Precision Medicine

CARDIOTOCOGRAPHIC IDENTIFICATION OF FETAL HEART CONDITIONS

Benchmark vs. Logistic Regression, Random Forest, Boosted Tree & Neural Network

UC#03 – 2024/03 (v3.0)   xtractis.ai

PROBLEM DEFINITION

GOAL

Design an AI-based decision system that accurately and instantly makes a rational medical diagnosis of fetal heart condition from signal characteristics of fetal heart rate and uterine contractions.

PROS & BENEFITS

- Identify parameters requiring increased vigilance and improve medical knowledge by helping cardiologists understand the causal relationships between specific cardiotocographic features, their combination, and the presence of an abnormality.
- Help the medical profession to make earlier and more personalized decisions by means of rapid, systematic, and explainable diagnoses. Extend access to high-level diagnoses even in medical deserts.
- Decrease prenatal mortality and avoid possible neurological sequelae for the fetus.

REFERENCE DATA

Source:
Dr. D. AYRES-DE-CAMPOS & J. BERNARDES – Faculty of Medicine, University of Porto

Dataset:
Dua, D. and Graff, C. (2019). UCI Machine Learning Repository http://archive.ics.uci.edu/ml. Irvine, CA: University of California, School of Information and Computer Science

Variable to Predict: Fetal heart behavior diagnosis among 10 conditions
[CS=Calm Sleep; REMS=Rapid Eye Movement Sleep; CV = Calm Vigilance; AV = Active Vigilance; SH = Shift Pattern (CS or Suspect with Shifts); AD = Accelerative/Decelerative Pattern (Stress Situation); DE = Decelerative Pattern (Vagal Stimulation); LD = Largely Decelerative Pattern; FS = Flat-Sinusoidal Pattern (Pathological State); SUSP = Suspect Pattern].

Potential Predictors: 21 parameters characterizing the fetal cardiotocograms and uterine contraction signals (Number of Uterine Contractions per Second, Fetal Heart Rate Baseline…).

Observations: 2,126 cardiotocographic signals related to different fetal cardiac conditions.

Data are divided into a Learning Dataset for model induction using Training and Validation Datasets, and an External Test Dataset to check the top model's performance on real data and for benchmarking.

Learning Dataset: 1,807 cases | 85% 80% for Training, 20% for Validation									
CS	REMS	CV	AV	SH	AD	LD	DE	FS	SUSP
18.04%	27.28%	2.49%	3.82%	3.38%	15.61%	11.84%	5.03%	3.27%	9.24%

External Test Dataset: 319 cases | 15%									
CS	REMS	CV	AV	SH	AD	LD	DE	FS	SUSP
18.18%	26.96%	2.51%	3.76%	3.45%	5.67%	11.91%	5.02%	3.14%	9.40%

MODEL TYPE   Regression   **Multinomial Classification**   Binomial Classification   Scoring

XTRACTIS-INDUCED DECISION SYSTEM

☑ **Intelligible Model, Explainable Decisions**

- The top-model is a decision system composed of 56 gradual rules without chaining, aggregated into 10 disjunctive rules.
- Each rule uses from 2 to 8 predictors among the 18 variables that XTRACTIS automatically identified as significant (out of the 21 potential predictors).
- Only a few rules are triggered at a time to compute the decision.

☑ **High Predictive Capacity**   It has a good Real Performance (on unknown data).

☑ **Ready to Deploy**   It computes real-time predictions up to 70,000 decisions/second, offline or online (API).

