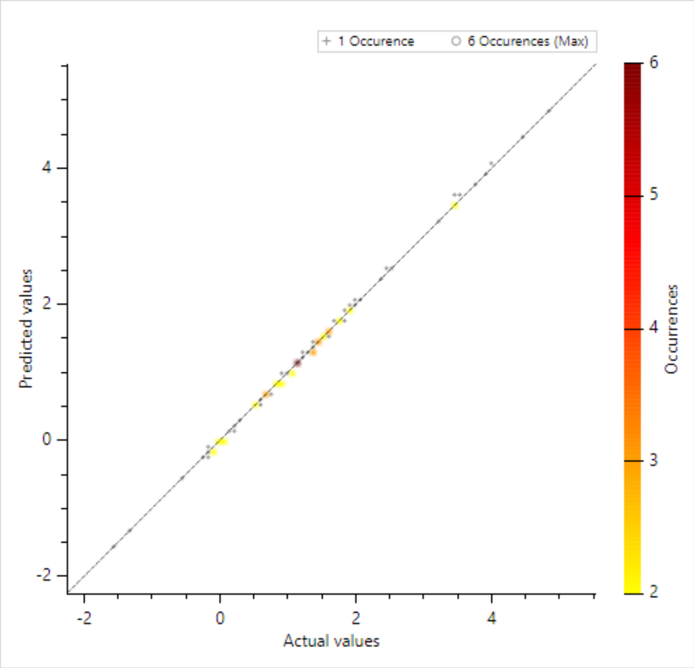
XTRACTIS®

Specificities of XTRACTIS® – Robustness through validation

UPD
2024.12

OVERFITTING PITFALL

⊙ Descriptive model **seems** perfect (regarding predictions on Training points),
but is **not necessarily** robust (as good at predicting External Test points as Training points)



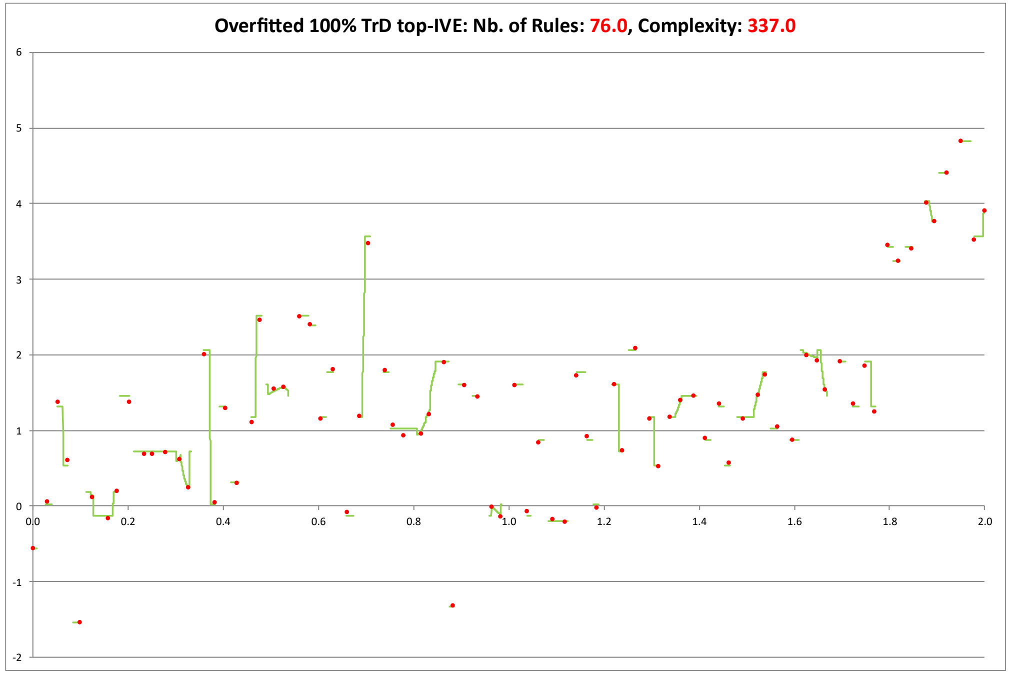
Descriptive top-IVE model
1 Variable, 76 Rules

Descriptiveness (TrD 100%: 80 known cases)
RMSE = **0.034** (**0.45%**)
Correlation = **1.000**
Refusal = 0.00%

Predicted Variable value = f(Predictor value)
- Red dots: 80 known cases of the Training Dataset TrD
- Green curve: model predictions over 20,000 unknown cases of the External Test Dataset ETD, distributed over the entire Definition Range

Discontinuities of the green curve represent the model's Refusals
(36.86% of Refusals on the 20,000 points of ETD)

A regression model without Refusal would be continuous



OVERFITTING INDICATORS

⊙ **Very High** number of Rules, compared to the number of known points
(**76 Rules** *vs* **80** Training Points)

⊙ **Very High** number of Refusals on unknown points
(**36.86% of Refusals** on ETD *vs* **0.00%** on TrD)

Scoring
Binomial Classification
Multinomial Classification
Regression

Results from XTRACTIS® REVEAL 13.2.54171

idagrai LABS

XTRACTIS®

Specificities of XTRACTIS® – Robustness through validation

UC#24 – Prediction of Von Mises Maximal Stress

Data: GSM, MOPS, Université de Technologie de Compiègne

- ⦿ Descriptive model is **not necessarily** robust
- ☑ Robust model: difficult to obtain but **mandatory**

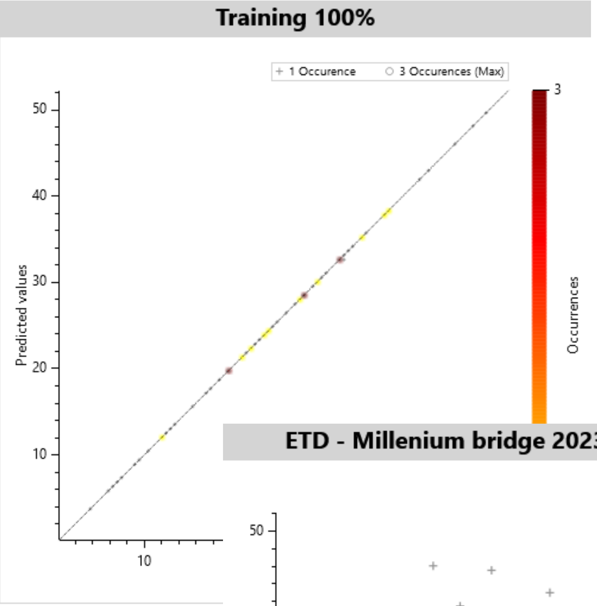
UPD 2024.10

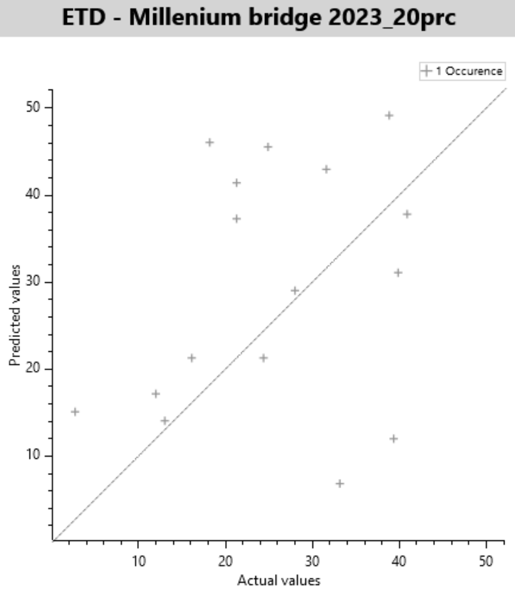
9 variables, 61 rules (Complexity 1,750; Intelligibility 3.62)

Descriptiveness (TrD 100%: 64 known cases)

RMSE = **0.036** (**0.07%**)

Correlation = **1.000**





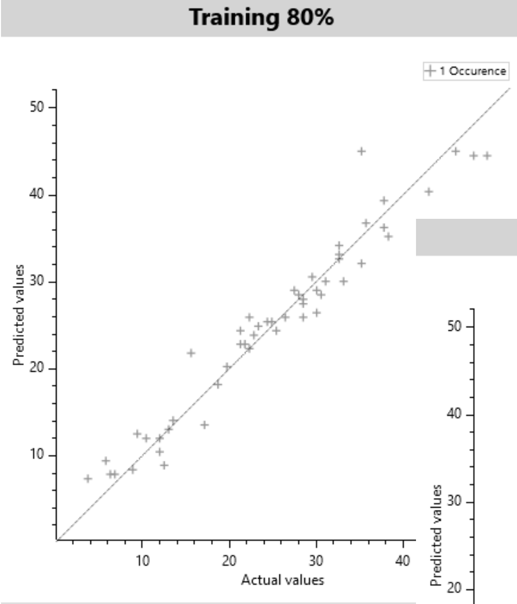
Real Performance

(External Test: 16 unknown cases)

RMSE = **15.258** (**29.87%**)

Correlation = **0.271**

Trivial result: descriptive model, but non-robust

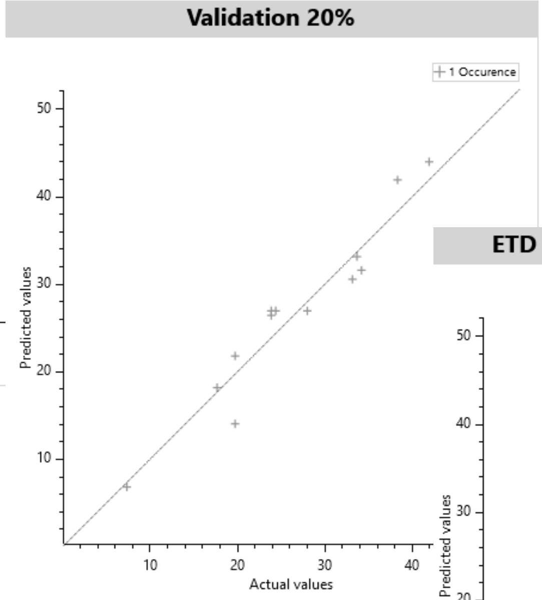


7 variables, 6 rules (Complexity 201; Intelligibility 4.24)

Descriptiveness (TrD 80%: 51 known cases)

RMSE = **2.568** (**5.03%**)

