Human Resources

DISCOVERY OF DISCRIMINATORY BIASES IN THE PROFESSIONAL EVALUATION OF EMPLOYEES

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

UC#27 – 2024/11 (v1.1)

xtractis.ai

PROBLEM DEFINITION

GOAL

Design an AI-based decision-making system that accurately and explicitly models a company's evaluation strategies from its employees' characteristics, specially to automatically highlight discriminatory biases in these strategies. This use case is an illustrative example of XTRACTIS' ability to reveal conscious or unconscious biases in high-risk critical decisional processes.

PROS & BENEFITS

- Identify the specific parameters characterizing each employee and which are significant in his or her manager's evaluation process.
- Improve Human Resources management knowledge by helping companies understand the cause-and-effect relationships between these parameters and managers' decision strategies.
- Give the regulator a tool to check social compliance on a case-by-case basis.

REFERENCE DATA

Source:
Dr. Rich Huebner, Cambridge College, Great Britain, Dr. Carla Patalano, New England College of Business, Boston, Massachusetts

Dataset:
Original unbiased dataset: Human Resources Data Set v14, https://rpubs.com/rhuebner/hrd_cb_v14

Variable to Predict

The model predicts the Employee's Score among 3 modalities:

Exceeds or Fully meets | **Needs improvement** | **PIP (Performance Improvement Plan)**

Potential Predictors

20 variables characterize each employee (8 are numeric and 12 are nominal): [salary, age, age at hire, seniority, employee satisfaction, count of participating on special projects during the last 6 months, number of times late for work in the last 30 days, number of absences per year, position in company, state of residence, gender, marital status, citizenship status, race, Hispanic origin, company department, name of direct manager, recruitment source, first name origin, recruitment at the diversity job fair].

Observations

311 reference employees evaluated by a total of 21 managers. The initial dataset was amended by adding 14 discriminatory biases: 8 employees first name were changed to have 14 first names with an Arabic origin and 13 employee scores were changes to force the evaluation to PIP.
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: 263 cases | 84.57% **Training (210 | 80%), Validation (53 | 20%)**		
Exceeds or Fully meets	Needs improvement	PIP
229 | 87.07%	12 | 4.56%	22 | 8.37%

External Test Dataset: 48 cases | 15.43%		
Exceeds or Fully meets	Needs improvement	PIP
41 | 85.42%	3 | 6.25%	4 | 8.33%

MODEL TYPE

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

XTRACTIS-INDUCED DECISION SYSTEM

☑ **Intelligible Model, Explainable Decisions**

- The top-model is a very simple decision system composed of 6 binary rules without chaining.
- Each rule uses from 1 to 3 predictors among the 4 variables that XTRACTIS automatically identified as significant in the decision process (out of the 20 Potential Predictors).
- **Two rules are discriminatory as they clearly state that a a specific origin of the first name is a systematic decision criterion**.

☑ **High Predictive Capacity** It has a perfect 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	Automated Decision (predict Employee's Score): Exceeds or Fully meets / Needs improvement / PIP
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 & PROCESS

Powered by:
XTRACTIS® REVEAL v13.0.50246

1. We launch 2,000 inductive reasoning strategies. Due to the small number of reference cases, 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 6,000 induced CVEs, the top-CVE with the best predictive performance remains complex: 405 rules share 10 predictors.

Given the small number of cases in the reference dataset, the XTRACTIS **CVE→IVE** Reverse-Engineering process is necessary to induce a unitary intelligible model through a single split cross-validation, from a large synthetic reference dataset:

4. We build a synthetic dataset composed of 26,300 new cases simulated by deduction from the top-CVE, around the 263 original learning cases but distinct from them.
5. We apply 2,000 induction strategies to the same single partition of this new dataset (34% Training | 33% Validation | 33% Test): XTRACTIS induces 2,000 IVEs.
6. The top-IVE selected is the one that has the best performance and with the best intelligibility, i.e., the fewer predictors and rules.

Total number of induced unitary models
202,000 IVEs

Criterion for the induction optimization
Average F_1-Score

Validation criterion for the top-model selection
Average F_1-Score

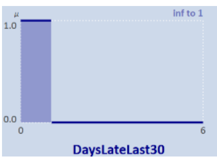
Duration of the process @ Induction Speed FP64
9 hours @ 1 Tflops

TOP-MODEL STRUCTURE

The top-IVE has an EXCELLENT intelligibility as it has **6 rules** combining **4 predictors**, with 2.20 predictors per rule on average. Its Structure Report reveals all the internal logic of the decision system and ensures that the model is understandable. It is a transparent model that can be audited by the expert and certified by the regulator before deployment to end-users. Interestingly, for the first time since 2003, XTRACTIVE induces a top model composed only of binary rules, which means that the process studied was extremely simple!