XTRACTIS PROCESS

STEPS	Reference Data	**INDUCTION**	XTRACTIS Top-Model	New Cases	**DEDUCTION**	CALM SLEEP / … / SUSPECT PATTERN Automated Decision (identify fetal heart condition)
SOFTWARE ROBOTS	XTRACTIS® **REVEAL** Delivers the decision system + its Structure & Performance Reports			XTRACTIS® **PREDICT** Delivers the decision + the Prediction Report explaining its reasoning		

TOP-MODEL INDUCTION

INDUCTION PARAMETERS

Powered by: XTRACTIS® REVEAL v12.1.41978

1. We launch 1,000 inductive reasoning strategies; each strategy is applied to 20 different 5-fold-partitions of the Learning Dataset to get a reliable assessment of the descriptive and predictive performances, respectively from Training and Validation Datasets.
2. Each strategy thus generates 100 unitary models called **Individual Virtual Expert** (IVE), whose decisions are aggregated with 3 possible operators into a **College of Virtual Experts** (CVE).
3. Among the 3,000 induced CVEs, the top-CVE with the best predictive performance remains complex: 2,095 rules sharing 21 predictors.

Given the small number of reference cases in the reference dataset, the XTRACTIS **CVE→IVE** Reverse-Engineering process is necessary to get a more intelligible model:

4. We build a synthetic dataset composed of 44,700 new cases simulated by deduction from the top-CVE, around the 149 original learning cases but distinct from them.
5. We apply 2,000 induction strategies to the same single 34% Training | 33% Validation | 33% Test partition of this new dataset: XTRACTIS induces 2,000 IVEs.
6. The top-IVE selected is the one that is the most intelligible while being as efficient as the top-CVE.

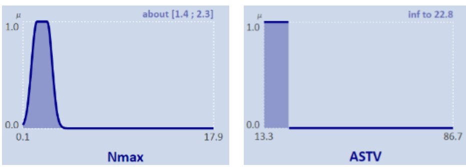
Total number of induced unitary models	Criterion for the induction optimization	Validation criterion for the top-model selection	Duration of the process (Induction Power FP64)
102,000 IVEs	Average F_1-Score	Average F_1-Score	9.5 days (1 Tflops)

TOP-MODEL STRUCTURE

The top-IVE model has a quite good intelligibility given the complexity of the phenomenon, as it combines 18 predictors into 56 rules with 4.2 predictors per rule on average. Its Structure Report reveals all the internal logic of the decision system and ensures that the model is understandable by the human expert. It is a transparent model that can be audited and certified before deployment to end-users.

PREDICTORS

- 18 signal characteristics out of 21
- 17 continuous variables + 1 nominal variable
- Ranked by individual contribution (3 strong signals, 7 medium & 8 weak): #1 Histogram Median / #2 Number of acc. per sec...
- Labeled by binary and fuzzy classes
  Examples: **binary interval** "inferior to 22.8"
  fuzzy interval "about [1.4 ; 2.3]"



RULES

- 56 connective fuzzy rules without chaining (aggregated into 10 disjunctive fuzzy rules)
- 2 to 8 predictors per rule (on average, 4.2 predictors per rule)
- Example: **fuzzy rule R17** uses 3 predictors and concludes "CALM VIGILANCE". 55 other fuzzy rules complete this model

IF	Number of accelerations per second (frequency)	IS	about [0.0021 ; 0.0030]
AND	Percentage of time with abnormal short-term variability	IS	inferior to 22.8
AND	Number of histogram peaks	IS	about [1.4 ; 2.3]
THEN	Fetal Heart Condition	IS	CALM VIGILANCE

Literally, the Heart Condition if the fetus gets a Calm Vigilance diagnosis if the Number of accelerations per second is about between 0.0021 and 0.0030 accelerations per second, and the Percentage of time with abnormal short-term variability is under around 22.8% and the Number of histogram peaks is between around 1.4 peaks and around 2.3 peaks.

TOP-MODEL PERFORMANCE

The top-IVE performances, measured in Training/Validation/Test on synthetic data, then in External Test on reference data, guarantee the model's predictive and real performances.

