UNIVERSITÉ DE LORRAINE

AfIA
Association française pour l'Intelligence Artificielle
Ecole Saisonnière en IA

Loria
Laboratoire lorrain de recherche en informatique et ses applications

IA & Energie

Human in the loop : Interactivity & visualization

Lydia Boudjeloud-Assala

Professeure des Universités

Université de Lorraine

LORIA - MosAIk

lydia.boudjeloud@loria.fr

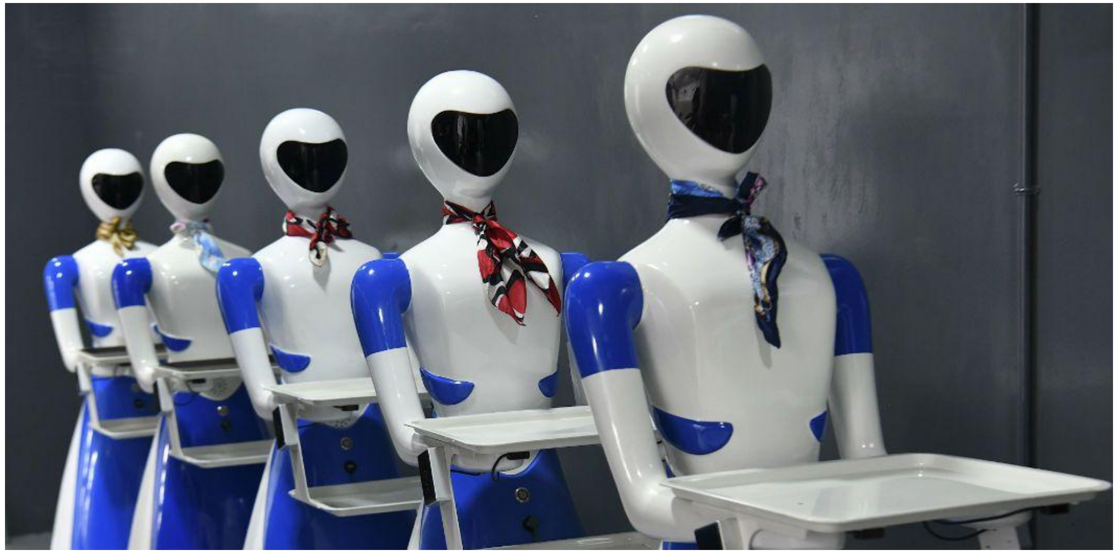
Plan

1. Introduction
2. Problematics
3. Approaches :
    1. Visual analytics
    2. Visual data mining
    3. Interactive machine learning
4. Application to French power grid :
    1. Grammatical evolutionary algorithm
    2. Reinforcement learning
    3. Interactive machine learning algorithms
5. Conclusion and perspectives

