



Cyber Security

LOG-BASED IDENTIFICATION OF CYBER INTRUSIONS (DARPA)

Benchmark vs. Logistic Regression, Random Forests, Boosted Trees & Neural Networks

2024/02 (v3.0)

xtractis.ai

PROBLEM DEFINITION

GOAL Design an AI-based decision system that accurately detects an intrusion on a computer network and identifies the type of attack from features of the connection logs, to instantly execute the appropriate rational action.

- PROS & BENEFITS**
- ▶ Identify the characteristics of logs defining a cyber intrusion. Enhance expert knowledge by helping cybersecurity specialists understand the causal relationships between specific log features, their combination, and the type of intrusion.
 - ▶ Help IT diagnose the type of the cyberattack as early as possible and understand the underlying strategy of the attacker in order to consider measures to thwart future attacks.
 - ▶ Avoid many false alarms thanks to transparent diagnosis, in a context of increasing number of attacks with the use of open-source AI algorithms.

REFERENCE DATA **Variable to Predict** The model predicts the network state among 5 possible classes: **NORMAL** | **DOS** = Denial of Service attacks | **PROBE** = Probing attacks | **U2R** = User to Root attack | **R2L** = Root to Local attacks

Predictive Variables 41 Potential Predictors characterizing each log: duration, protocol type, network service, number of data bytes from source to destination, flag status of connection...

Source:
Cyber Systems and Technology group of MIT Lincoln Laboratory, DARPA ITO, Air Force Research Laboratory [UCI Machine Learning Repository].

Observations 1,074,983 connection logs on the US Air Force military computer network. Each log is associated with a normal activity or a type of attack. Data are divided into

- a Learning Dataset for model induction using Training, Validation and Test Datasets,
- and an External Test Dataset (ETD#1) with an environment close to the learning one to check the top model's performance on real data and for benchmarking.

An additional dataset of 70,874 connections corresponding to a network environment that has strongly changed is used as a second External Test Dataset (ETD#2).

All duplicates were removed from the reference dataset to avoid biasing performance assessment.

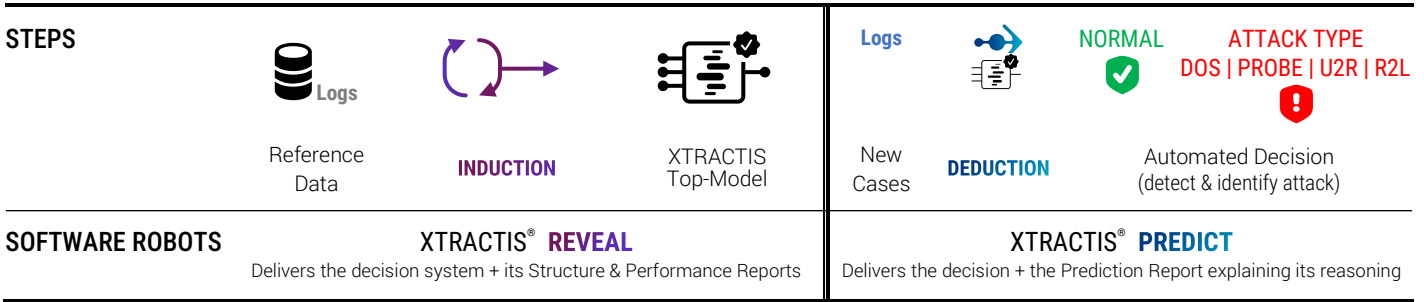
Learning Dataset: 859,984 logs 80%					ETD#1: 214,999 logs 20%					ETD#2: 70 874 logs				
70% for Training, 15% for Validation, 15% for Test														
NORMAL	DOS	PROBE	U2R	R2L	NORMAL	DOS	PROBE	U2R	R2L	NORMAL	DOS	PROBE	U2R	R2L
75.61%	23.00%	1.29%	0.01%	0.09%	75.61%	23.00%	1.29%	0.01%	0.09%	67.13%	24.77%	3.68%	0.30%	4.11%

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 **36 gradual rules without chaining, each rule uses some of the 27 variables that XTRACTIS identified as predictors.** Moreover, 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), slightly degraded in ETD#2.
- ☑ **Efficient AI System** It computes real-time predictions up to 70,000 decisions/second, offline or online (API).