Performance / Dataset	Synthetic Data: DESCRIPTIVE (34% Training)	Synthetic Data: PREDICTIVE (33% Validation)	Synthetic Data: REAL (33% Test)	Reference Data: REAL (External Test)
Average F_1-Score	98.72%	98.48%	98.40%	87.07%
Classification Error	1.26%	1.41%	1.54%	12.42%

EXPLAINED PREDICTIONS FOR 4 UNKNOWN CASES

Powered by: XTRACTIS® PREDICT version 12.1.41978

CASE (from the External Dataset, i.e., not included in the Learning Dataset)	DEDUCTIVE INFERENCE OF RULES	AUTOMATED DECISION

PATIENT #CTG0381

actual value = REMS	
Fetal Heart Rate Baseline	141.0
Number of accelerations per second	0.00800
Number of fetal movements per second	0.024000
Number of Uterine contractions per second	0.0016
Number of light decelerations per second	0.00160
Percentage of time with abnormal short term variability	50.0
Mean value of short term variability	0.90
Percentage of time with abnormal long term variability	1.0
Mean Value of Long Term Variability	4.300
Width of FHR Histogram	114
Minimum of FHR Histogram	58
Maximum of FHR histogram	172
Number of Histogram Peaks	7.0
Histogram Mode	148
Histogram Mean	147
Histogram Median	151
Histogram Variance	7
Histogram Tendency	Right Asymmetric

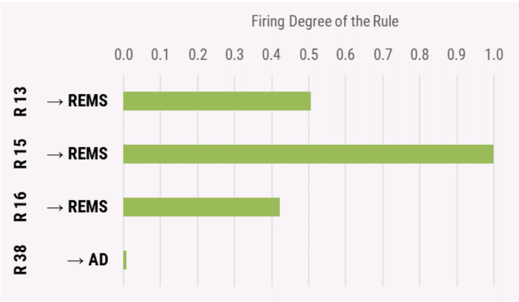
Real Time

For this patient, 4 rules are triggered:

R15 at 1.000, **R13** at 0.505, and **R16** at 0.422
to conclude "Rapid Eye Movement Sleep".

R38 is fired at 0.008
to conclude "Accelerative / Decelerative Pattern".

The other 52 rules are not activated.



NUMBER OF TRIGGERED RULES
4 / 56

FUZZY PREDICTION
{ REMS | 1.000,
AD | 0.008 }

FINAL PREDICTION
{ REMS }

The system delivers a correct diagnosis of the heart condition compared to that given by the cardiologist:

RAPID EYE MOVEMENT SLEEP

PATIENT #CTG1926

actual value = DE	
Fetal Heart Rate Baseline	140.0
Number of accelerations per second	0.00092
Number of fetal movements per second	0.001845
Number of Uterine contractions per second	0.0055
Number of light decelerations per second	0.00554
Percentage of time with abnormal short term variability	63.0
Mean value of short term variability	1.50
Percentage of time with abnormal long term variability	1.0
Mean value of long term variability	9.500
Width of FHR Histogram	163
Minimum of FHR Histogram	65
Maximum of FHR histogram	228
Number of Histogram Peaks	5.0
Histogram Mode	142
Histogram Mean	118
Histogram Median	141
Histogram Variance	27
Histogram Tendency	Symmetric

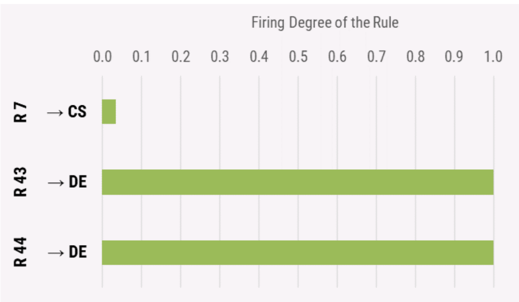
Real Time

For this patient, 3 rules are triggered:

R43 and **R44** at 1.000
to conclude "Decelerative Pattern".

R7 is fired at 0.034
to conclude "Calm Sleep".

The other 53 rules are not activated.