1

Introduction

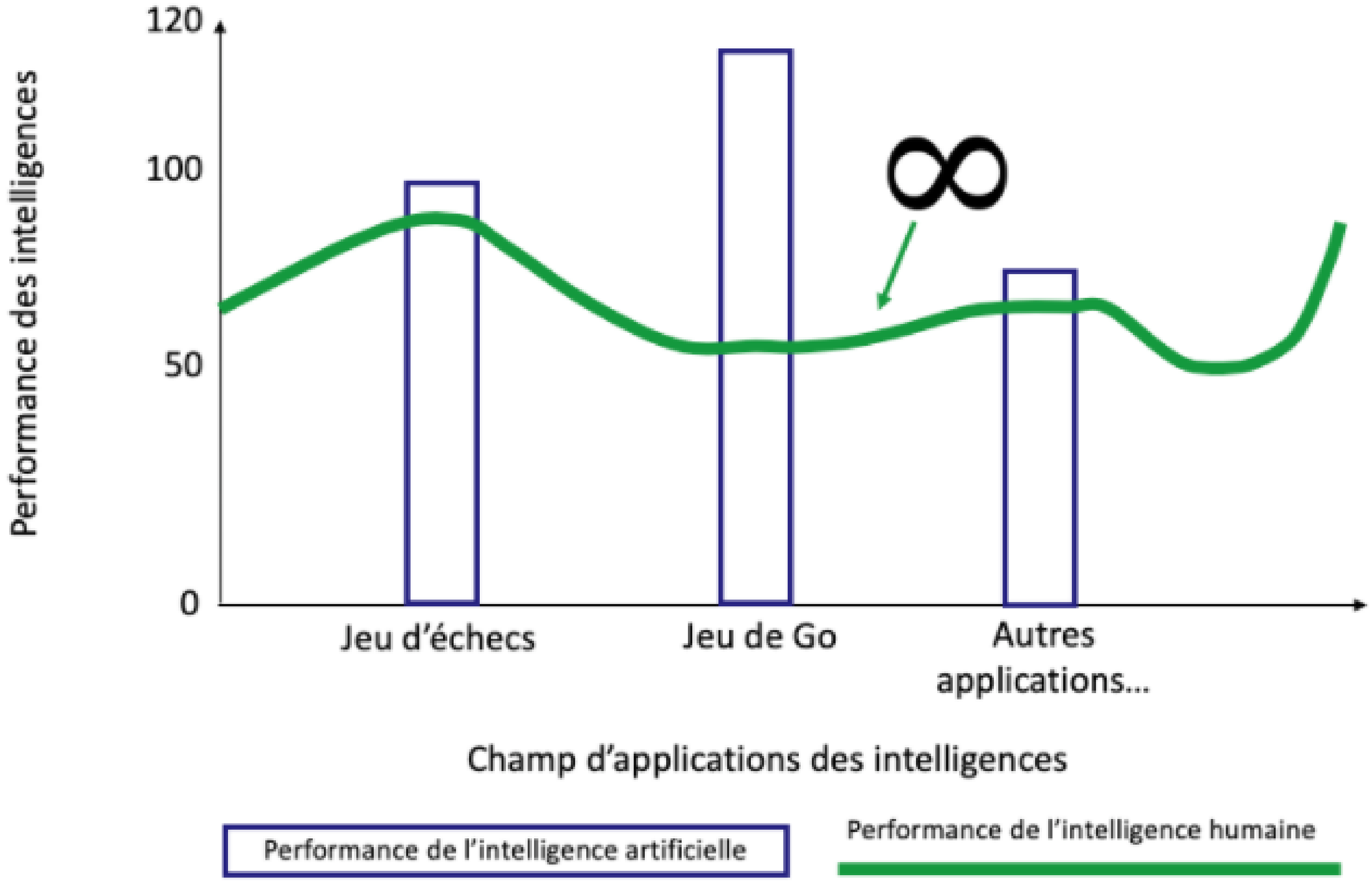
AI in the news


Ouverture du restaurant Robot à Bangalore, le 17 août 2019. Manjunath Kiran / AFP

Artificial Intelligence

L'intelligence artificielle et l'intelligence humaine sont différentes

L'intelligence humaine est capable d'appréhender un nombre infini d'enjeux tandis que l'intelligence artificielle n'est capable que de traiter que des sujets pour lesquels elle fut programmée



Source : Luc Julia

Artificial Intelligence

Artificial Intelligence (AI) :

- thanks to the cross-referencing of an increasing amount of data (big data)
- a rapidly expanding field of research
- applications that concern all human activities,

→ **Medical:** assisted interventions, remote monitoring of patients, personalized treatment,…

→ **Environnement:** prevent the risk of forest fires, analyze the migratory flow of storks, geese,…

→ **Energy:** consumption prediction, predicting or explaining network failures

Multiple approaches & techniques :

- ontology construction,
- automatic language processing - NLP,
- data mining,
- machine learning.

Artificial Intelligence

Definitions

What is Data Mining?

Data mining is considered the process of extracting useful information from large amounts of data.
It is used to discover new, accurate and relevant patterns in the data.

What is machine learning?

It is the design, study, and development of algorithms that allow machines to learn from data.
It is a tool that makes machines a little smarter.
It Automates a large part of the manual data mining tasks, mainly featuring, at the expense of model interpretability.

2

Why human-in-the-loop approaches ?

Data Mining

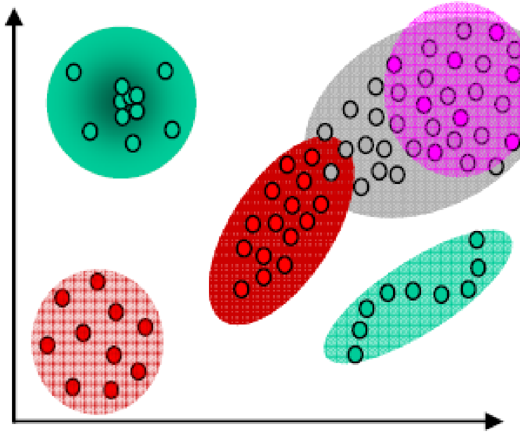
Attributs

Observations

	V_1	V_2
I_1	2	3.4
I_2	1	-0.7
I_3	0.33	4
...	...	...