XTRACTIS PROCESS



TOP-MODEL INDUCTION

INDUCTION PARAMETERS

- We launch 860 inductive reasoning strategies; each strategy is applied to the same single partition of the learning dataset (70% Training / 15% Validation / 15% Test) to get a reliable assessment of the descriptive and predictive performances, respectively from Training and Validation Datasets.
- Each strategy thus generates one unitary model called **Individual Virtual Expert (IVE)**.
- Among the 860 induced models, the top-IVE is the one that has the best predictive performance, close to its descriptive performance, and with the fewer predictors and rules: **36 rules sharing 27 predictors**.

Powered by:



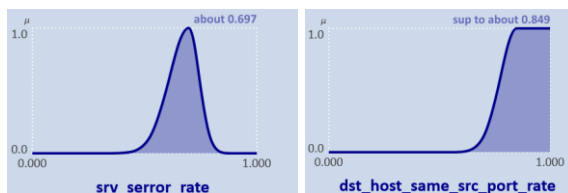
Total number of induced unitary models	Criterion for the induction optimization	Validation criterion for the top-model selection	Duration of the process (Induction Power FP64)
860 IVEs	Average F₁-Score	Average F₁-Score	14.3 days (24 Tflops)

TOP-MODEL STRUCTURE

The top-model has an excellent intelligibility, given the problem complexity, as it has 36 rules combining the 27 predictors that XTRACTIS automatically selected out of 41 variables. The 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

- 27 log characteristics (out of 41)
- 22 continuous + 5 nominal variables
- Ranked by impact significance (4 strong, 7 medium & 16 weak signals):
#1 `service` / #2 `src_bytes_1450Clip` / ...
- Labeled by fuzzy and binary classes
Examples: **binary interval** "inf to 0.697";
fuzzy interval "inf to about 0.849"



RULES

- 36 connective fuzzy rules without chaining (aggregated into 5 disjunctive fuzzy rules)
- 1 to 6 predictors per rule (on average, 2.8 predictors per rule)
- Example: **fuzzy rule R29** uses 2 predictors to conclude "R2L". 35 other fuzzy rules complete this model.

```

IF  srv_error_rate           IS  ~0.697
AND dst_host_same_src_port_rate IS  sup to ~0.849
THEN Connection             IS  R2L
    
```

Literally, the connection is a R2L Attack if the percentage of connections with SYN errors, among connections to the same service, during the last 2 seconds is around 70% and the percentage of connections from the same source port number, among connections to the same service during the last connections preceding the current connection, to the same service is superior to around 85%.

TOP-MODEL PERFORMANCE

The top-IVE performances, measured in Training/Validation/Test, then in External Test on ETD#1 and ETD#2, guarantee the model's predictive and real performances.

Performance Dataset
Average F₁-Score
Classification Error

DESCRIPTIVE	PREDICTIVE	REAL
70% Training	15% Validation	15% Test
96.27%	96.67%	86.69%
0.21%	0.20%	0.22%



EXPLAINED PREDICTIONS FOR 4 UNKNOWN CASES

NEW CASE

(from the External Dataset, i.e., not included in the Learning Dataset)

LOG V_289657

actual value = NORMAL

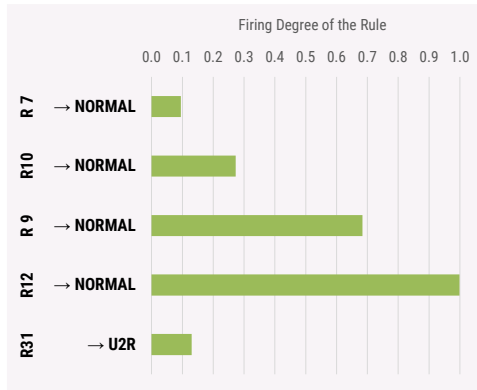
Urgent	0.0
num_access_files	0.00
svr_serror_rate	0.000
svr_rerror_rate	0.000
same_srv_rate	1.000
diff_srv_rate	0.000
dst_host_count	22
dst_host_srv_count	255
dst_host_same_srv_rate	1.000
dst_host_diff_srv_rate	0.000
dst_host_same_src_port_rate	0.050
dst_host_srv_diff_host_rate	0.010
dst_host_serror_rate	0.000
dst_host_srv_serror_rate	0.000
dst_host_rerror_rate	0.000
duration_3Clip	0.00
src_bytes_1450Clip	293
dst_bytes_11111Clip	1,628
hot_6Clip	0.00
num_compromised_5Clip	0.00
num_root_10Clip	0.00
svr_count_35Clip	23.0
service	http
flag	SF
logged_in	Yes
root_shell	No
guest_login	No



DEDUCTIVE INFERENCE OF RULES

For this connection, 5 rules are triggered:

R12 at 0.999, R9 at 0.684, R10 at 0.273, R7 at 0.095 to conclude {NORMAL}, R31 at 0.130 to conclude {U2R}. The 31 other rules are not activated.