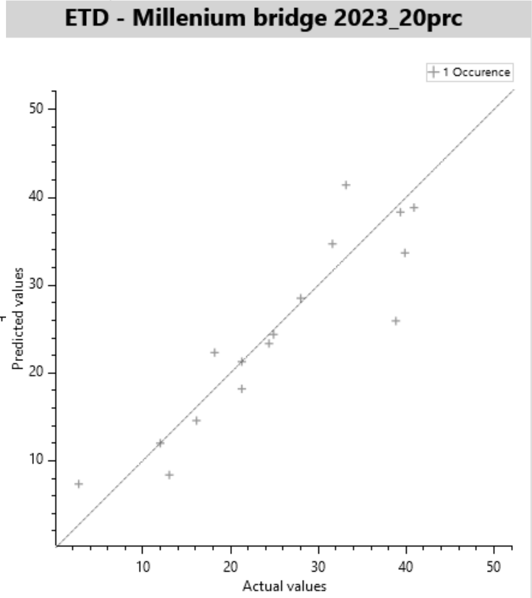
Correlation = **0.974**



Robustness (VD 20%: 13 unknown cases)

RMSE = **2.557** (**5.00%**)

Correlation = **0.966**



Real Performance

(External Test: 16 unknown cases)

RMSE = **4.653** (**9.11%**)

Correlation = **0.909**

Use of robustness estimators: less descriptive model, but robust

Results from XTRACTIS® REVEAL 13.2.53357

Scoring | Binomial Classification | Multinomial Classification | Regression

idagrai LABS

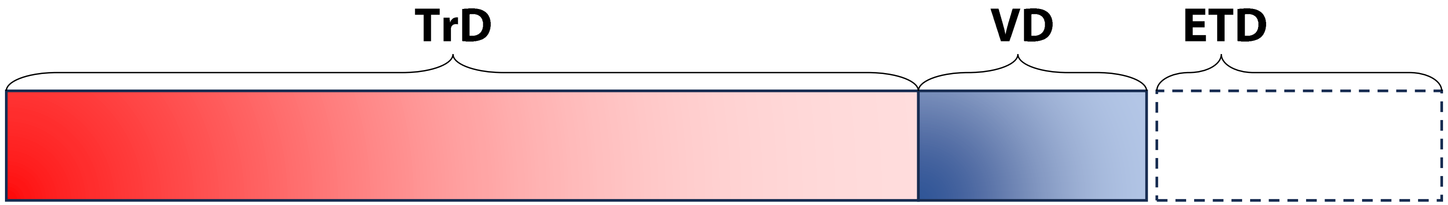
XTRACTIS®

Specificities of XTRACTIS® – Robustness through validation

DATASET PARTITIONING

- **Reference Dataset RD**: All available reference cases, representative of the process to be modeled. ➡ **RD = LD ∪ ETD**

- **External Test Dataset ETD**: Unknown reference cases of the robot, and used to evaluate the real performance of the model created from TrD and to compare the real performance of models created by competing AI techniques.
- **Learning Dataset LD**: Reference cases available after possible setting aside of points kept in ETD. ➡ **LD = TrD ∪ VD**



- **Training Dataset TrD**: Known reference cases of the robot, and used to create the model and evaluate its descriptive performance
- **Validation Dataset VD**: Unknown reference cases of the robot, and used to evaluate the predictive performance of the model created from TrD

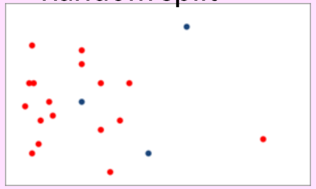
Robustness – Training/Validation with 1 split (if more than ~m × 10^3 learning points)

1 IVE induced by the reasoning strategy S

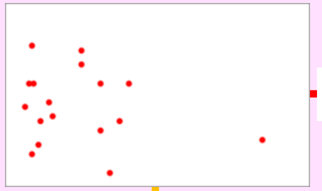
Induction
Training Predictions
Validation Predictions

Random split

LD = random split of
L Learning points



L × (100-z)%
Training points



IVE M_1

L × (100-z)% predictions

Induction of the IVE M_1 by applying
strategy S to L × (100-z)% Training points
from random split

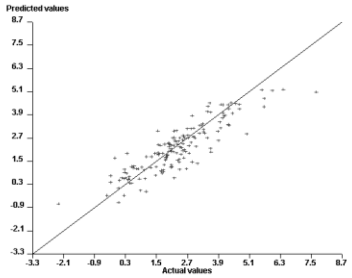
IVE M_1

L × z%
Validation points
(unknown situations)

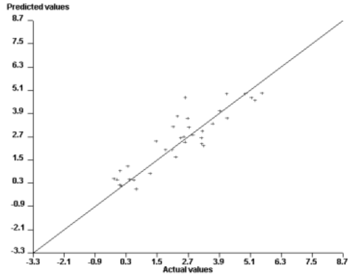
IVE M_1

L × z% predictions

Descriptiveness (Training performance) =
predictions made by M_1 on
L × (100-z)% Training points



Robustness (Validation performance) =
predictions made by M_1 on
L × z% Validation points



Scoring
Binomial Classification
Multinomial Classification
Regression

Robustness – Training/Validation with N splits: Monte-Carlo (if less than ~m × 10^3 learning points)

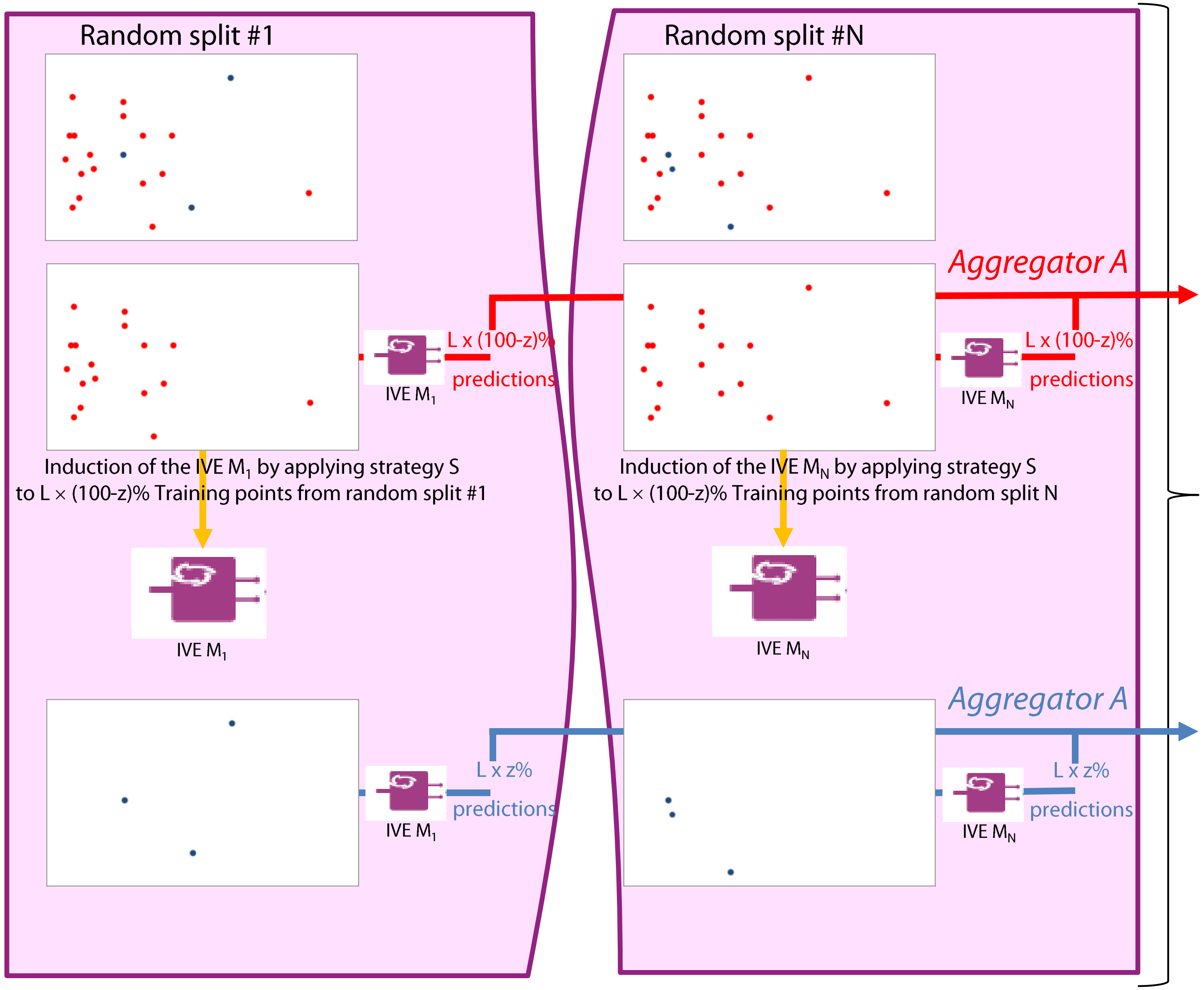
CVE2: college of N IVEs induced by the reasoning strategy S and aggregation operator A

Induction
Training Predictions
Validation Predictions

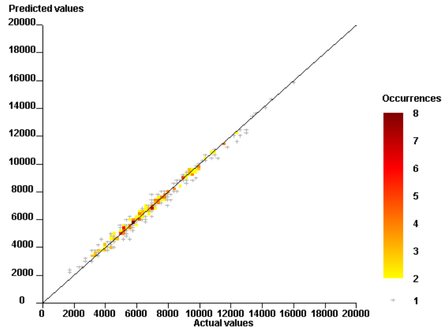
N random splits of L Learning points

L x (100-z)% Training points

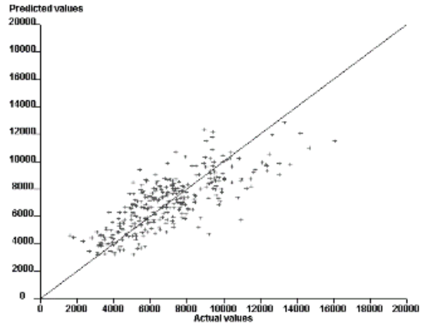
L x z% Validation points (unknown situations)



Descriptiveness (Training performance) = aggregation of predictions made by CVE2 on the L × (100-z)% Training points from each of the N IVEs



Robustness (Validation performance) = aggregation of predictions made by CVE2 on the L × z% Validation points from each of the N IVEs