Data

Algorithm

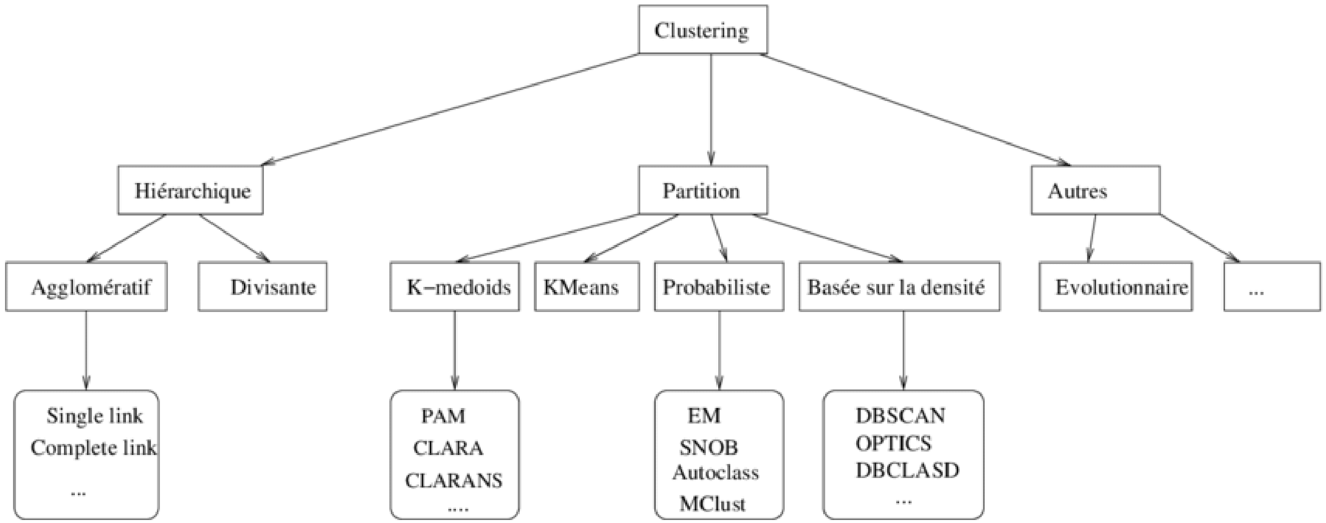


Results

Clustering

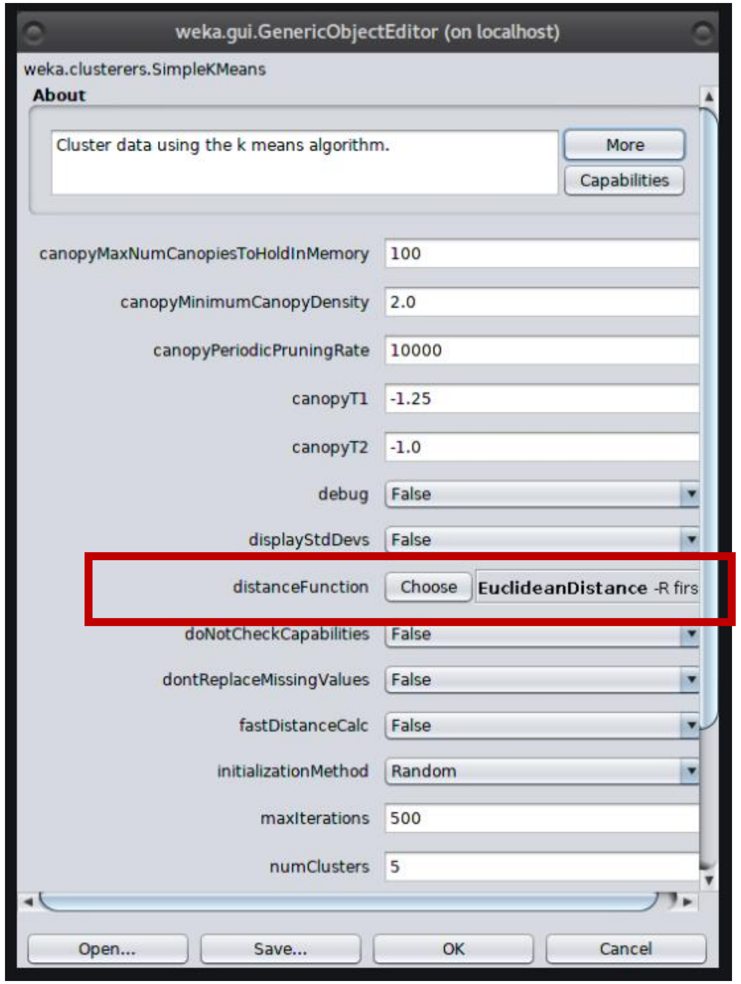
Goal:
Homogeneous sets of individuals sharing same characteristics