AUTOMATED DECISION

NUMBER OF TRIGGERED RULES

5 / 36

FUZZY PREDICTION

{ NORMAL | 0.999, U2R | 0.130 }

FINAL PREDICTION

{ NORMAL }

The system delivers the correct diagnosis compared to that given by the cyber expert:

NORMAL 

LOG V_307879

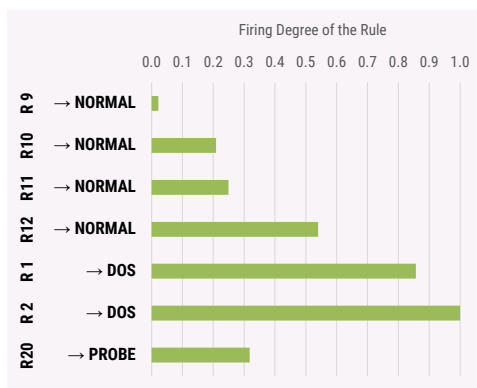
actual value = DOS ATTACK

Urgent	0.0
num_access_files	0.00
svr_serror_rate	1.000
svr_rerror_rate	0.000
same_srv_rate	0.040
diff_srv_rate	0.060
dst_host_count	255
dst_host_srv_count	8
dst_host_same_srv_rate	0.030
dst_host_diff_srv_rate	0.070
dst_host_same_src_port_rate	0.000
dst_host_srv_diff_host_rate	0.000
dst_host_serror_rate	1.000
dst_host_srv_serror_rate	1.000
dst_host_rerror_rate	0.000
duration_3Clip	0.00
src_bytes_1450Clip	0
dst_bytes_11111Clip	0
hot_6Clip	0.00
num_compromised_5Clip	0.00
num_root_10Clip	0.00
svr_count_35Clip	8.0
service	http
flag	S0
logged_in	No
root_shell	No
guest_login	No



For this connection, 7 rules are triggered:

R2 at 1.000, R1 at 0.857 to conclude {DOS}, R12 at 0.540, R11 at 0.249, R10 at 0.209, R9 at 0.022 to conclude {NORMAL}, R20 at 0.318 to conclude {PROBE}. The other 29 rules are not triggered.



NUMBER OF TRIGGERED RULES

11 / 36

FUZZY PREDICTION

{ DOS | 1.0, NORMAL | 0.540, PROBE | 0.318 }

FINAL PREDICTION

{ DOS }

The system delivers the correct diagnosis compared to that given by the cyber expert:

DOS ATTACK 

NEW CASE

(from the External Dataset, i.e., not included in the Learning Dataset)

log V_98996

actual value = R2L ATTACK

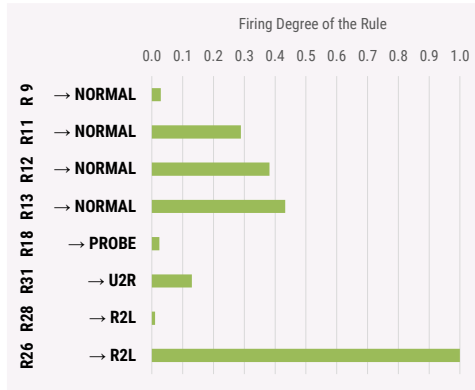
urgent	0
num_access_files	0
srv_serror_rate	0
svr_rerror_rate	0
same_srv_rate	1
diff_srv_rate	0
dst_host_count	4
dst_host_srv_count	10
dst_host_same_srv_rate	1
dst_host_diff_srv_rate	0
dst_host_same_src_port_rate	1
dst_host_srv_diff_host_rate	0.3
dst_host_serror_rate	0
dst_host_srv_serror_rate	0
dst_host_rerror_rate	0
duration_3Clip	0
src_bytes_1450Clip	12
dst_bytes_11111Clip	0
hot_6Clip	0
num_compromised_5Clip	0
num_root_10Clip	0
srv_count_35Clip	3
service	ftp_data
flag	SF
logged_in	No
root_shell	No
guest_login	No



Real Time

For this connection, 8 rules are triggered:

R26 at 1.000, R28 at 0.010 to conclude {R2L},
 R13 at 0.433, R12 at 0.382, R11 at 0.289
 and R9 at 0.029 to conclude {NORMAL},
 R31 at 0.130 to conclude {U2R},
 R18 at 0.024 to conclude {PROBE},
 The 28 other rules are not activated.