NUMBER OF TRIGGERED RULES
3 / 56

FUZZY PREDICTION
{ DE | 1.000,
CS | 0.034 }

FINAL PREDICTION
{ DE }

The system delivers a correct diagnosis of the heart condition compared to that given by the cardiologist:

DECELERATIVE PATTERN

CASE (from the External Dataset, i.e., not included in the Learning Dataset)	DEDUCTIVE INFERENCE OF RULES	DECISION

PATIENT #CTG1754

actual value = LD	
Fetal Heart Rate Baseline	134.0
Number of accelerations per second	0.00421
Number of fetal movements per second	0.001404
Number of Uterine contractions per second	0.0014
Number of light decelerations per second	0.00421
Percentage of time with abnormal short term variability	60.0
Mean value of short term variability	1.60
Percentage of time with abnormal long term variability*	0.0
Mean value of long term variability*	0.000
Width of FHR Histogram	113
Minimum of FHR Histogram	71
Maximum of FHR histogram	184
Number of Histogram Peaks	7.0
Histogram Mode	89
Histogram Mean	118
Histogram Median	113
Histogram Variance	195
Histogram Tendency	Left Asymmetric

Real Time

For this patient, 4 rules are triggered:

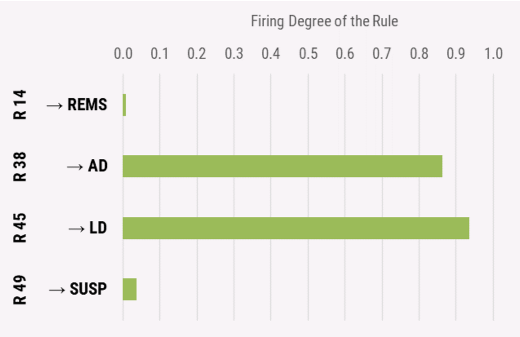
R45 is fired at 0.935
to conclude “Largely Decelerative Pattern”.

R38 is fired at 0.864
to conclude “Accelerative/Decelerative Pattern”.

R49 is fired at 0.036
to conclude “Suspect Pattern”.

R14 is fired at 0.008
to conclude “Rapid Eye Movement Sleep”.

The other 52 rules are not activated.



NUMBER OF TRIGGERED RULES
4 / 56

FUZZY PREDICTION
{ LD | 0.935,
AD | 0.864,
SUSP | 0.036,
REMS | 0.008 }

FINAL PREDICTION
{ LD }

The system delivers a correct diagnosis of the heart condition compared to that given by the cardiologist:

LARGELY DECELERATIVE PATTERN

PATIENT #CTG0450

actual value = CS	
Fetal Heart Rate Baseline	135.0
Number of accelerations per second	0.00123
Number of fetal movements per second	0.002466
Number of Uterine contractions per second*	0.0000
Number of light decelerations per second*	0.00000
Percentage of time with abnormal short term variability	58.0
Mean value of short term variability	0.60
Percentage of time with abnormal long term variability	15.0
Mean value of long term variability	8.000
Width of FHR Histogram	95
Minimum of FHR Histogram	70
Maximum of FHR histogram	165
Number of Histogram Peaks	5.0
Histogram Mode	139
Histogram Mean	140
Histogram Median	141
Histogram Variance	2
Histogram Tendency	Right Asymmetric

Real Time

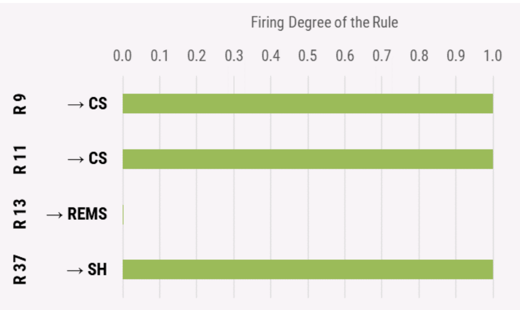
For this patient, 4 rules are triggered:

R9 and **R11** at 1.000
to conclude “Calm Sleep”.

R37 is fired at 1.000
to conclude “Shift Pattern: Calm Sleep or Suspect With Shifts”.

R13 is fired at 0.002
to conclude “Rapid Eye Movement Sleep”.

The other 52 rules are not activated.