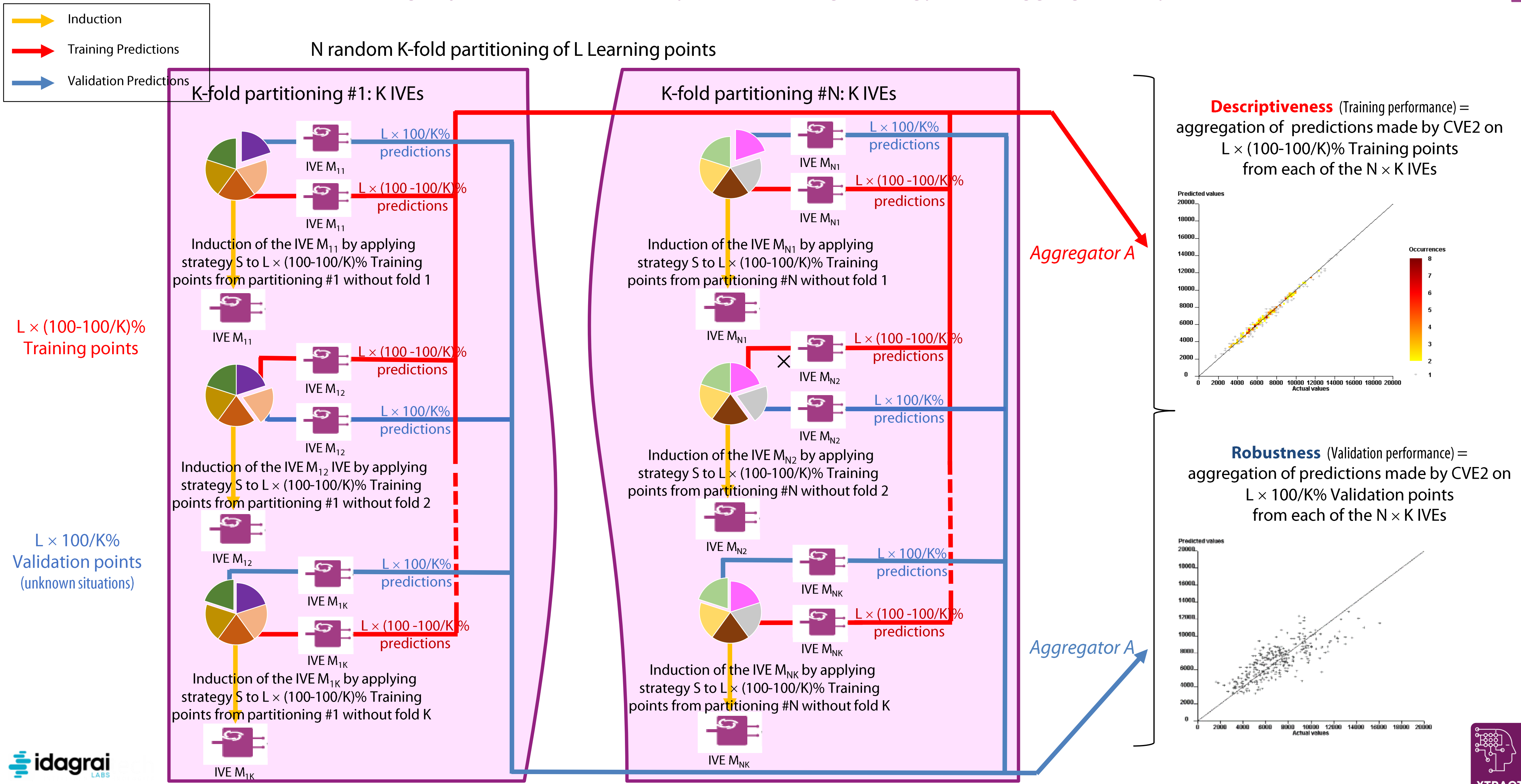


idagrai LABS

XTRACTIS

Robustness – Training/Validation with N × K splits: **N × K folds** (if less than ~m × 10³ leaning points)

CVE2: college of N × K IVEs induced by the reasoning strategy S and aggregation operator A



Scoring

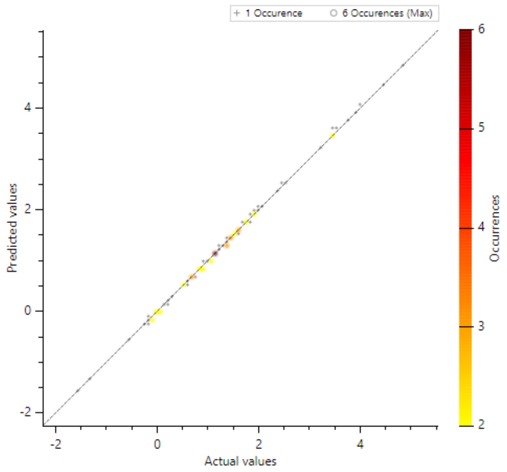
Binomial Classification

Multinomial Classification

Regression

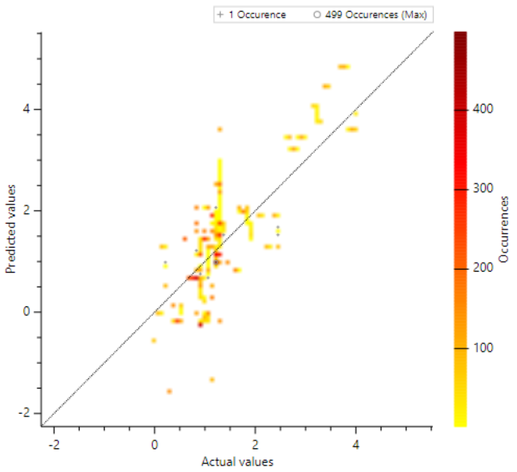
Corrective Induction – Noise detection

- 80 Learning Points
- Hidden Law: $y = \sin(\pi x) + x^2$
- xtractis®



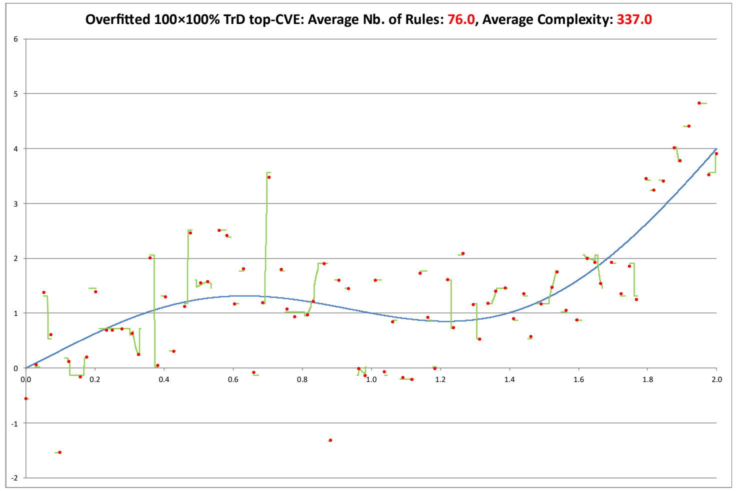
Training 100 × 100%

Descriptive Performance:
r = **1.000**. RMSE = 0.034 (**0.45%**)



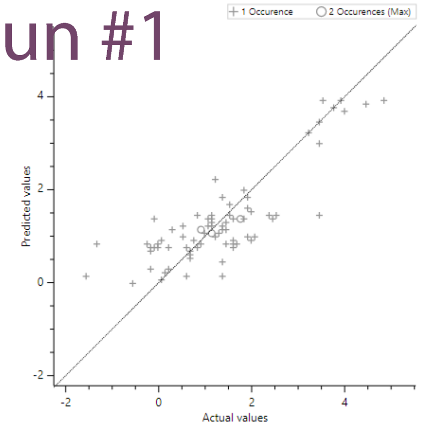
20,000 pts from Noiseless Hidden Law

Real Performance:
r = **0.793**. RMSE = 0.736 (**9.63%**)



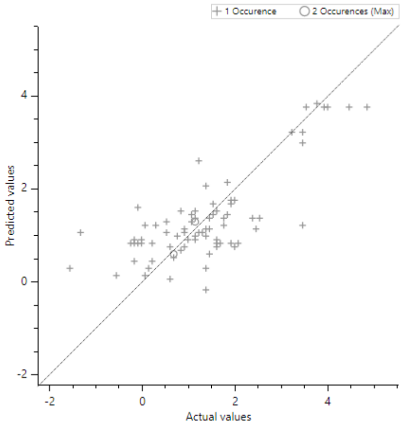
Gaussian noise, σ=20% × width([0;4])
RMSE (Learning Points / Hidden Values) = 0.773 (12.88%)
Correlation (Learning Points / Hidden Values) = 0.782

Run #1



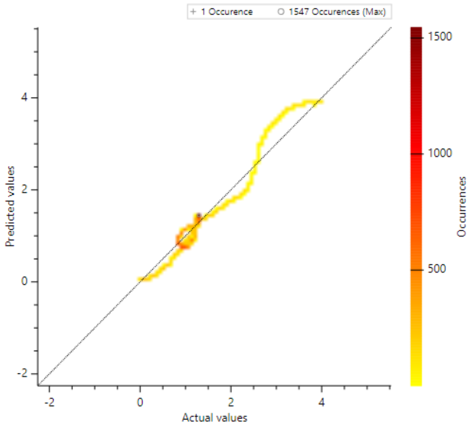
Training 20 × 5 Folds

Descriptive Performance:
r = 0.830. RMSE = 0.688 (9.00%)



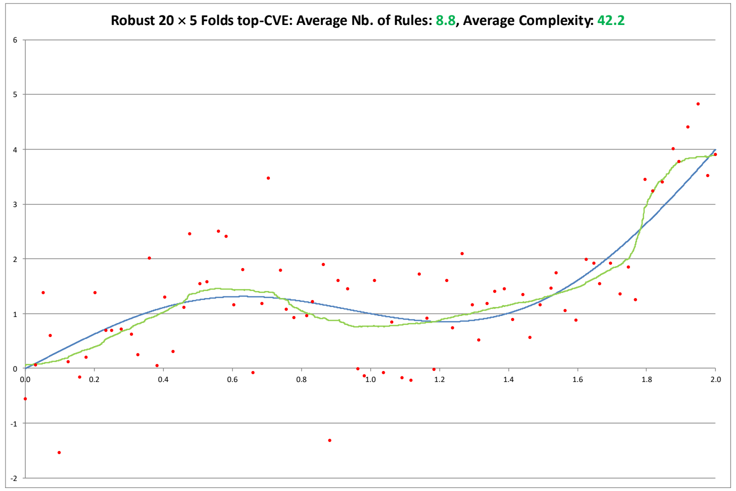
Validation 20 × 5 Folds

Predictive Performance:
r = 0.771. RMSE = 0.781 (10.21%)



20,000 pts from Noiseless Hidden Law

Real Performance:
r = **0.981**. RMSE = 0.199 (**2.60%**)



