HEALTH / PHARMA

SEROLOGICAL DIAGNOSIS OF CHRONIC KIDNEY DISEASE

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

Use Case 2023/02 (v1.1) • xtractis.ai

PROBLEM DEFINITION

PROBLEM

How to make an automated — yet totally transparent — medical diagnosis of chronic kidney disease, from the patient record and its blood measures?

GOALS & BENEFITS

- ☑ Identify the parameters involved in the kidney disease and enhance medical knowledge by helping nephrologists understand the causal relationships between these parameters, their combination, and the disease.
- ☑ Help the medical profession to make earlier and more personalized decisions through rapid, systematic, and explainable diagnoses.
- ☑ Use a model with few predictors to limit medical data that can be expensive to collect.

REFERENCE DATA

- **Observations**: 340 reference of patients for Training/Validation, without (128 | 37.7%) or with chronic kidney disease (212 | 62.3%), and 60 reference of patients for External Test, without (22 | 36.7%) or with chronic kidney disease (38 | 63.3%).
  Source: Dr. P.Eswaran, Department of Computer Science and Engineering, Alagappa University, Karaikudi, Tamilnadu, India [http://archive.ics.uci.edu/ml] (2015)
- **Predictive Variables**: 24 Potential Predictors characterizing each patient [Age, appetite, blood measures (Pressure, Specific Gravity, Albumin Sugar, Red Blood Cells, Pus Cell, Pus Cell Clumps, Bacteria, Blood Glucose, Blood Urea, Serum, Sodium, Potassium, Hemoglobin, Packed Cell Volume, White Blood Cell count, Red Blood Cell count), Hypertension, Diabetes, Coronary Artery Disease, Pedal Edema, Anemia].
- **Variable to Predict**: Chronic Kidney Disease diagnosis [No CKD / CKD].

MODEL TYPE

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

XTRACTIS SOLUTION

PROCESS

Reference Data → **INDUCTION + Reverse-Engineering**[1] → XTRACTIS Top-Model | New Cases → **DEDUCTION** → Automated Decision (supporting experts' diagnosis): No CKD / CKD

SOFTWARE ROBOTS

XTRACTIS® **GENERATE** | XTRACTIS® **PREDICT** | + **PREDICT**ion Report (for decision explainability)

RESULTS

- ☑ **Intelligible Predictive Top-Model**: Decision system composed of 4 unchained gradual rules, each rule using some of the 8 variables that XTRACTIS identified as predictors.
- ☑ **Robust Predictive Top-Model**: Perfect Real Performance in External Test.
- ☑ **Operational Efficient System**: Real-time predictions up to 70,000 decisions/s., offline or online (API).

TOP-MODEL INDUCTION

INDUCTION PARAMETERS

We launch 1,000 inductive reasoning strategies; each strategy is applied to 20 different 5-fold-partitions of the Training/Validation dataset to get a reliable assessment of the descriptive and predictive performances. 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)**. Among the 3,000 CVE, the top-CVE with the best predictive performance remains complex (12 predictors shared by 200 rules).

We then apply 2,000 induction strategies to the same single Training (34%)/Validation (33%)/Test (33%) partition of a synthetic dataset: 34,000 new cases simulated by deduction from the top-CVE, around the 340 cases but distinct from these original cases. This XTRACTIS Reverse-Engineering[1] process induces 2,000 IVE. The top-IVE selected is as efficient as the top-CVE, but intelligible (8 predictors shared by 4 rules).

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 IVE	**F_1-Score**	**F_1-Score**	**2 hours** (1 Tflops)

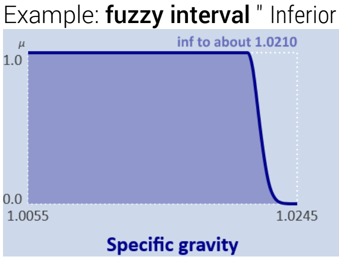
STRUCTURE

Intelligibility

The top-IVE model is very simple as it combines the 8 predictors automatically selected by XTRACTIS into 4 rules. Its Structure Report reveals all the internal decision logic and ensures that the human expert understands the model. This decision system is a *white-box* model that can be audited by the domain expert and certified by the regulator before its deployment to end-users.