NUMBER OF TRIGGERED RULES
4 / 56

FUZZY PREDICTION
{ CS | 1.000,
SH | 1.000,
REMS | 0.002 }

FINAL PREDICTION
REFUSAL

The decision system cannot choose between “Calm Sleep” and “Shift Pattern” so it refuses to decide.

This warning means that a thorough opinion of the cardiologist is required.

More training data with situations near this patient profile should strengthen the model in this decision space area.

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. It is one situation of the “Refusal” prediction.

TOP-MODELS BENCHMARK: DECISION STRUCTURE & INTELLIGIBILITY × PERFORMANCE SCORES

MODELING PARAMETERS

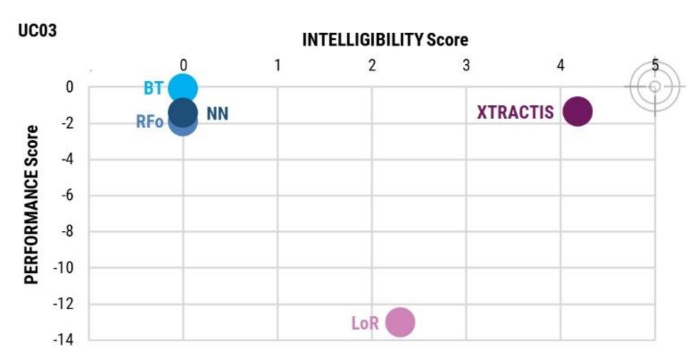
	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREE	NEURAL NETWORK
MODELS RELEASE	2022/05	2022/09	2022/04	2022/04	2022/04
ALGORITHM VERSION	XTRACTIS **REVEAL** 12.1.41978	Python 3.9 | Scikit-Learn 1.0.2	Python 3.6 | LightGBM 2.2.2	Python 3.6 | LightGBM 2.2.2	Python 3.6 | TensorFlow 2.6.2 | Keras 2.6.0
CROSS-VALIDATION TECHNIQUE	20×5 folds for each CVE model. Then 1-Split Validation for each IVE model: 34% Training | 33% Validation | 33% Test	20×5 folds for each CVE model	20×5 folds for each CVE model	20×5 folds for each CVE model	20×5 folds for each CVE model
NUMBER OF EXPLORED STRATEGIES(1)	1,000 induction strategies for the CVE on Training / Validation data. 2,000 induction strategies for the IVE on synthetic data	2,000 data analysis strategies on Training / Validation data	2,000 ML strategies on Training / Validation data	2,000 ML strategies on Training / Validation data	619 ML strategies on Training / Validation data
TOP-MODEL SELECTION(2)	Top-CVE among 3,000 CVEs. Then Top-IVE among 2,000 IVEs	Top-CVE selected among 2,000 CVEs	Top-CVE selected among 2,000 CVEs	Top-CVE selected among 2,000 CVEs	Top-CVE selected among 619 CVEs
		Then single model obtained by applying best CVE strategy on 100% of the Learning Dataset			

TOP-MODEL STRUCTURE

	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREE	NEURAL NETWORK
NUMBER OF PREDICTORS (out of 21 Potential Predictors)	**18**	**20**	**21**	**20**	**21**
AVERAGE NUMBER OF PREDICTORS PER RULE OR EQUATION	**4.2** per rule	**8.6** per equation	**5.6** per rule	**4.8** per rule	**28.8** per equation
STRUCTURE OF THE DECISION SYSTEM	**56** fuzzy rules without chaining (aggregated into 10 disjunctive fuzzy rules) Only some rules are triggered at a time to compute a prediction	**10** linear equations	**470** trees without chaining **26,435** binary rules	**10** chains of **75** trees **15,932** binary rules Tree #N corrects the error of the N-1 previous trees	**3** hidden layers | **90** hidden nodes **100** equations 90 unintelligible synthetic variables