Classical process:
1- the user sets the number of clusters
2- a partitioning is generated
3- evaluation:
by the user himself,
by homogeneity criteria,
4- validation:
results can be challenged



Clustering

Cluster algorithm	Complexity	Capability of tackling high dimensional data
K-means	$O(NKd)$ (time) $O(N+K)$ (space)	No
Fuzzy c-means	Near $O(N)$	No
Hierarchical clustering*	$O(N^2)$ (time) $O(N^2)$ (space)	No
CLARA	$O(K(40+K)^2+K(N-K))^+$ (time)	No
CLARANS	Quadratic in total performance	No
BIRCH	$O(N)$ (time)	No
DBSCAN	$O(N \log N)$ (time)	No
CURE	$O(N_{sample}^2 \log N_{sample})$ (time) $O(N_{sample})$ (space)	Yes
WaveCluster	$O(N)$ (time)	No
DENCLUE	$O(N \log N)$ (time)	Yes
FC	$O(N)$ (time)	Yes
CLIQUE	Linear with the number of objects, Quadratic with the number of dimensions	Yes
OptiGrid	Between $O(Nd)$ and $O(Nd \log N)$	Yes
ORCLUS	$O(K_0^3+K_0Nd+K_0^2d^3)$ (time) $O(K_0d^2)$ (space)	Yes



Measures	Forms
Minkowski distance	$D_{ij}=\left(\sum_{l=1}^{d}\lvert x_{il}-x_{jl}\rvert^{1/n}\right)^{n}$
Euclidean distance	$D_{ij}=\left(\sum_{l=1}^{d}\lvert x_{il}-x_{jl}\rvert^{1/2}\right)^{2}$
City-block distance	$D_{ij}=\sum_{l=1}^{d}\lvert x_{il}-x_{jl}\rvert$
Sup distance	$D_{ij}=\max_{1\le l\le d}\lvert x_{il}-x_{jl}\rvert$
Mahalanobis distance	$D_{ij}=(\mathbf{x}_i-\mathbf{x}_j)^T\mathbf{S}^{-1}(\mathbf{x}_i-\mathbf{x}_j)$, where $\mathbf{S}$ is the within-group covariance matrix.
Pearson correlation	$D_{ij}=(1-r_{ij})/2$, where $r_{ij}=\frac{\sum_{l=1}^{d}(x_{il}-\overline{x_i})(x_{jl}-\overline{x_j})}{\sqrt{\sum_{l=1}^{d}(x_{il}-\overline{x_i})^2\sum_{l=1}^{d}(x_{jl}-\overline{x_j})^2}}$
Point symmetry distance	$D_{ir}=\min_{\substack{j=1,\dots,N\\ \text{and } j\neq i}}\frac{\lVert(\mathbf{x}_i-\mathbf{x}_r)+(\mathbf{x}_j-\mathbf{x}_r)\rVert}{\lVert(\mathbf{x}_i-\mathbf{x}_r)\rVert+\lVert(\mathbf{x}_j-\mathbf{x}_r)\rVert}$
Cosine similarity	$S_{ij}=\cos\alpha=\frac{\mathbf{x}_i^T\mathbf{x}_j}{\lVert\mathbf{x}_i\rVert\lVert\mathbf{x}_j\rVert}$