NUMBER OF TRIGGERED RULES

8/ 36

FUZZY PREDICTION

{ R2L | 1.000,
 NORMAL | 0.433,
 U2R | 0.130,
 PROBE | 0.024 }

FINAL PREDICTION

{ R2L }

The system delivers the correct diagnosis compared to that given by the cyber expert:

R2L ATTACK !

log V_91583

actual value = U2R ATTACK

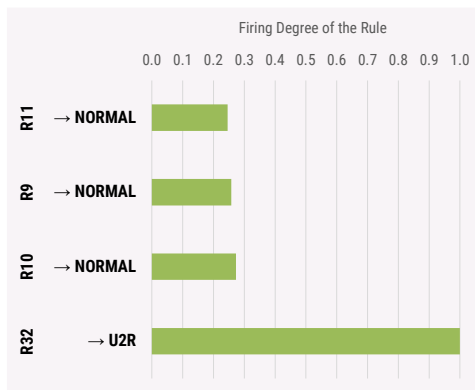
urgent *	3.0
num_access_files	0
srv_serror_rate	0
svr_rerror_rate	0
same_srv_rate	1
diff_srv_rate	0
dst_host_count	18
dst_host_srv_count	4
dst_host_same_srv_rate	0.22
dst_host_diff_srv_rate	0.17
dst_host_same_src_port_rate	0.06
dst_host_srv_diff_host_rate	0
dst_host_serror_rate	0
dst_host_srv_serror_rate	0
dst_host_rerror_rate	0.06
duration_3Clip	3
src_bytes_1450Clip	1,450
dst_bytes_11111Clip	11,111
hot_6Clip	6
num_compromised_5Clip	5
num_root_10Clip	10
srv_count_35Clip	1
service	telnet
flag	SF
logged_in	Yes
root_shell	Yes
guest_login	No



Real Time

For this connection, 4 rules are triggered:

R32 at 1.000 to conclude {U2R},
 R10 at 0.273, R9 at 0.258 and
 R11 at 0.246 to conclude {NORMAL},
 The other 32 rules are not triggered.



NUMBER OF TRIGGERED RULES

4/ 36

FUZZY PREDICTION

{ U2R | 1.0,
 NORMAL | 0.273 }

FINAL PREDICTION


{ U2R }

The system delivers the correct diagnosis compared to that given by the cyber expert:

U2R ATTACK !

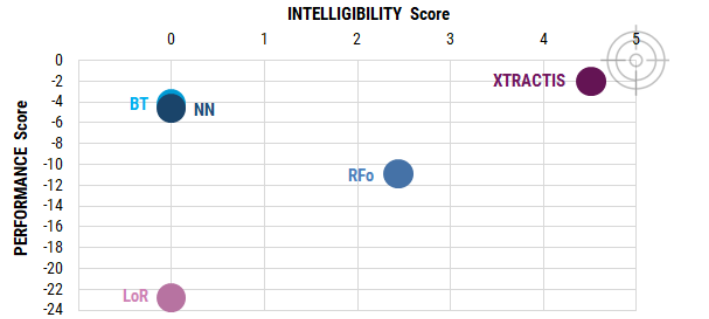
*Predictor value outside the variation range of the model but inside the allowed extrapolation range. Xtractis will refuse to give a result for an extrapolation far from the allowed extrapolation range. It is one case of the "Refusal" prediction.