TOP-MODEL SCORES

	Random(3)	XTRACTIS	LoR	RFo	BT	NN
INTELLIGIBILITY Score(4)		4.18	2.30	0.00	0.00	0.00
CVE Real Performance (Average F_1-Score) in External Test		86.35	73.95	86.70	89.05	87.75
Gap to CVE Leader in External Test		-2.70	-15.10	-2.35	0.00	-1.30
IVE Real Performance (Average F_1-Score) in External Test	*15.96*	87.07	76.10	85.59	86.92	85.54
Gap to IVE Leader in External Test		0.00	-10.97	-1.48	-0.15	-1.53
Average Real Performance in External Test	15.96	86.71	75.03	86.15	87.99	86.65
PERFORMANCE Score(4)		-1.35	-13.04	-1.92	-0.08	-1.42



(1) For all algos: on the same Learning Dataset. All Models are optimized according to their Validation Average F_1-Score.
(2) All top-models are selected according to their Validation Average F_1-Score while checking that it remains close to their Training Average F_1-Score.
(3) Baseline performances that models must exceed to perform better than chance (P-value = 0.001; 100,000 models generated by random permutation of the output values). The value of each performance criterion is generally achieved by a different random model.
(4) See Appendices for explanations and detailed results. Performance Scores are calculated on all available unknown data.

More Use Cases:
xtractis.ai/use-cases/

APPENDIX 1 — Calculation of the Intelligibility × Performance Scores

AI Technique #i	T_i	$i \in [1 ; n]$ n = number of AI Techniques benchmarked in terms of data-driven modeling = 5
Benchmark #k	B_k	$k \in [1 ; p]$ p = number of Benchmarks for the Use Case ∈ {1, 2, 3}

Remarks:

- In case of a small number of reference data, a CVE model (College of Virtual Experts) is generated by each explored strategy of T_i, generally via an N×K-fold cross validation. In this case, a Benchmark is led with the top-CVE on the External Test Dataset (ETD, composed of unknown reference cases). Then, a top-IVE model (Individual Virtual Expert) is generated from the top-CVE, through the XTRACTIS® reverse-engineering process, or for the other T_i, by applying the top-strategy, which has generated the top-CVE, on the Training and Validation Datasets. And a second Benchmark is led with this top-IVE on the same ETD.
- In case of a huge number of reference data, an IVE is generated by each explored strategy of T_i, via a 1-split validation. In this case, Benchmarks are led with the top-IVE on the Test Dataset (TD, composed of unknown reference cases) and on the available ETDs.
- Each Benchmark uses the latest versions of the following algorithms available at the date of the benchmark. XTRACTIS®: REVEAL; Logistic Regression: Python, Scikit-Learn; Random Forest & Boosted Tree: Python, LightGBM; Neural Network: Python, TensorFlow, Keras.
- Each B_k uses exactly the same TD and ETD for each T_i model.
- No Regression models can be obtained by Logistic Regression. So, this Data Analysis technique is benchmarked only for Classification or Scoring problems.
- The Holy Grail for critical AI-based decision systems is to obtain a model with the highest Performance and the highest Intelligibility scores (top-right corner of the graph).

PERFORMANCE Score

For each B_k, we calculate the values of the Performance Criterion (PC) on the same ETD for all the T_i top-CVEs; and on the same TD and ETDs for all the T_i top-IVEs. The PC is: RMSE in percentage for a Regression; F_1-Score for a Binomial Classification; Average F_1-Score or Average F_2-Score for a Multinomial Classification; Gini index for a Scoring.
Then, we compare the value of the PC of each T_i top-CVE (resp. top-IVE) to the best value of this PC reached by the best T_i top-CVE (resp. top-IVE) on ETD (resp. on TD and ETDs).

For Regression, we calculate for each T_i top-model (CVE and IVE): $PS(T_i, B_k) = Best\_PC(B_k) - PC(T_i, B_k)$.

For Classification and Scoring, we calculate for each T_i top-model: $PS(T_i, B_k) = PC(T_i, B_k) - Best\_PC(B_k)$.

Performance Score of T_i

$$\boxed{PS(T_i) = \text{Mean}\left(PS(T_i, B_k)\right)_{k \in [1 ; p]}}$$

Remark:

- Each PS varies theoretically from -100 (Lowest Score) to 0 (Highest Score), but practically between -50 and 0.