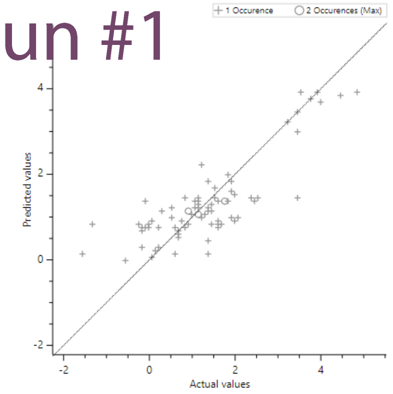
Modeling with noisy data
(**without** and **with** robustness analysis)

Results from XTRACTIS® REVEAL 13.2.54199

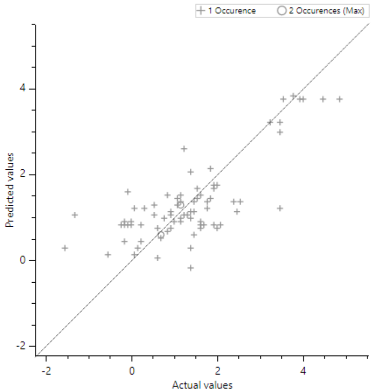
Corrective Induction – Dataset retro-improvements: round 1

80 Learning Points ▪ 23 corrected Learning Points
Hidden Law: $y = \sin(\pi x) + x^2$
xtractis®

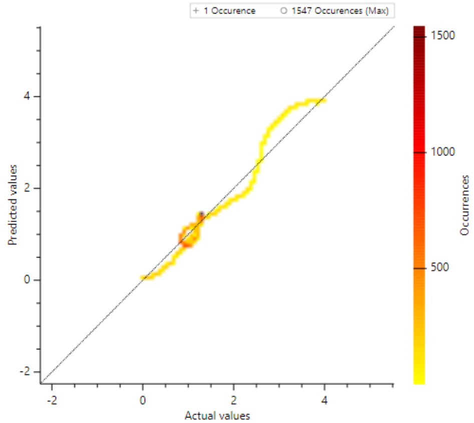
Run #1



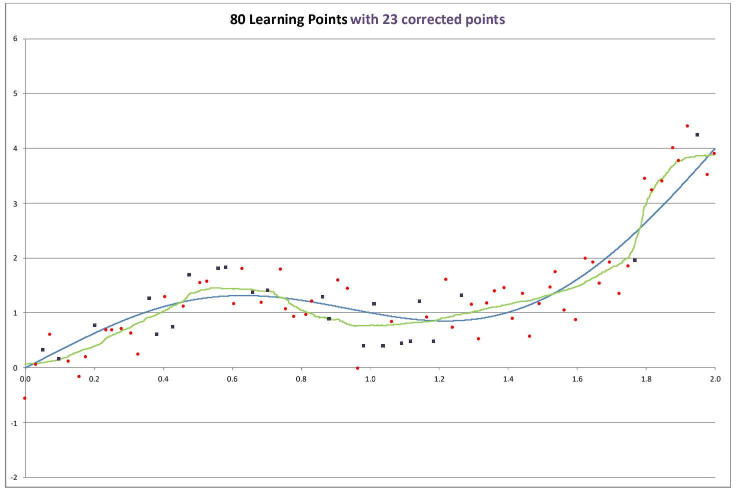
Training 20 × 5 Folds
Descriptive Performance:
r = 0.830. RMSE = 0.688 (9.00%)



Validation 20 × 5 Folds
Predictive Performance:
r = 0.771. RMSE = 0.781 (10.21%)



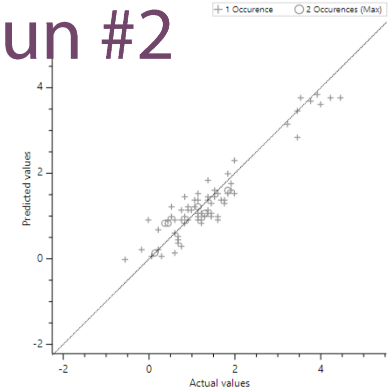
20,000 pts from Noiseless Hidden Law
Real Performance:
r = **0.981**. RMSE = 0.199 (**2.60%**)



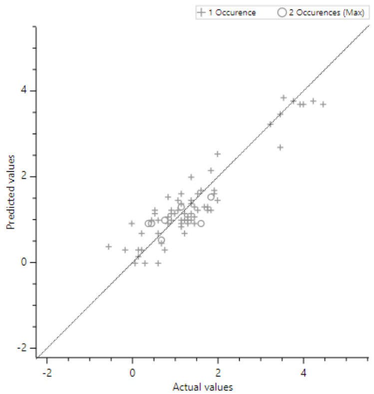
Noise estimation, after 1 round of corrective induction

RMSE (Learning Points / Hidden Values) = ~~0.773 (12.88%)~~ 0.428 (7.13%)
Correlation (Learning Points / Hidden Values) = ~~0.782~~ 0.915

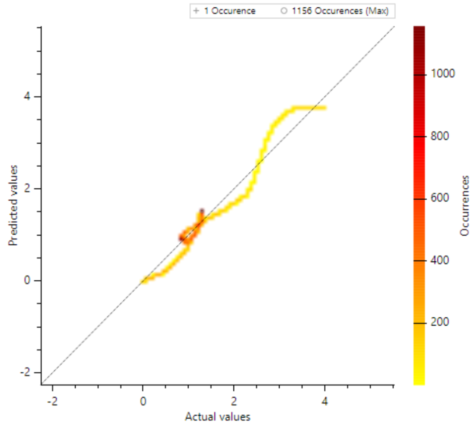
Run #2



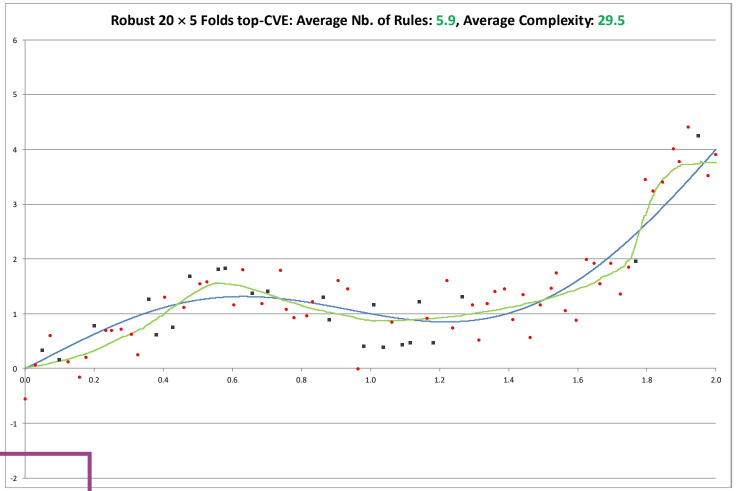
Training 20 × 5 Folds
Descriptive Performance:
r = **0.945**. RMSE = 0.344 (4.50%)



Validation 20 × 5 Folds
Predictive Performance:
r = **0.921**. RMSE = 0.402 (5.27%)



20,000 pts from Noiseless Hidden Law
Real Performance:
r = **0.979**. RMSE = 0.197 (**2.58%**)



- Learning Points (after Corrective Induction) are **closer** to unknown points (RMSE : 7.13% vs 12.88%)
- Model is **more Compact** (Rules/IVE: 5.9 vs 8.8) and slightly **more Efficient** on unknown points (RMSE : 2.58% vs 2.60%)

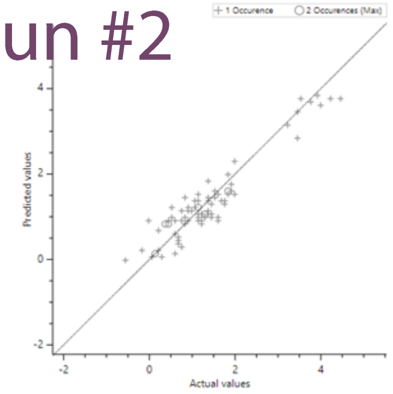
Modeling with noisy data
(**before** and **after 1 round** of corrective induction)

Results from XTRACTIS® REVEAL 13.2.54199

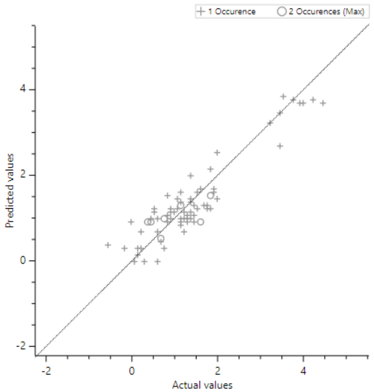
Scoring
Classification Binomiale
Classification Multinomiale
Régression

Corrective Induction – Dataset retro-improvements: round 2

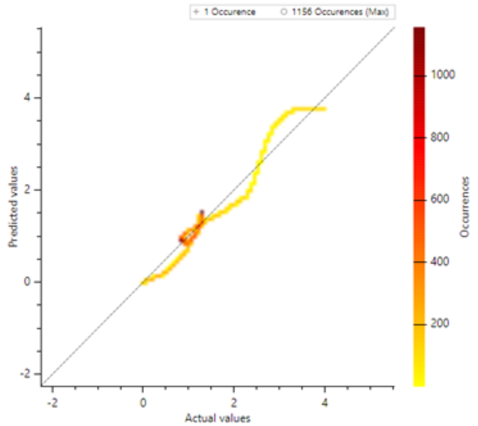
Run #2



Training 20 × 5 Folds
Descriptive Performance:
r = **0.945**. RMSE = 0.344 (4.50%)

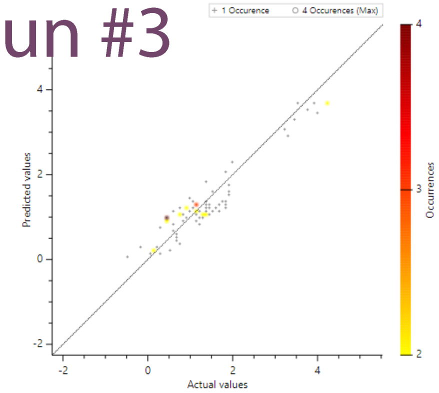


Validation 20 × 5 Folds
Predictive Performance:
r = **0.921**. RMSE = 0. 402 (5.27%)