TOP-MODELS BENCHMARK

	XTRACTIS 	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREES	NEURAL NETWORK	
MODELING PARAMETERS	MODELS RELEASE	2022/08	2022/10	2022/08	2022/08	2022/09
	ALGORITHM VERSION	XTRACTIS REVEAL 12.2.43016	Python 3.7 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	All explored strategies for all algorithms use the same single-split of the Learning Dataset: 70% Training 15% Validation 15% Test				
	NUMBER OF EXPLORED STRATEGIES⁽¹⁾	860 induction strategies	1,000 data analysis strategies	1,000 ML strategies	1,000 ML strategies	1,000 ML strategies
	TOP-MODEL SELECTION⁽²⁾	Top-IVE among 860 IVEs	Top-IVE among 1,000 IVEs	Top-IVE among 1,000 IVEs	Top-IVE among 1,000 IVEs	Top-IVE among 1,000 IVEs
TOP-MODEL STRUCTURE	NUMBER OF PREDICTORS (out of 41 Potential Predictors)	27	39	30	23	122 3 nominal variables are decomposed into 84 binary variables
	AVERAGE NUMBER OF PREDICTORS PER RULE / EQUATION	2.8 per rule	35.8 per equation	6.7 per rule	3.7 per rule	107.2 per equation
	STRUCTURE OF THE DECISION SYSTEM	36 fuzzy rules without chaining (aggregated into 5 disjunctive fuzzy rules) Only a few rules are triggered at a time to compute a decision	5 linear equations	5 trees without chaining 252 binary rules	5 chains of 325 trees each 16,982 binary rules Tree #N corrects the error of the N-1 previous trees	1 hidden layer 27 hidden nodes 32 equations 27 unintelligible synthetic variables

INTELLIGIBILITY × PERFORMANCE SCORES (Performance Score is calculated on all available unknown data)

	Random ⁽³⁾	XTRACTIS	LoR	RFo	BT	NN
INTELLIGIBILITY Score⁽⁴⁾		4.51	0.00	2.44	0.00	0.00
IVE Real Perf. (F ₁ -Score) in Test		86.69	61.35	73.05	85.00	85.94
Gap to Leader in Test		0.00	-25.34	-13.64	-1.69	-0.75
IVE Real Perf. (F ₁ -Score) in External Test #1	20.33	92.72	62.02	73.54	91.87	88.28
Gap to Leader in External Test #1		0.00	-30.7	-19.18	-0.85	-4.44
IVE Real Perf. (F ₁ -Score) in External Test #2	20.51	60.77	54.29	66.77	56.62	58.05
Gap to Leader in External Test #2		-6.00	-12.48	0.00	-10.15	-8.72
IVE Average Real Performance	20.42	80.06	59.22	71.12	77.83	77.42
PERFORMANCE Score⁽⁴⁾		-2.00	-22.84	-10.94	-4.23	-4.64



(1) For all algos: on the same Learning Dataset. All Models are optimized according to their validation Average F₁-Score.
 (2) All top-models are selected according to their validation Average F₁-Score while checking that it remains close to their training Average F₁-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.

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

APPENDIX 1 – Calculation of the Intelligibility × Performance

AI Technique #i	T _i	i ∈ [1 ; n] n = number of AI Techniques benchmarked in terms of data-driven modeling = 5
Benchmark #k	B _k	k ∈ [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 & Boost Trees: 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 target is to obtain 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; F₁-Score for a Binomial Classification; Average F₁-Score or Average F₂-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).

$$\text{Performance Score of } T_i$$

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

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):
 $\text{Pen1}(T_i) = \min(0, 1 - \log_{10} \text{number of predictors})$
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):
 $\text{Pen2}(T_i) = \min(0, 0.01 - \frac{\text{average number of rules or equations per modality to predict}}{100})$
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):
 $\text{Pen3}(T_i) = \min(0, \frac{9 - 3 \times \text{average number of predictors per rule or equation}}{7})$
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):
 $\text{Pen4}(T_i) = \min(0, 1 - \text{number of chained trees})$
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):
 $\text{Pen5}(T_i) = -5$