INTELLIGIBILITY Score

We consider the T_i top-IVE. Its Intelligibility Score $IS(T_i)$ is valued from 0.00 to 5.00 regarding the structure of the model: number of predictors, classes, rules, equations, trees, synthetic variables, modalities to predict for classifications (or numeric variables to predict for regressions or scoring). The more compact the model, the higher its IS.

The IS of each T_i is obtained by accumulating the following five penalty values to the ideal IS value of 5.00 (each penalty has a null or a negative value):

- Penalty 1 (logarithmic penalty regarding the number of predictors):
  $$Pen1(T_i) = \min(0\ , 1 - \log_{10} number\ of\ predictors)$$
  Examples: *Pen1 = 0.00 for up to 10 predictors*
  Pen1 = −3.00 for 10.000 predictors

- Penalty 2 (linear penalty regarding the average number of rules or equations per modality to predict):
  $$Pen2(T_i) = \min\left(0\ , 0.01 - \frac{average\ number\ of\ rules\ or\ equations\ per\ modality\ to\ predict}{100}\right)$$
  Examples: *Pen2 = 0.00 for 1 rule or equation per modality to predict on average*
  Pen2 = −3.00 for 301 rules or equations per modality to predict on average

- Penalty 3 (linear penalty regarding the average number of predictors per rule or equation):
  $$Pen3(T_i) = \min\left(0\ , \frac{9 - 3 \times average\ number\ of\ predictors\ per\ rule\ or\ equation}{7}\right)$$
  Examples: *Pen3 = 0.00 for up to 3.0 predictors per rule or equation on average*
  Pen3 = −3.00 for 10.0 predictors per rule or equation on average

- Penalty 4 (linear penalty regarding the number of chained trees, here for BT only):
  $$Pen4(T_i) = \min(0\ , 1 - number\ of\ chained\ trees)$$
  Examples: *Pen4 = 0.00 for 1 tree*
  Pen4 = −3.00 for 4 chained trees

- Penalty 5 (maximum penalty due to unintelligibility of synthetic variables, here for NN only):
  $$Pen5(T_i) = -5$$

Intelligibility Score of T_i

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

Remarks:

- For the difference between the Intelligibility and the Explainability of a model, please see the XTRACTIS® Brochure, page 7.
- The real complexity of the process/phenomenon under study is intrinsic, i.e., it could not be reduced or simplified, but only discovered; thus, the top-model will be complex if the process/phenomenon turns out to be complex [Zalila 2017]. Consequently, for some complex process/phenomenon, IS can be equal to 3.00 or less, even if Ti natively produces intelligible models (XTRACTIS, Random Forest).
- For similar structures, the Boosted Tree model is always less intelligible than the Random Forest one, as it is composed of chains of trees, instead of a college of trees (see Penalty 4).
- Neural Network model has always the lowest IS of 0.00, because it uses synthetic unintelligible variables (hidden nodes) in addition to all the potential predictors (see Penalty 5).

APPENDIX 2 — Use Case Results (all Performance criteria of all Top-Models)