PREDICTORS

- 8 features out of 24
- Ranked by impact significance (1 strong signal, 1 medium signal, 6 weak signals):
  #1 Specific gravity / #2 Hemoglobin
  #3 …
- Labeled by fuzzy classes.
  Example: **fuzzy interval** " Inferior to about 1.0210"



RULES

- 4 connective fuzzy rules without chaining
- 2 to 6 predictors per rule
  (on average, 3 predictors per rule)
- Example: **fuzzy rule R2** uses 2 predictors and concludes "Chronic Kidney Disease".
  3 other fuzzy rules complete this model.

IF	Specific gravity	IS	Inferior to about 1.0210
AND	Hypertension	IS	Yes
THEN	Diagnosis	IS	Chronic Kidney Disease

PERFORMANCE

Robustness

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.

	Synthetic Data			Reference Data
Performance	DESCRIPTIVE	PREDICTIVE	REAL	REAL
Dataset	34% Training	33% Validation	33% Test	External Test
F_1-Score	99.92 %	99.94%	99.90%	100.00%
Classification Error	0.11%	0.09%	0.14%	0.00%

Xtractis Top-Model: Intelligible AND Perfect Predictive Capacity

EXPLAINED PREDICTIONS FOR 2 CASES FROM THE EXTERNAL TEST SET

CASE
(not used in Training/Validation)

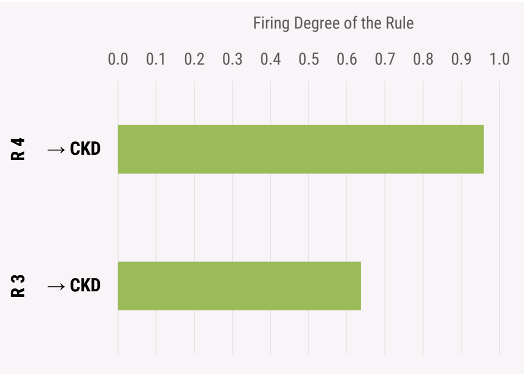
DEDUCTIVE INFERENCE OF RULES

AUTOMATED DECISION

PATIENT #30 (actual value = Chronic Kidney Disease)	
Specific gravity*	1.0050
Blood Glucose Random (mg/dL)	Missing Value
Hemoglobin (g/dL)	12.9
Packed Cell Volume (%)	38.0
Hypertension	No
Diabetes Melitus	No
Appetite	Yes
Pedal Edema	No

Real Time

For this patient, 2 rules are triggered
to conclude {Chronic Kidney Disease}:
R4 is fired at 0.959.
R3 is fired at 0.637.



NUMBER OF TRIGGERED RULES
2 / 4

FUZZY PREDICTION
{ Chronic Kidney Disease | 0.959 }

FINAL PREDICTION
{ Chronic Kidney Disease }

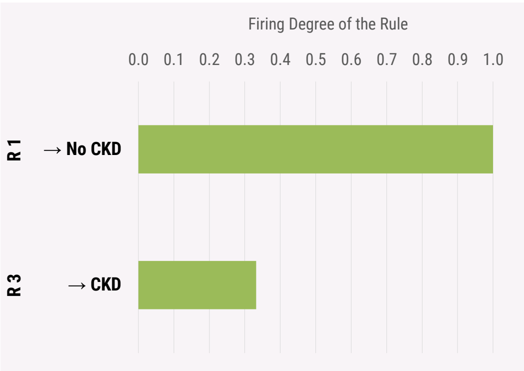
The system delivers a correct diagnosis compared to that given by the nephrologists:

Chronic Kidney Disease

PATIENT #305 (actual value = No Chronic Kidney Disease)	
Specific gravity*	1.0250
Blood Glucose Random (mg/dL)	128
Hemoglobin (g/dL)	13.1
Packed Cell Volume (%)	45.0
Hypertension	No
Diabetes Melitus	No
Appetite	Yes
Pedal Edema	No

Real Time

For this patient, 2 rules are triggered:
R1 is fired at 1.000
to conclude {No Chronic Kidney Disease}.
R3 is fired at 0.332
to conclude {Chronic Kidney Disease}.