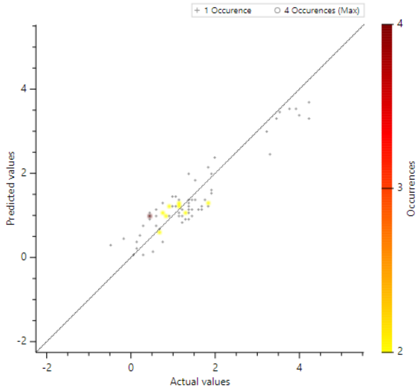


20,000 pts from Noiseless Hidden Law
Real Performance:
r = **0.979**. RMSE = 0.197 (**2.58%**)

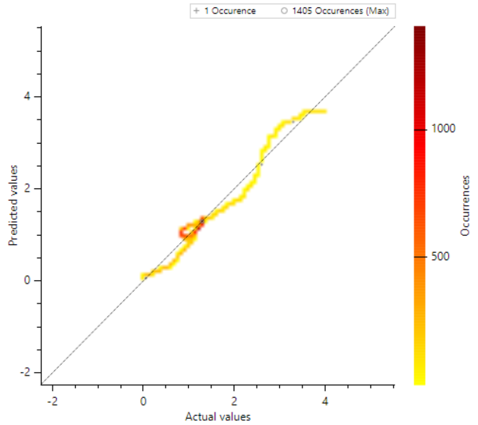
Run #3



Training 20 × 5 Folds
Descriptive Performance:
r = **0.949**. RMSE = 0.327 (4.28%)



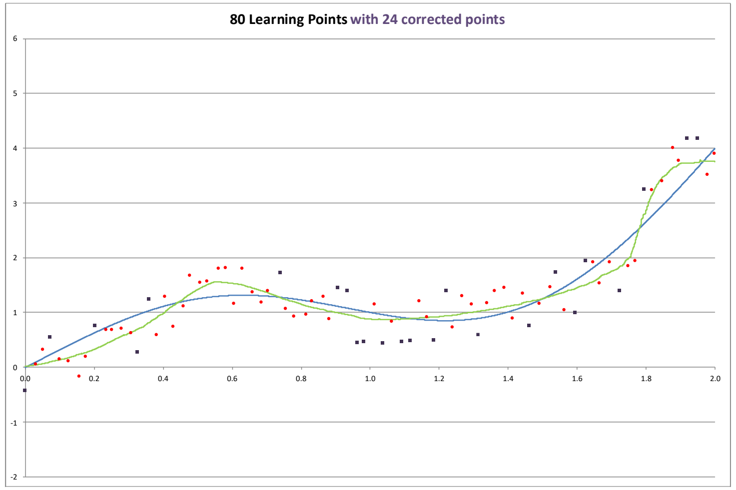
Validation 20 × 5 Folds
Predictive Performance:
r = **0.931**. RMSE = 0.381 (4.99%)



20,000 pts from Noiseless Hidden Law
Real Performance:
r = **0.982**. RMSE = 0.161 (**2.11%**)

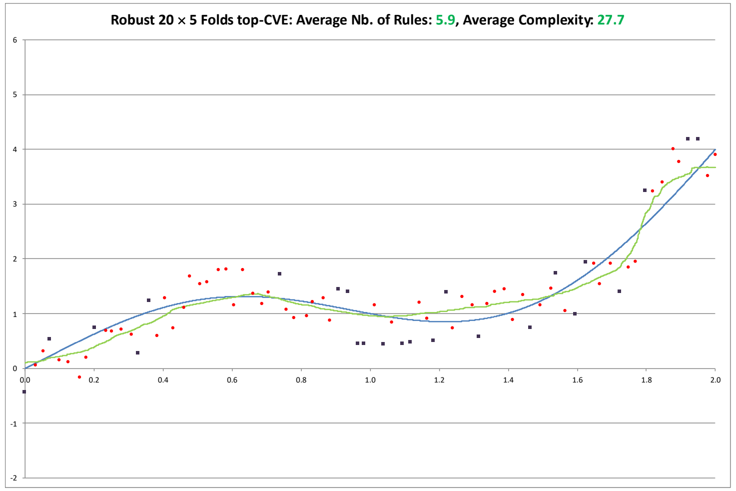
- Learning Points (after Corrective Induction) are **closer** to unknown points (RMSE : 6.39% *vs* 7.13%)
- Model is slightly **more Efficient** on unknown points (RMSE : 2.11% *vs* 2.58%)

80 Learning Points — 24 corrected Learning Points
Hidden Law: $y = \sin(\pi x) + x^2$
xtractis®



Noise estimation, after 2 rounds of corrective induction

RMSE (Learning Points / Hidden Values) = ~~0.428 (7.13%)~~ 0.383 (6.39%)
Correlation (Learning Points / Hidden Values) = ~~0.915~~ 0.928



Modeling with noisy data
(**after 1** and **after 2 rounds** of corrective induction)

Results from XTRACTIS® REVEAL 13.2.54199

Weak signals – Importance of predictors with low individual contribution

Breast Cancer Diagnosis (Zalilian approach: 30 potential predictors)

Top-IVE: 9 predictors, 8 rules (Intelligibility Score 4.87; Complexity 113)

Actual class / Predicted class

Diagnosis	Benign	Malignant		
Benign	5,033	16	99.68%	0.32%
Malignant	26	2,895	0.89%	99.11%
	99.49%	0.55%		
	0.51%	99.45%		
Refusals	0 (0.00%)	0 (0.00%)		

Quality of CVE Copy/Synthetic Validation 33%

Classification error: 0.53%

Actual class / Predicted class

Diagnosis	Benign	Malignant		
Benign	303	4	98.70%	1.30%
Malignant	0	176	0.00%	100.00%
	100.00%	2.22%		
	0.00%	97.78%		
Refusals	0 (0.00%)	0 (0.00%)		

Quality of CVE Copy/483 Original Points

Classification error: 0.83%

Actual class / Predicted class

Diagnosis	Benign	Malignant		
Benign	53	0	100.00%	0.00%
Malignant	1	32	3.03%	96.97%
	98.15%	0.00%		
	1.85%	100.00%		
Refusals	0 (0.00%)	0 (0.00%)		

Real Performance/External Test

Classification error: 1.16%

Original Dataset - 483 Original Points

Individual Contribution = Impact on Model Performance when Predictor is removed
Performance (F1-Score): 98.88%. Number of Refusals: 0 (0.00%)

Rank	Predictor	Impact on Model Performance	Performance Absolute Variation	Performance Relative Variation	Number of Refusals	Missing Values
1	area Cell - mean_3_largest	85.86%	-13.01	-13.16%	71 (14.70%)	0 (0.00%)
2	perimeter Cell - mean_3 largest	94.12%	-4.76	-4.81%	7 (1.45%)	0 (0.00%)
3	texture Cell - mean	94.33%	-4.55	-4.60%	3 (0.62%)	0 (0.00%)
4	concave points Cell - mean	94.97%	-3.91	-3.95%	20 (4.14%)	0 (0.00%)
5	texture Cell - mean_3_largest	97.18%	-1.70	-1.72%	13 (2.69%)	0 (0.00%)
6	concave points Cell - mean_3_largest	97.65%	-1.23	-1.24%	26 (5.38%)	0 (0.00%)
7	symmetry Cell - mean_3_largest	97.73%	-1.15	-1.16%	0 (0.00%)	0 (0.00%)
8	fractal dimension Cell - mean_3_largest	98.31%	-0.57	-0.58%	0 (0.00%)	0 (0.00%)
9	smoothness Cell - mean_3_largest	98.60%	-0.28	-0.28%	0 (0.00%)	0 (0.00%)

Percentages calculated on the 483 Patients of the Original Dataset - 483 Original Points dataset

Predictors with a rather low individual contribution

Reference dataset limited to the **4 predictors** with the highest individual contributions (Cartesian approach)

Top-IVE: 4 predictors, 7 rules (Intelligibility Score 4.94; Complexity 59)

Actual class / Predicted class

Diagnosis	Benign	Malignant		
Benign	5,020	32	99.37%	0.63%
Malignant	20	2,897	0.69%	99.31%
	99.60%	1.09%		
	0.40%	98.91%		
Refusals	0 (0.00%)	0 (0.00%)		

Quality of CVE Copy/Synthetic Validation 33%

Classification error: 0.65%

Actual class / Predicted class

Diagnosis	Benign	Malignant		
Benign	298	9	97.07%	2.93%
Malignant	5	171	2.84%	97.16%
	98.35%	5.00%		
	1.65%	95.00%		
Refusals	0 (0.00%)	0 (0.00%)		

Quality of CVE Copy/483 Original Points

Classification error: 2.90%

Actual class / Predicted class

Diagnosis	Benign	Malignant		
Benign	49	1	98.00%	2.00%
Malignant	5	31	13.89%	86.11%
	90.74%	3.13%		
	9.26%	96.88%		
Refusals	0 (0.00%)	0 (0.00%)		

Real Performance/External Test

Classification error: 6.98%

Original Dataset - 483 Original Points

Individual Contribution = Impact on Model Performance when Predictor is removed
Performance (F1-Score): 96.07%. Number of Refusals: 0 (0.00%)

Rank	Predictor	Impact on Model Performance	Performance Absolute Variation	Performance Relative Variation	Number of Refusals	Missing Values
1	area Cell - mean_3_largest	90.15%	-5.92	-6.16%	5 (1.04%)	0 (0.00%)
2	texture Cell - mean	92.97%	-3.10	-3.23%	13 (2.69%)	0 (0.00%)
3	concave points Cell - mean	94.31%	-1.75	-1.82%	42 (8.70%)	0 (0.00%)
4	perimeter Cell - mean_3 largest	96.76%	0.69	0.72%	308 (63.77%)	0 (0.00%)

Percentages calculated on the 483 Patients of the Original Dataset - 483 Original Points dataset

Scoring
Binomial Classification
Multinomial Classification
Regression

Data: Dr. William H. Wolberg, W. Nick Street, Olvi L. Mangasarian – University of Wisconsin., UCI Machine Learning Repository [Dua, D. and Graff, C. 2019]