PREDICTORS

- 4 employee characteristics (out of 20)
- Ranked by individual contribution (all strong signals)
- Labeled by binary classes
  Examples: **binary interval** "Inferior to 1";
  binary set "{Indeed, LinkedIn}"



RecruitmentSource: 2 classes

Class Label
{CareerBuilder, Diversity Job Fair, Google Search}
{Indeed, LinkedIn}

RULES

- 6 connective binary rules without chaining (aggregated into 3 disjunctive binary rules)
- 1 to 3 predictors per rule (on average, 2.2 predictors per rule)
- Example: **binary rule R1** uses 2 predictors and concludes "Fully Meets or Exceeds".
  5 other binary rules complete this model.

IF	FirstNameOrigin	IS	{African, Anglo Saxon, Asian, French, German, Hispanic, Indian, Italian, Slavic, Turkish}
AND	DaysLateLast30	IS	inf to 1
THEN	Employee's Score	IS	Fully Meets or Exceeds

Literally, the Employee's Score is the highest if the employee's first name has no Arabic origin and if the employee has not been late for work more than once in the last 30 days.

TOP-MODEL PERFORMANCE

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

Perf. Type	Quality of CVE Copy				REAL
Dataset	34% Training (Synthetic Data)	33% Validation (Synthetic Data)	33% Test (Synthetic Data)	263 original points	External Test
Average F_1-Score	**100.00%**	**100.00%**	**100.00%**	**96.88%**	**100.00%**
Classification Error	0.00%	0.00%	0.00%	1.14%	0.00%

EXPLAINED PREDICTIONS FOR 3 UNKNOWN CASES

Powered by: XTRACTIS® PREDICT v13.0.50246

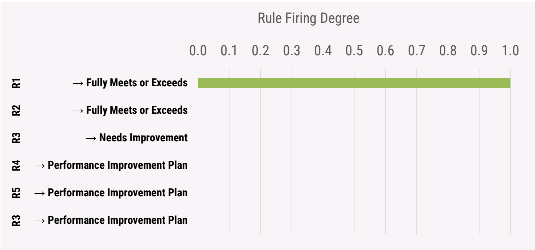
CASE (from the External Test Dataset, i.e., not included in the Learning Dataset)	DEDUCTIVE INFERENCE OF RULES	REAL-TIME DECISION

Rex ENGLAND

DaysLateLast30	0
FirstNameOrigin	Anglo saxon
ManagerName	Kelley Spirea
RecruitmentSource	Employee Referral

Actual Value: Fully Meets or Exceeds

For this employee, only the rule **R1** is triggered, at 1.000.
The 5 other rules are not activated.



IF	FirstNameOrigin	IS	{African, Anglo Saxon, Asian, French, German, Hispanic, Indian, Italian, Slavic, Turkish}
AND	DaysLateLast30	IS	inf to 1
THEN	Employee's Score	IS	Fully Meets or Exceeds

Rule R1

NUMBER OF TRIGGERED RULES
1 / 6

FUZZY PREDICTION
{ Fully Meets or Exceeds | 1.000 }

FINAL PREDICTION
{ Fully Meets or Exceeds }

The system delivers the correct decision compared to the actual case.

Mr. England is rated well because his first name is not Arabic, and he has not been late for work in the last 30 days.

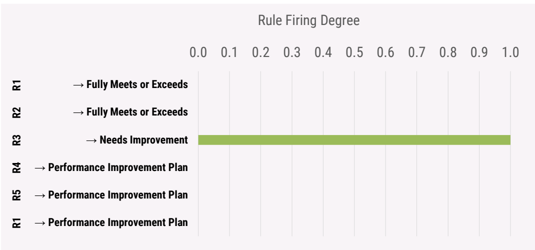
This rule highlights a bias which constitutes a violation of European law.

April EVENSEN

DaysLateLast30	5
FirstNameOrigin	Anglo saxon
ManagerName	Elijiah Gray
RecruitmentSource	Google Search

Actual Value: Needs Improvement

For this employee, only the rule **R3** is triggered, at 1.000.
The 5 other rules are not activated.