NUMBER OF TRIGGERED RULES
2 / 4

FUZZY PREDICTION
{ No Chronic Kidney Disease | 1.000, Chronic Kidney Disease | 0.332 }

FINAL PREDICTION
{ No Chronic Kidney Disease }

The system delivers a correct diagnosis compared to that given by the nephrologists:

No Chronic Kidney Disease

**Predictor value is out of the variation range of the model (<2.63 % OOR for case #30 and case #305) 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-IVE BENCHMARK

	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREES	NEURAL NETWORK
MODELS RELEASE	2023/02	2023/02	2023/02	2023/02	2023/02
ALGO VERSION	XTRACTIS **GENERATE** 13.0.44983	Python 3.9, Scikit-Learn 1.1.2	Python 3.9, LightGBM 3.3.2	Python 3.9, LightGBM 3.3.2	Python 3.9, TensorFlow 2.10.0, Keras 2.10.0
CROSS-VALIDATION TECHNIQUE	20x5 folds for each CVE model Then 1-Split Validation for each IVE model (for the reverse engineering of top-CVE): 34% Training; 33% Validation; 33% Test	20x5 folds for each CVE model	20x5 folds for each CVE model	20x5 folds for each CVE model	20x5 folds for each CVE model
NUMBER OF EXPLORED STRATEGIES[2]	1,000 induction strategies for the CVE on Training / Validation data 2,000 induction strategies for the IVE on synthetic data	1,000 data analysis strategies on Training / Validation data	1,000 ML strategies on Training / Validation data	1,000 ML strategies on Training / Validation data	1,000 ML strategies on Training / Validation data
NUMBER OF MODELS	3,000 CVE + selection of the top-CVE 2,000 IVE (for the reverse engineering of top-CVE) + selection of the top-IVE	1,000 CVE + selection of the top-CVE 1 top-IVE	1,000 CVE + selection of the top-CVE 1 top-IVE	1,000 CVE + selection of the top-CVE 1 top-IVE	1,000 CVE + selection of the top-CVE 1 top-IVE
TOP-IVE STRUCTURE					
NUMBER OF PREDICTORS (out of 24 Potential Predictors)	8	10 out of 34 (2 nominal variables are decomposed into 12 binary variables)	9	9	34 out of 34 (2 nominal variables are decomposed into 12 binary variables)
DECISION STRUCTURE	System with 4 unchained fuzzy rules	1 linear equation with 10 coefficients.	9 trees; 51 binary rules	11 chained trees; 45 binary rules	1 hidden layer; 3 hidden nodes
MODEL INTELLIGIBILITY (& DECISION EXPLAINABILITY)	Only 3 predictors per rule on average Only a few rules are triggered at a time	A few predictors and coefficients	Many predictors and too many rules	Tree #N corrects the error of the N-1 previous trees	Unintelligible synthetic variables

TOP-IVE REAL PERFORMANCE (External Test)

	Random[3]	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREES	NEURAL NETWORK
Classification Error	23.33%	**0.00%**	**0.00%**	**1.67%**	**0.00%**	**0.00%**
Sensitivity		100.00%	100.00%	97.37%	100.00%	100.00%
Specificity		100.00%	100.00%	100.00%	100.00%	100.00%
PPV		100.00%	100.00%	100.00%	100.00%	100.00%
NPV		100.00%	100.00%	95.65%	100.00%	100.00%
F_1-Score	81.58%	**100.00%**	**100.00%**	**98.67%**	**100.00%**	**100.00%**
Refusals	N/A	0.00%	N/A	N/A	N/A	N/A
MODEL ROBUSTNESS[4]		#1	#1	#5	#1	#1

1 Given the small number of reference cases of this dataset, the XTRACTIS Reverse-Engineering (CVE→IVE) is necessary to get a robust AND intelligible model.

2 All CVE and IVE models are optimized according to their validation F_1-Score. The XTRACTIS top-CVE and top-IVE are selected according to their validation F_1-Score while checking that it remains close to their training F_1-Score. The ML/LoR CVE top-models are selected according to their F_1-Score mean value in validation. Each ML/LoR top-IVE is obtained by applying the respective ML/LoR top-CVE strategy on 100% of the Training/Validation data.

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).

4 The perfect results of the XTRACTIS, Logistic Regression, BT and Neural Network Top-IVE on External Test could be explained by a low number of reference points compared to the large number of potential predictors.

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