Results from XTRACTIS® REVEAL 14.0.57285

Fuzzy Mathematics – suitable for real world

slide 18

Probability theory (Boolean lattice)	**Possibility** theory (non-complemented distributive lattice)
$P(A) = \int_{u \in A} p(u)du$ crisp event	$\Pi(A) = \sup_{u \in A}(\pi(u))$ binary event $\Pi(A) = \sup_{u \in U}(\min(\mu_A(u), \pi(u)))$ fuzzy event
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$	$\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$
Additive cardinal frequentist approach	Non-additive ordinal subjectivist approach
Stochastic uncertainty	Epistemic uncertainty

slide 49
slide 50

slide 48

Fuzzy measures of binary events (∨–∧–possibility, ∧–∨–necessity)

$$\forall A,B \in \{0,1\}^U,\quad \Pi(A\cup B)=\Pi(A)\vee\Pi(B) \qquad\qquad \forall A,B \in \{0,1\}^U,\quad \mathrm{N}(A\cap B)=\mathrm{N}(A)\wedge\mathrm{N}(B)$$

with A and B non-necessarily disjoints

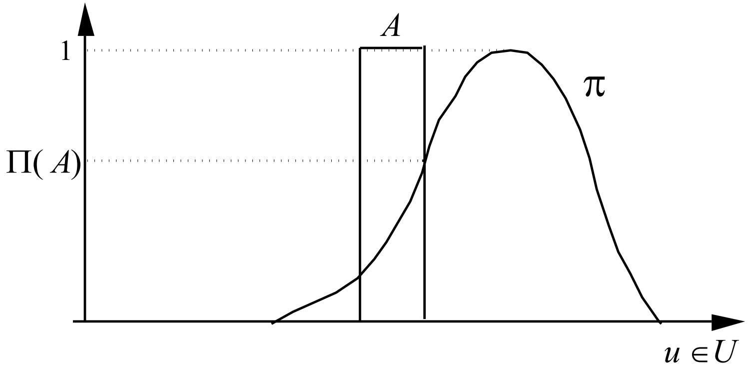
$$\forall A \in \{0,1\}^U,\quad \Pi(A)=\bigvee_{u\in A}\pi(u) \qquad \text{où} \qquad \pi(u)=\Pi(\{u\})$$

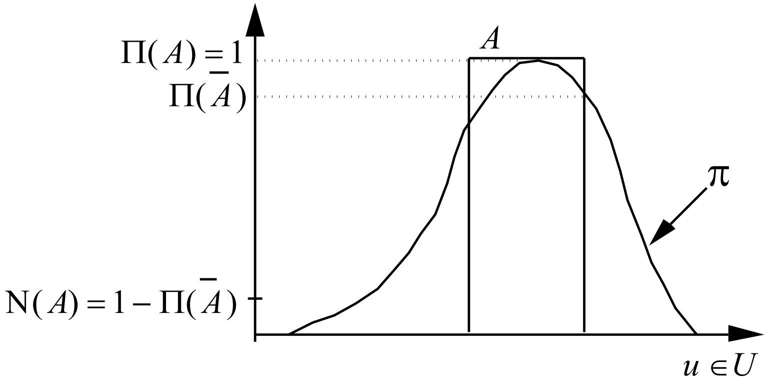
$$\forall A \in \{0,1\}^U,\quad \mathrm{N}(A)=1-\Pi(\overline{A}) \qquad\qquad \forall A \in \{0,1\}^U,\quad \mathrm{N}(A)=\bigwedge_{u\notin A}(1-\pi(u))=\bigwedge_{u\notin A}\overline{\pi}^{c_K}(u)$$

$$0\le \mathrm{N}(A)\le \mathrm{P}(A)\le \Pi(A)\le 1$$

slide 51

$$\forall A \in \{0,1\}^U,\quad \Pi(A)\vee\Pi(\overline{A})=1 \qquad\qquad \mathrm{N}(A)\wedge\mathrm{N}(\overline{A})=0$$





Fuzzy measures of fuzzy events (∨–∧–possibility, ∧–∨–necessity)

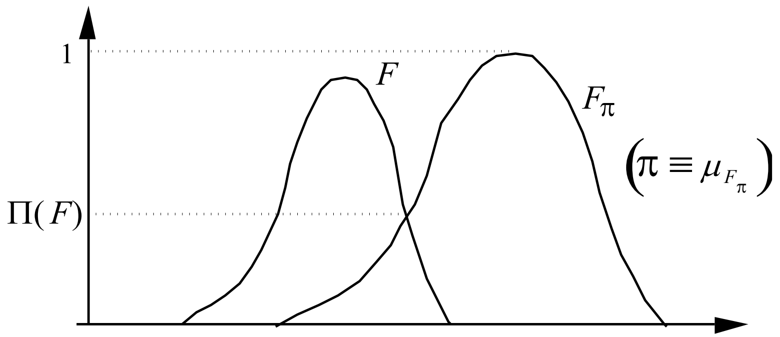
$$\forall F,G \in [0,1]^U, \quad \Pi(F \cup G) = \Pi(F) \vee \Pi(G) \qquad \mathrm{N}(F \cap G) = \mathrm{N}(F) \wedge \mathrm{N}(G)$$

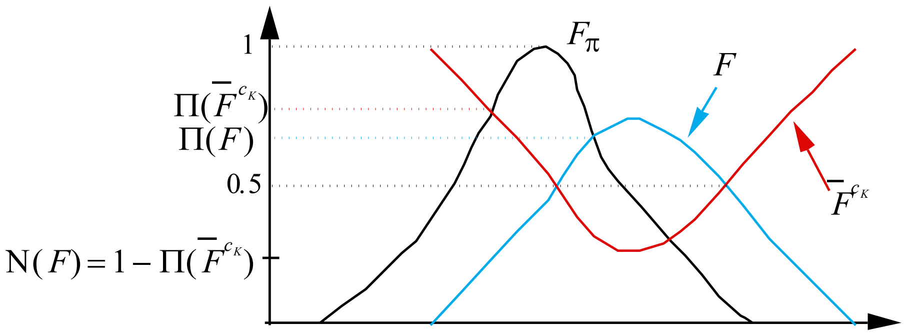
$$\forall F \in [0,1]^U, \quad \Pi(F) = \bigvee_{u \in U} \left(\mu_F(u) \wedge \pi(u)\right)$$

$$\forall F \in [0,1]^U, \quad \mathrm{N}(F) = 1 - \Pi(\overline{F}^{c_K}) \qquad \mathrm{N}(F) = \bigwedge_{u \in U} \left(\mu_F(u) \vee (1 - \pi(u))\right)$$

$$0 \le \mathrm{N}(F) \le \Pi(F) \le 1$$

$$\forall F \in [0,1]^U, \quad \Pi(F) \vee \Pi(\overline{F}^{c_K}) \ge 0.5 \qquad \mathrm{N}(F) \wedge \mathrm{N}(\overline{F}^{c_K}) \le 0.5$$





Fuzzy Mathematics – Possibility/Necessity *vs* Probability *4/4*

Helen wishes to measure/assess the confidence of the occurrence of the **binary event *A***: "Her boyfriend Peter ate *X* eggs for breakfast" with $x \in \{0, 1, 2, 3, 4, 5\}$.

Two informational situations can appear depending on the quantity and quality of information she has:

Situation 1: Helen takes daily breakfast with Peter and counts the number *x*. After about 6 months, she has sufficient accurate and certain information to propose the following probability distribution **p** of Peter's appetite for eggs for breakfast (additive cardinal frequentist approach):

X	0	1	2	3	4	5
p(x)	0	0.1	0.3	0.4	0.2	0

Situation 2 : For professional reasons, Peter and Helen do not share breakfast often during the week. However, after a few years together, Helen knows from experience that Peter cannot do without eggs before starting his day and that he loves to eat 2 or 3 eggs, regardless of cooking method (omelet, scrambled or boiled). Therefore, she can only propose a possibility distribution $\boldsymbol{\pi}$ of Peter's appetite for eggs for breakfast (non-additive ordinal subjectivist approach):

X	0	1	2	3	4	5
$\pi(x)$	0	0.5	1	1	0.2	0.1

According to the informational situation, Helen will therefore be able to answer the question by calculating, either

P(*A*) (Situation 1)

or

Π(*A*) and **N(*A*)** (Situation 2)

where *A* is any binary set of $X = \{0, 1, 2, 3, 4, 5\}$

(64 possible values of *A*)

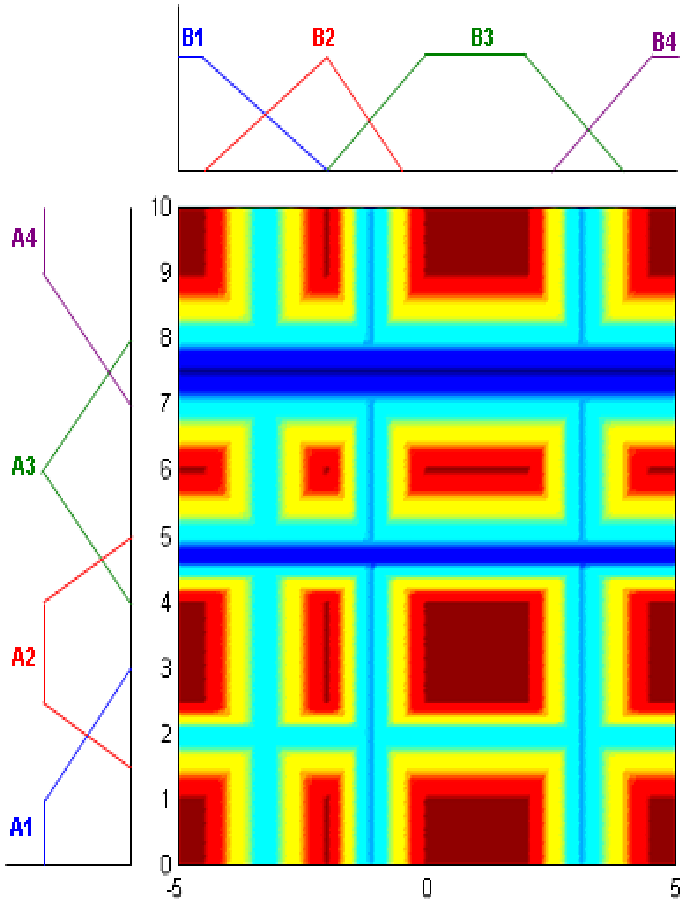
$A = \{0\}$	$\Pi(A) = 0$	$P(A) = 0$	$N(A) = 0$
$A = \{3\}$	$\Pi(A) = 1$	$P(A) = 0.4$	$N(A) = 0$
$A = \{0,5\}$	$\Pi(A) = 0.1$	$P(A) = 0$	$N(A) = 0$
$A = \{1,4\}$	$\Pi(A) = 0.5$	$P(A) = 0.3$	$N(A) = 0$
$A = \{1,5\}$	$\Pi(A) = 0.5$	$P(A) = 0.1$	$N(A) = 0$
$A = \{2,5\}$	$\Pi(A) = 1$	$P(A) = 0.3$	$N(A) = 0$
$A = \{4,5\}$	$\Pi(A) = 0.2$	$P(A) = 0.2$	$N(A) = 0$
$A = \{0,1,3\}$	$\Pi(A) = 1$	$P(A) = 0.5$	$N(A) = 0$
$A = \{0,1,3,4\}$	$\Pi(A) = 1$	$P(A) = 0.7$	$N(A) = 0$
$A = \{0,2,5\}$	$\Pi(A) = 1$	$P(A) = 0.3$	$N(A) = 0$
$A = \{3,4,5\}$	$\Pi(A) = 1$	$P(A) = 0.6$	$N(A) = 0$
$A = \{1,3,4\}$	$\Pi(A) = 1$	$P(A) = 0.7$	$N(A) = 0$
$A = \{2,3,5\}$	$\Pi(A) = 1$	$P(A) = 0.7$	$N(A) = 0.5$
$A = \{2,3,4,5\}$	$\Pi(A) = 1$	$P(A) = 0.9$	$N(A) = 0.5$
$A = \{1,2,3\}$	$\Pi(A) = 1$	$P(A) = 0.8$	$N(A) = 0.8$
$A = \{1,2,3,4\}$	$\Pi(A) = 1$	$P(A) = 1$	$N(A) = 0.9$
$A = \{1,2,3,4,5\}$	$\Pi(A) = 1$	$P(A) = 1$	$N(A) = 1$
$A = \{0,1,2,4,5\}$	$\Pi(A) = 1$	$P(A) = 0.6$	$N(A) = 0$

$$0 \le N(A) \le P(A) \le \Pi(A) \le 1$$

Fuzzy Deductive Inference System – Locality

	B1	B2	B3	B4
A4				
A3				
A2			C1	
A1				

IF [x is A2] AND [y is B3] **THEN** [z is C1]