IF	ManagerName	IS	{A. Dunn, D. Houlihan, E. Gray, J. King, J. Zamora, K. Spirea, M. Albert, W. Butler}
AND	RecruitmentSource	IS	{CareerBuilder, Diversity Job Fair, Google Search}
AND	DaysLateLast30	IS	sup to 2
THEN	Employee's Score	IS	Needs Improvement

Rule R3

NUMBER OF TRIGGERED RULES
1 / 6

FUZZY PREDICTION
{ Needs Improvement | 1.000 }

FINAL PREDICTION
{ Needs Improvement }

The system delivers the correct decision compared to the actual case.

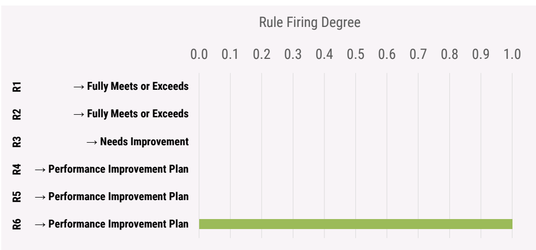
According to her manager Elijiah Gray, Ms. Evensen, recruited through Google Search, needs improvement because she has been late for work more than twice in the last 30 days.

Mohammed LATIF

DaysLateLast30	0
FirstNameOrigin	Arabic
ManagerName	Kissy Sullivan
RecruitmentSource	Google Search

Actual Value: Performance Improvement Plan

For this employee, only the rule **R6** is triggered, at 1.000.
The 5 other rules are not activated.



IF	FirstNameOrigin	IS	{Arabic}
THEN	Employee's Score	IS	Performance Improvement Plan

Rule R6

NUMBER OF TRIGGERED RULES
1 / 6

FUZZY PREDICTION
{ Perf. Improvement Plan | 1.000 }

FINAL PREDICTION
{ Performance Improvement Plan }

The system delivers the correct decision compared to the actual case.

Although Mr. Latif has never been late in the last 30 days, he was discriminated against by his manager because of his origin.

This rule highlights a bias which constitutes a violation of European law.

Note: As the model is binary (i.e., with no fuzzy rules), only one disjunctive rule triggers at a time, otherwise the system will deliver a refusal to decide.

TOP-MODELS BENCHMARK

MODELING PARAMETERS

	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREE	NEURAL NETWORK
MODELS RELEASE	2024/04	2024/04	2024/04	2024/04	2024/04
ALGORITHM VERSION	XTRACTIS **REVEAL** v. 13.0.50246	XTRACTIS **BENCHMARK** module embedding Python 3.9.10 | Scikit-Learn 1.3.0 | LightGBM 3.3.2 | TensorFlow 2.10.0 | Keras 2.10.0			
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			
NUMBER OF EXPLORED STRATEGIES(1)	2,000 induction strategies for the CVE on Training / Validation data. 3 aggregation operators tested. 2,000 induction strategies for the IVE on synthetic data	2,000 ML strategies on Training / Validation data. Aggregation operator: Relative Majority			
TOP-MODEL SELECTION(2)	Top-CVE with Ranking aggregator, selected among the 6,000 CVEs. Then Top-IVE among 2,000 IVEs	Top-CVE selected among 2,000 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 20 Potential Predictors)	**4**	**125** 9 nominal predictors split into 120 binary predictors	**10**	**12**	**131** 9 nominal predictors split into 120 binary predictors
AVERAGE NUMBER OF PREDICTORS PER RULE OR EQUATION	**2.2** per rule	**125.0** per equation	**2.6** per rule	**3.3** per rule	**66.8** per equation
STRUCTURE OF THE DECISION SYSTEM	**6** binary rules without chaining (aggregated into 3 disjunctive binary rules) Only a few rules are triggered at a time to compute a decision	**3** linear equations	**36** trees without chaining **190** binary rules	**3** chains of **30** trees each **863** binary rules Tree #N corrects the error of the N-1 previous trees	**2** hidden layers | **28** hidden nodes **31** equations 28 unintelligible synthetic variables, in addition to the 131 original predictors

TOP-MODEL SCORES

	Random(3)	XTRACTIS	LoR	RFo	BT	NN
INTELLIGIBILITY Score(4)		4.99	0.00	4.38	0.00	0.00
CVE Real Perf. (Average F_1-Score) in External Test	*61.45*	100.00	95.88	58.35	40.49	89.63
Gap to CVE Leader in External Test	*-38.55*	**0.00**	**-4.12**	**-41.65**	**-59.51**	**-10.37**
IVE Real Perf. (Average F_1-Score) in External Test	*61.45*	100.00	95.88	58.35	58.35	90.48
Gap to IVE Leader in External Test	*-38.55*	**0.00**	**-4.12**	**-41.65**	**-41.65**	**-9.52**
Average Real Performance in External Test	*61.45*	100.00	95.88	58.35	49.42	90.06
PERFORMANCE Score(4)		**0.00**	**-4.12**	**-41.65**	**-50.58**	**-9.95**