$$\text{Intelligibility Score of } T_i$$

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

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 T_i natively produces intelligible models (XTRACTIS, Random Forests).
- For similar structures, the Boosted Trees model is always less intelligible than the Random Forest one, as it is composed of chains of trees, instead of a college of trees (cf. 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 (cf. 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. F ₁ -Score	Average F ₁ -Score	Weighted Av. F ₁ -Score	Refusal
RANDOM MODEL									
<i>Nb of Random Permutations (P-value) = 100,000 (0.001%)</i>									
Performance against chance in External Test 1	37.29%	0.50%	20.33%	0.50%	20.33%	0.50%	20.33%	62.71%	
Performance against chance in External Test 2	48.05%	2.33%	20.51%	2.33%	20.51%	2.33%	20.51%	51.95%	
XTRACTIS TOP-MODEL									
Descriptive Performance (Training)	0.21%	79.31%	93.91%	97.47%	99.15%	88.46%	96.27%	99.79%	0 (0.00%)
Predictive Performance (Validation)	0.20%	90.83%	97.58%	85.71%	96.04%	92.31%	96.67%	99.80%	0 (0.00%)
Real Performance (Test)	0.22%	50.00%	87.43%	37.50%	86.36%	42.86%	86.69%	99.78%	0 (0.00%)
Real Performance (External Test 1)	0.22%	63.64%	89.79%	87.50%	96.51%	73.68%	92.72%	99.78%	1 (0.00%)
Real Performance (External Test 2)	6.46%	5.58%	57.62%	80.00%	90.11%	10.43%	60.77%	92.11%	0 (0.00%)
LOGISTIC REGRESSION TOP-MODEL									
Descriptive Performance (Training)	7.94%	73.35%	90.71%	0.07%	57.96%	0.14%	62.08%	95.34%	
Predictive Performance (Validation)	8.02%	71.67%	90.32%	0.07%	57.24%	0.14%	61.37%	95.28%	
Real Performance (Test)	8.06%	70.00%	90.05%	0.07%	57.29%	0.14%	61.35%	95.26%	
Real Performance (External Test 1)	7.97%	70.00%	90.14%	0.08%	58.05%	0.15%	62.02%	95.33%	
Real Performance (External Test 2)	11.62%	3.19%	62.98%	2.22%	67.63%	4.23%	54.29%	89.36%	
RANDOM FOREST TOP-MODEL									
Descriptive Performance (Training)	0.25%	93.10%	98.16%	6.59%	69.63%	12.30%	74.42%	99.80%	
Predictive Performance (Validation)	0.28%	16.67%	82.47%	0.90%	68.57%	1.70%	72.27%	99.78%	
Real Performance (Test)	0.28%	66.67%	92.67%	4.49%	68.60%	8.42%	73.05%	99.77%	
Real Performance (External Test 1)	0.26%	63.64%	91.62%	5.18%	69.15%	9.59%	73.54%	99.79%	
Real Performance (External Test 2)	7.24%	10.23%	65.10%	37.29%	74.96%	16.06%	66.77%	92.57%	
BOOSTED TREES TOP-MODEL									
Descriptive Performance (Training)	0.00%	96.55%	99.31%	99.99%	100.00%	98.25%	99.65%	100.00%	
Predictive Performance (Validation)	0.01%	95.00%	98.91%	85.71%	97.13%	92.31%	97.90%	99.99%	
Real Performance (Test)	0.02%	33.33%	85.45%	25.00%	84.82%	28.57%	85.00%	99.98%	
Real Performance (External Test 1)	0.02%	54.55%	89.46%	75.00%	94.97%	63.16%	91.87%	99.98%	
Real Performance (External Test 2)	7.93%	3.26%	53.52%	84.83%	92.13%	6.28%	56.62%	90.29%	
NEURAL NETWORK TOP-MODEL									
Descriptive Performance (Training)	0.05%	34.48%	84.26%	83.33%	95.19%	48.78%	87.68%	99.95%	
Predictive Performance (Validation)	0.06%	50.00%	87.01%	75.00%	93.47%	60.00%	89.72%	99.94%	
Real Performance (Test)	0.06%	50.00%	85.96%	37.50%	86.57%	42.86%	85.94%	99.94%	
Real Performance (External Test 1)	0.06%	45.45%	85.15%	71.43%	92.71%	55.56%	88.28%	99.94%	
Real Performance (External Test 2)	7.84%	3.72%	55.58%	70.76%	83.10%	7.08%	58.05%	90.79%	

The entirety of this document is protected by copyright. All rights are reserved, particularly the rights of reproduction and distribution. Quotations from any part of the document must necessarily include the following reference:
Zalila, Z., Intellictech & Xtractis (2014-2024). XTRACTIS® the Reasoning AI for Trusted Decisions. Use Case | Cyber Security: Log-based Identification of Cyber Intrusions (DARPA) – Benchmark vs. Logistic Regression, Random Forests, Boosted Trees & Neural Networks. INTELLITECH [intelligent technologies], February 2024, v3.0, Compiègne, France, 7p.