➡ **Cells with fuzzy borders**

(Fuzzy Relations of order *N* – [Zalila 1993])

Gradual deductive reasoning / by analogy (Fuzzy Modus Ponens [Zadeh 1975])

⊥–T and T–⊥ Composition by anchoring of Fuzzy Relations of order *N*
(Generalized Fuzzy Modus Ponens [Zalila 1993])

IF the tomato is red **AND** soft **THEN** the tomato is ripe

This tomato is very red **AND** soft

This tomato is very ripe

$$\frac{\begin{array}{l}(A_1 \text{ AND } A_2) \rightarrow B \\ A_1' \text{ AND } A_2'\end{array}}{B'}$$

$$B'(y) = A_1' \times_{T'} A_2'(x_1,x_2) \circ_{\perp-T} R_{(A_1 \times_{T'} A_2) \rightarrow B}(x_1,x_2,y)$$

Fuzzy Deductive Inference System – Zalila Generalized algorithm

$$B' = \bigcup_{i}^{\vee}\left(\left(\top\, h(A'_j \mathbin{\overset{\frown}{\top}} A_{ji})\right) \langle\top\rangle\, B_i\right)$$

$$= \bigcup_{i}^{\vee}\left(\left(\top\, \Pi_{\top A'_j}(A_{ji})\right) \langle\top\rangle\, B_i\right)$$

where $\Pi_\top$ is the $\vee\!-\!\top$ generalized possibility [Zalila 1993]

idagrai LABS

XTRACTIS®

Reasoning modes – Human or Robotic

INDUCTION (11th c.) A → **Model?** → B

DEDUCTION (-4th c.) A’ → Model → **B’?**

ABDUCTION (20th c.) **A”?** → Model → B”

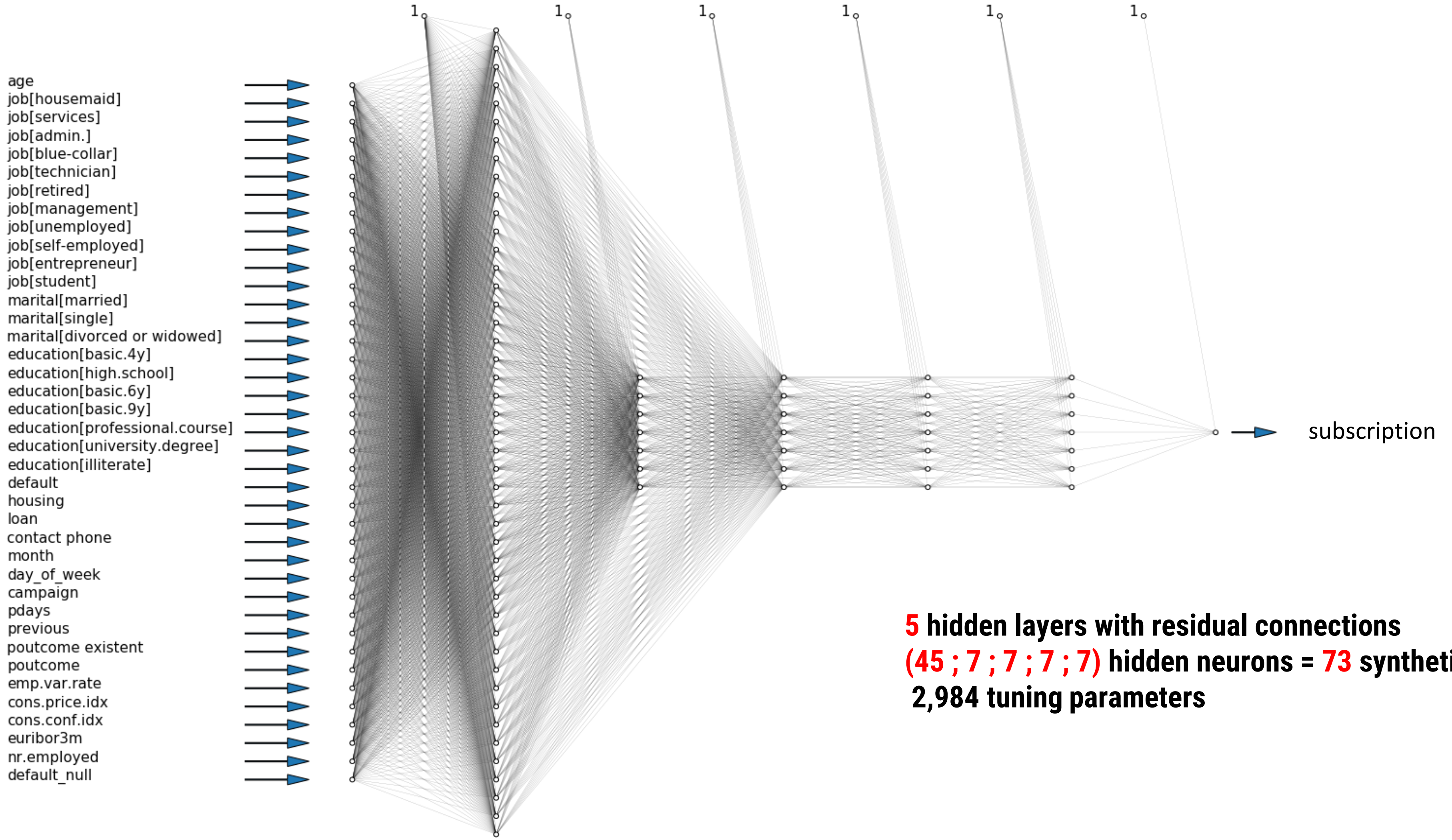
Possible reasons of **non-robustness**

- Absence of important predictors in the learning dataset
- Very few number of learning points
- Lot of noise (measurement, reading or transcription errors, conflicting cases) or missing values in the learning dataset
- Necessity to segment
- Random phenomenon

Recommended procedure

- ☑ Update (change / complete) the learning dataset in relation with field experts
- ☑ Run XTRACTIS® REVEAL on the updated learning dataset
- ☑ Assess the robustness of the new induced models

Structure of a Neural Network – Bank Marketing use case



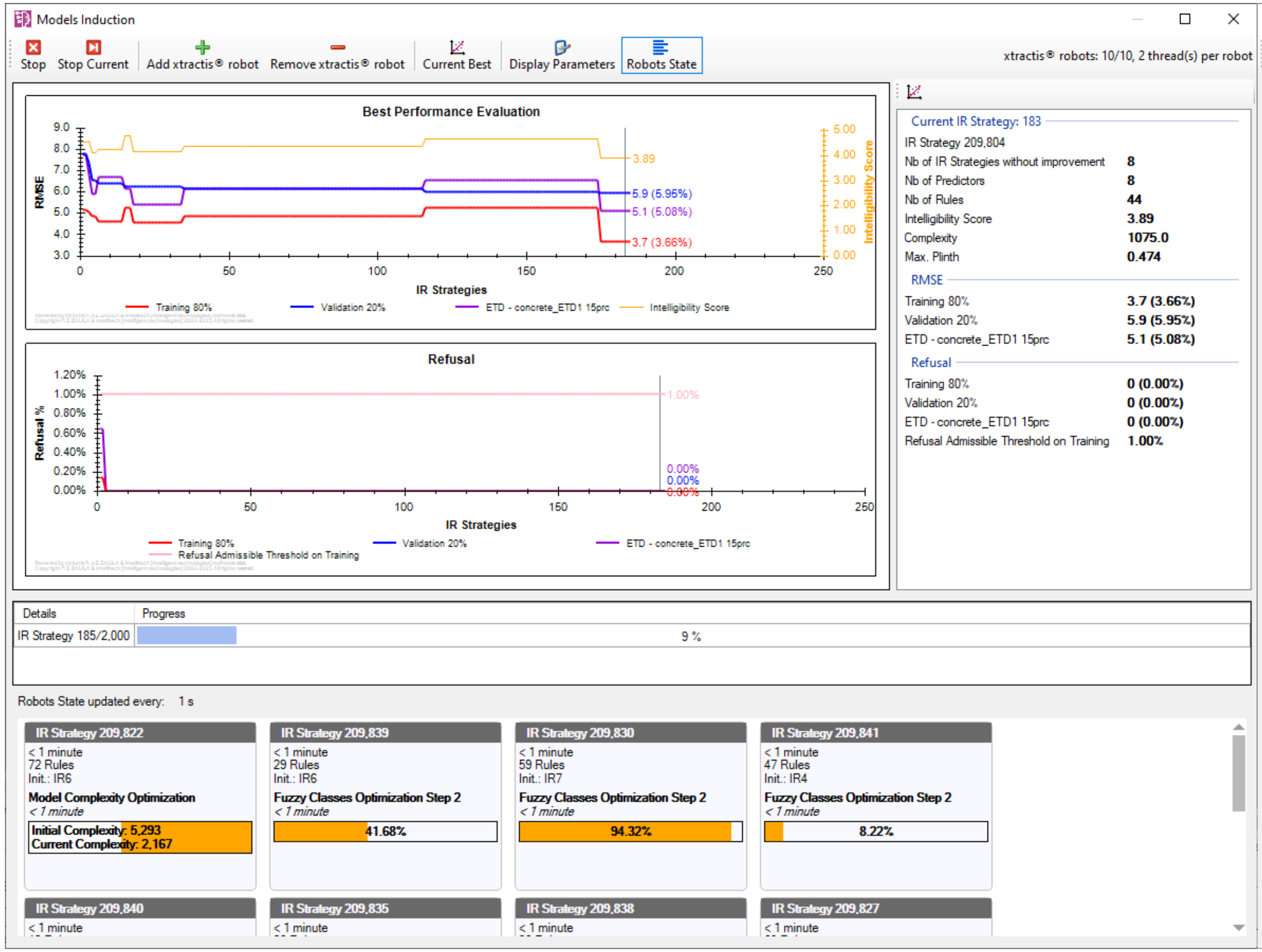
5 hidden layers with residual connections
(45 ; 7 ; 7 ; 7 ; 7) hidden neurons = 73 synthetic variables
2,984 tuning parameters