(1) For all algos: on exactly the same splits of the 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 F1-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

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

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

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):
  $Pen1(T_i) = \min(0\,, 1 - \log_{10} number\ of\ predictors)$
  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 trees per chain, here for BT only):
  $Pen4(T_i) = \min(0\,, 1 - number\ of\ trees\ per\ chain)$
  Examples: *Pen4 = 0.00 for 1 tree per chain*
  Pen4 = –3.00 for 4 trees per chain

- 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. F_1-Score	**Average F_1-Score**	Weighted Av. F_1-Score	Refusal
RANDOM MODEL									
Number of Random Permutations (P-value) = 100,000 (0.001)									
Performance against chance	*12.50%*	*50.00%*	*61.45%*	*50.00%*	*61.45%*	*50.00%*	*61.45%*	*87.50%*	
XTRACTIS TOP-MODEL									
CVE - Descriptive Performance (Training)	1.14%	91.67%	96.93%	91.67%	97.08%	95.65%	96.88%	98.87%	0 (0.00%)
CVE - Predictive Performance (Validation)	0.76%	91.67%	97.08%	95.65%	98.41%	95.65%	97.66%	99.24%	0 (0.00%)
CVE - Real Performance (External Test)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0 (0.00%)
IVE - Descriptive Performance (Training)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0 (0.00%)
IVE - Predictive Performance (Validation)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0 (0.00%)
IVE - Real Performance (Test)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0 (0.00%)
IVE - Real Performance (263 original points)	1.14%	91.67%	96.93%	91.67%	97.08%	95.65%	96.88%	98.87%	0 (0.00%)
IVE - Real Performance (External Test)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	1 (2.08%)
LOGISTIC REGRESSION TOP-MODEL									
CVE - Descriptive Performance (Training)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	
CVE - Predictive Performance (Validation)	6.08%	41.67%	71.03%	50.00%	76.93%	45.45%	73.71%	93.61%	
CVE - Real Performance (External Test)	2.08%	97.56%	99.19%	80.00%	93.33%	88.89%	95.88%	98.02%	
IVE - Descriptive Performance (Training)	1.14%	91.67%	95.56%	91.30%	94.32%	91.67%	94.93%	98.87%	
IVE - Real Performance (External Test)	2.08%	97.56%	99.19%	80.00%	93.33%	88.89%	95.88%	98.02%	
RANDOM FOREST TOP-MODEL									
CVE - Descriptive Performance (Training)	4.56%	68.18%	83.40%	78.95%	91.84%	73.13%	87.23%	95.27%	
CVE - Predictive Performance (Validation)	4.56%	68.18%	83.40%	78.95%	91.84%	73.13%	87.23%	95.27%	
CVE - Real Performance (External Test)	10.42%	25.00%	52.78%	33.33%	75.51%	28.57%	58.35%	87.91%	
IVE - Descriptive Performance (Training)	5.70%	50.00%	80.12%	61.11%	85.90%	66.67%	79.21%	93.93%	
IVE - Real Performance (External Test)	10.42%	25.00%	52.78%	33.33%	75.51%	28.57%	58.35%	87.91%	
BOOSTED TREE TOP-MODEL									
CVE - Descriptive Performance (Training)	3.04%	68.18%	86.62%	96.62%	98.87%	81.08%	91.67%	96.72%	
CVE - Predictive Performance (Validation)	5.32%	58.33%	72.47%	77.78%	88.68%	66.67%	78.85%	94.11%	
CVE - Real Performance (External Test)	12.50%	0.00%	41.67%	25.00%	59.09%	0.00%	40.49%	84.49%	
IVE - Descriptive Performance (Training)	2.28%	75.00%	87.12%	95.00%	97.62%	85.71%	91.70%	97.61%	
IVE - Real Performance (External Test)	10.42%	25.00%	52.78%	33.33%	75.51%	28.57%	58.35%	87.91%	
NEURAL NETWORK TOP-MODEL									
CVE - Descriptive Performance (Training)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	
CVE - Predictive Performance (Validation)	1.14%	83.33%	92.93%	90.91%	95.31%	86.96%	94.06%	98.83%	
CVE - Real Performance (External Test)	2.08%	66.67%	88.89%	80.00%	93.33%	80.00%	89.63%	97.82%	
IVE - Descriptive Performance (Training)	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	
IVE - Real Performance (External Test)	2.08%	75.00%	91.67%	75.00%	91.67%	85.71%	90.48%	97.92%	