Performance Criterion	Classification Error	Min. Sensitivity	Average Sensitivity	Min. PPV	Average PPV	Min. F1-Score	**Average F_1-Score**	Weighted Average F_1-Score	Refusal
RANDOM MODEL									
Nb of Random Permutations (P-value) = 100,000 (0.001%)									
Performance against chance (External Test)	*77.43%*	*6.67%*	*15.96%*	*6.67%*	*15.96%*	*6.67%*	***15.96%***	*22.57%*	
XTRACTIS TOP-MODEL									
CVE - Descriptive Performance (Training)	10.51%	75.41%	88.66%	68.66%	87.93%	71.88%	88.26%	89.53%	0 (0.00%)
CVE - Predictive Performance (Validation)	10.63%	75.41%	88.20%	68.66%	87.92%	71.88%	88.02%	89.42%	0 (0.00%)
CVE - Real Performance (External Test)	12.54%	72.73%	88.04%	61.54%	85.25%	66.67%	86.35%	87.62%	0 (0.00%)
IVE - Descriptive Performance (Training)	1.26%	97.59%	98.98%	96.42%	98.46%	97.47%	98.72%	98.74%	119 (0.48%)
IVE - Predictive Performance (Validation)	1.41%	97.30%	98.59%	96.69%	98.38%	97.37%	98.48%	98.59%	121 (0.51%)
IVE - Real Performance (Test)	1.54%	96.54%	98.50%	96.14%	98.32%	97.47%	98.40%	98.46%	132 (0.55%)
IVE - Real Performance (1,807 original points)	12.42%	75.00%	86.58%	65.22%	86.07%	69.77%	86.23%	87.64%	12 (0.66%)
IVE - Real Performance (External Test)	12.22%	70.00%	88.78%	70.00%	85.98%	70.00%	87.07%	87.82%	8 (2.51%)
LOGISTIC REGRESSION TOP-MODEL									
CVE - Descriptive Performance (Training)	22.13%	52.45%	83.08%	47.27%	70.70%	52.76%	73.92%	78.47%	
CVE - Predictive Performance (Validation)	24.41%	51.53%	78.95%	30.83%	67.96%	44.85%	70.75%	76.44%	
CVE - Real Performance (External Test)	21.94%	55.17%	83.64%	35.29%	71.01%	48.00%	73.95%	78.81%	
IVE - Descriptive Performance (Training)	21.58%	53.37%	82.73%	37.17%	71.15%	53.16%	74.32%	78.99%	
IVE - Real Performance (External Test)	20.06%	55.17%	84.64%	37.50%	73.04%	50.00%	76.10%	80.42%	
RANDOM FOREST TOP-MODEL									
CVE - Descriptive Performance (Training)	0.83%	97.55%	99.60%	84.91%	97.69%	91.84%	98.56%	99.19%	
CVE - Predictive Performance (Validation)	12.62%	72.13%	86.72%	61.97%	83.20%	66.67%	84.81%	87.47%	
CVE - Real Performance (External Test)	11.60%	72.73%	89.55%	61.54%	84.72%	66.67%	86.70%	88.54%	
IVE - Descriptive Performance (Training)	0.89%	97.55%	99.58%	86.54%	97.55%	92.78%	98.50%	99.13%	
IVE - Real Performance (External Test)	12.23%	72.73%	88.09%	61.54%	83.83%	66.67%	85.59%	87.96%	
BOOSTED TREE TOP-MODEL									
CVE - Descriptive Performance (Training)	0.05%	99.69%	99.97%	97.83%	99.78%	99.90%	99.87%	99.94%	
CVE - Predictive Performance (Validation)	10.24%	63.93%	86.05%	83.33%	90.14%	72.90%	87.87%	89.69%	
CVE - Real Performance (External Test)	9.40%	54.55%	88.47%	70.00%	90.65%	66.67%	89.05%	90.51%	
IVE - Descriptive Performance (Training)	0.11%	99.39%	99.94%	97.83%	99.62%	98.90%	99.78%	99.89%	
IVE - Real Performance (External Test)	9.72%	50.00%	84.95%	66.67%	89.98%	57.14%	86.92%	90.07%	
NEURAL NETWORK TOP-MODEL									
CVE - Descriptive Performance (Training)	0.61%	94.92%	99.14%	96.53%	99.32%	97.39%	99.22%	99.39%	
CVE - Predictive Performance (Validation)	10.63%	73.33%	86.08%	78.57%	89.17%	75.86%	87.54%	89.34%	
CVE - Real Performance (External Test)	10.34%	62.50%	85.90%	77.78%	90.30%	70.00%	87.75%	89.54%	
IVE - Descriptive Performance (Training)	3.21%	91.53%	95.74%	90.00%	95.84%	90.76%	95.78%	96.79%	
IVE - Real Performance (External Test)	14.11%	62.50%	84.72%	58.33%	87.56%	60.87%	85.54%	85.94%	