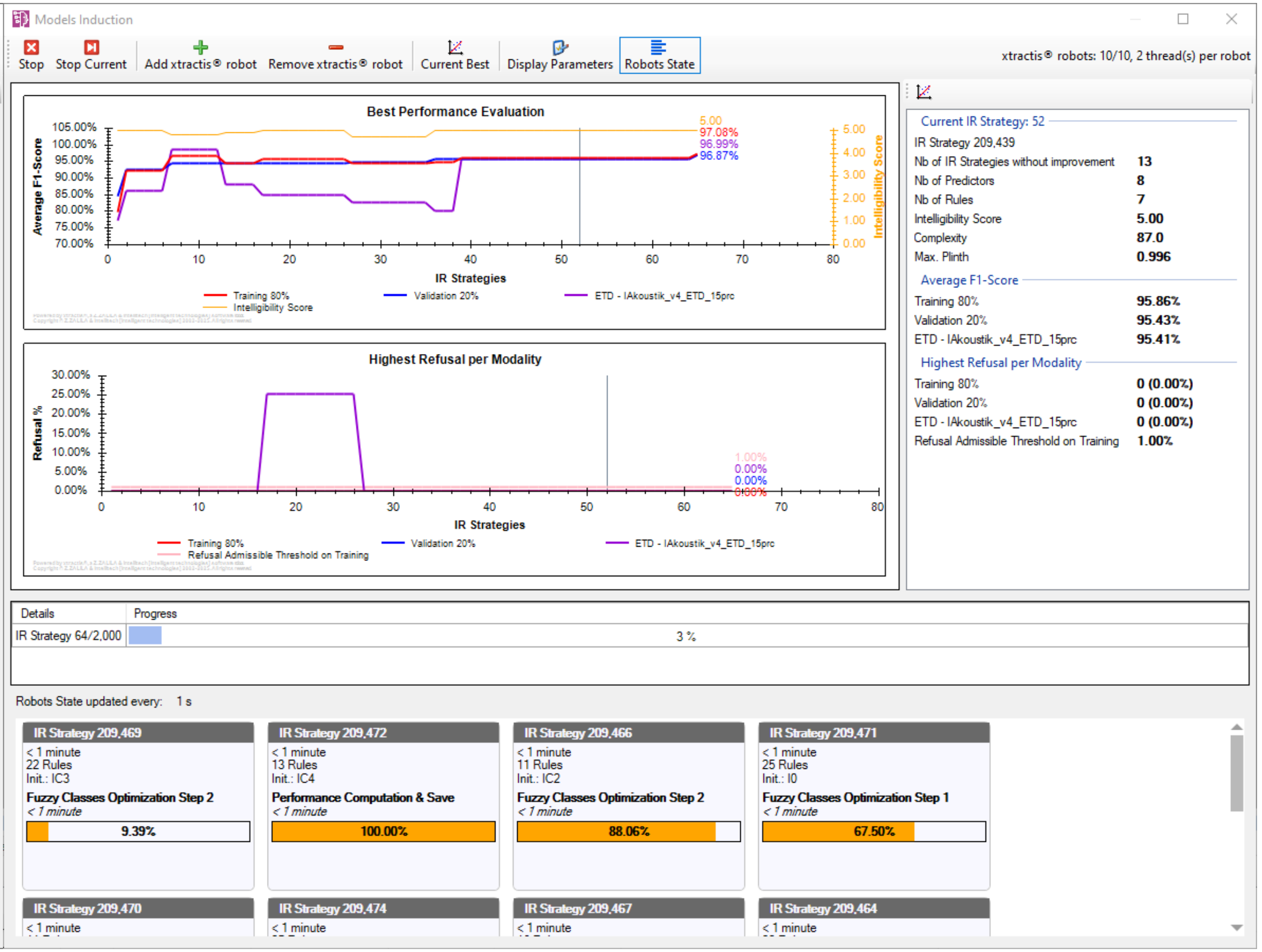
UPD 05/2018

Results from Python 3.6 packages: TensorFlow 1.7, Keras 2.1.4

Scoring
Binomial Classification
Multinomial Classification
Regression

Screenshots – Induction Convergence: Popper's refutability

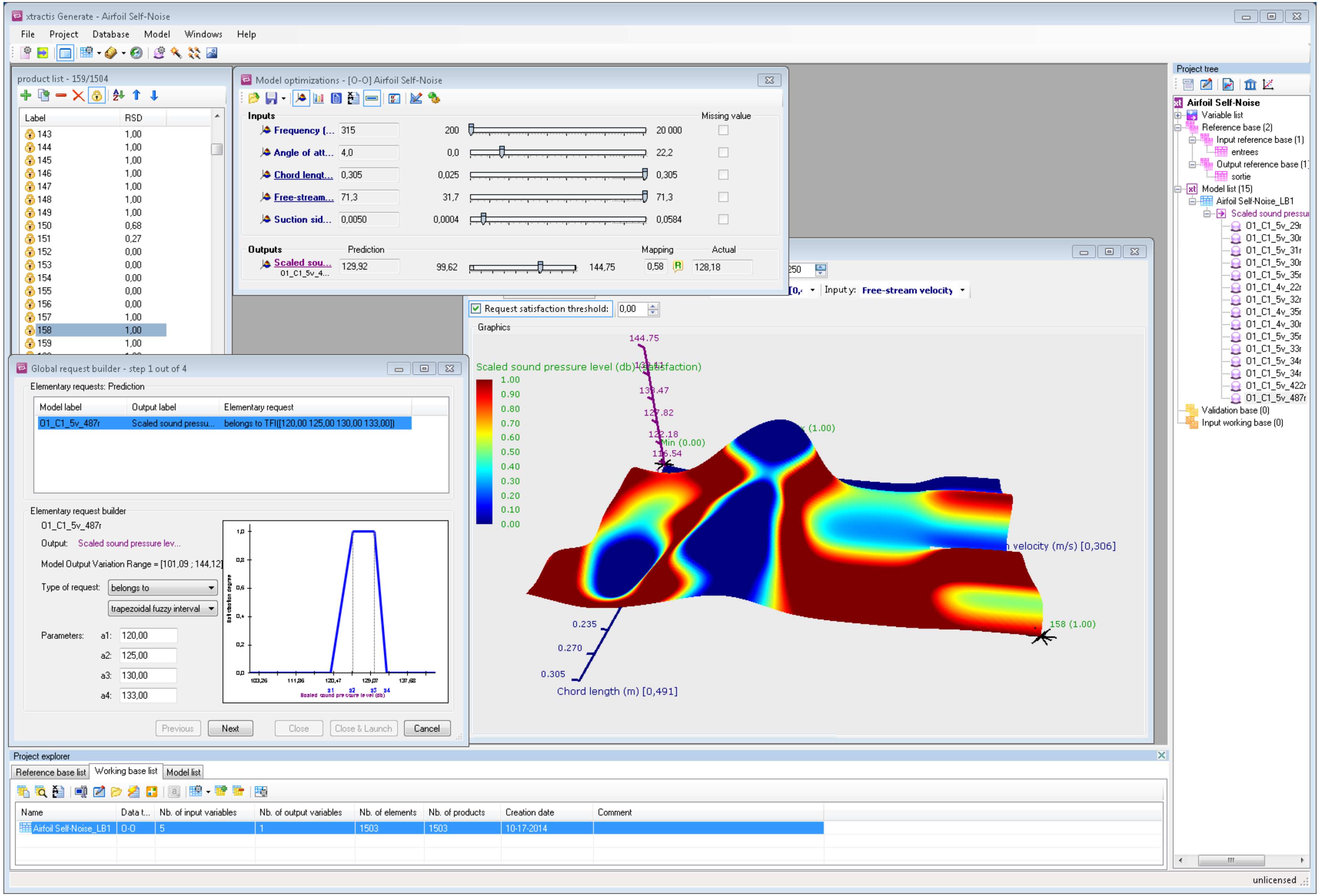




Regression

Multinomial Classification

Screenshots – The most optimal solutions: abduction



Name	Data t...	Nb. of input variables	Nb. of output variables	Nb. of elements	Nb. of products	Creation date	Comment
Airfoil Self-Noise_LB1	O-O	5	1	1503	1503	10-17-2014	

XTRACTIS®

Intelligibility Score [Zalila & Leroux 2023-2025]

$$IS(T_i) = \max(0.00\,,\, 5.00 + (Pen1 + Pen2 + Pen3 + Pen4 + Pen5))$$

- Penalty 1 (logarithmic penalty regarding the number of predictors):

$$Pen1(T_i) = \min(0\,,\ \ 1 - \log_{10} number\ of\ predictors)$$

- Penalty 2 (linear penalty regarding the average number of rules or equations per variable or modality to predict):

$$Pen2(T_i) = \min\left(0\,,\ \ \frac{1 - average\ number\ of\ rules\ or\ equations\ per\ variable\ or\ modality\ to\ predict}{40}\right)$$

- Penalty 3 (linear penalty regarding the average number of predictors per rule or equation):

$$Pen3(T_i) = \min\left(0\,,\ \ \frac{9 - 3 \times average\ number\ of\ predictors\ per\ rule\ or\ equation}{7}\right)$$

- Penalty 4 (linear penalty regarding the number of trees per chain, here for BT only):

$$Pen4(T_i) = \min(0\,,\ \ 1 - number\ of\ trees\ per\ chain)$$

- Penalty 5 (maximum penalty due to unintelligibility of synthetic variables, here for NN only):

$$Pen5(T_i) = -5$$

idagrai LABS

XTRACTIS®