R. Xu and D.Wunsch. Survey of clustering algorithms. IEEE Transactions on Neural Networks, 16(3):645–678, May 2005.

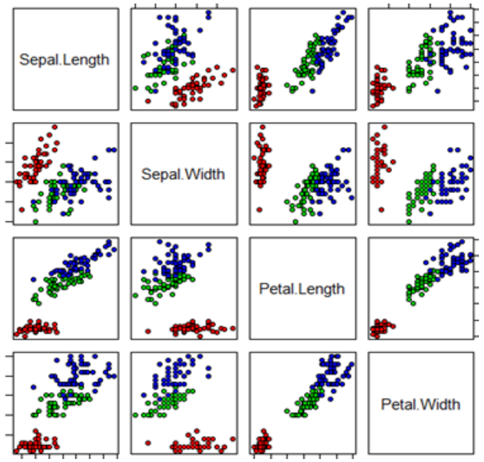
Clustering

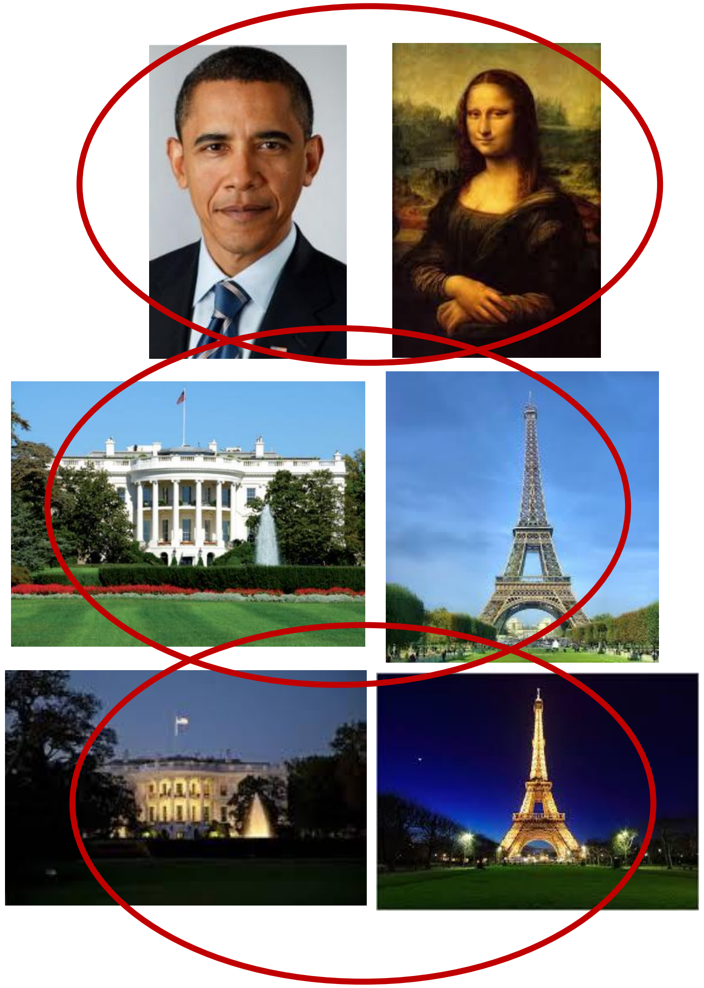
Evaluation: different point of views

Index	*Rule*
Ball_Hall	*max diff*
Banfeld_Raftery	*min*
C_index	*min*
Calinski_Harabasz	*max*
Davies_Bouldin	*min*
Det_Ratio	*min diff*
Dunn	*max*
GDI	*max*
Gamma	*max*
G_plus	*min*
Ksq_DetW	*max diff*
Log_Det_Ratio	*min diff*
Log_SS_Ratio	*min diff*
McClain_Rao	*min*
PBM	*max*
Point_biserial	*max*
Ratkowsky_Lance	*max*
Ray_Turi	*min*
Scott_Symons	*min*
SD	*min*
S_Dbw	*min*
Silhouette	*max*
Tau	*max*
Trace_W	*max diff*
Trace_WiB	*max diff*
Wemmert_Gancarski	*max*
Xie_Beni	*min*

Iris Fischer (150 obs., 4 var., 3 classes)

The Czekanowski-Dice index
The Folkes-Mallows index .
The Hubert $\hat{\Gamma}$ index . . . .
The Jaccard index . . . . .
The Kulczynski index . . .
The McNemar index . . . .
The Phi index . . . .
The Rand index . . . . . .
The Rogers-Tanimoto index
The Russel-Rao index . . .
The Sokal-Sneath indices .





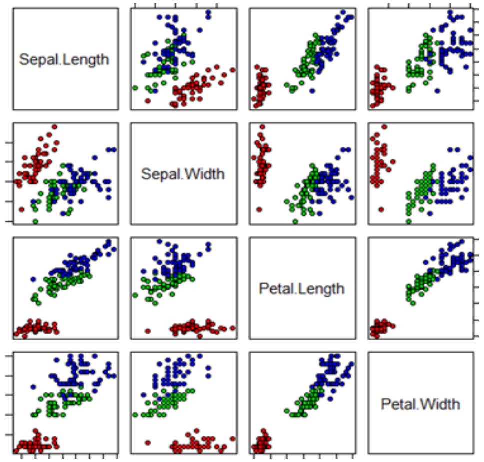
Clustering

Evaluation: different point of views

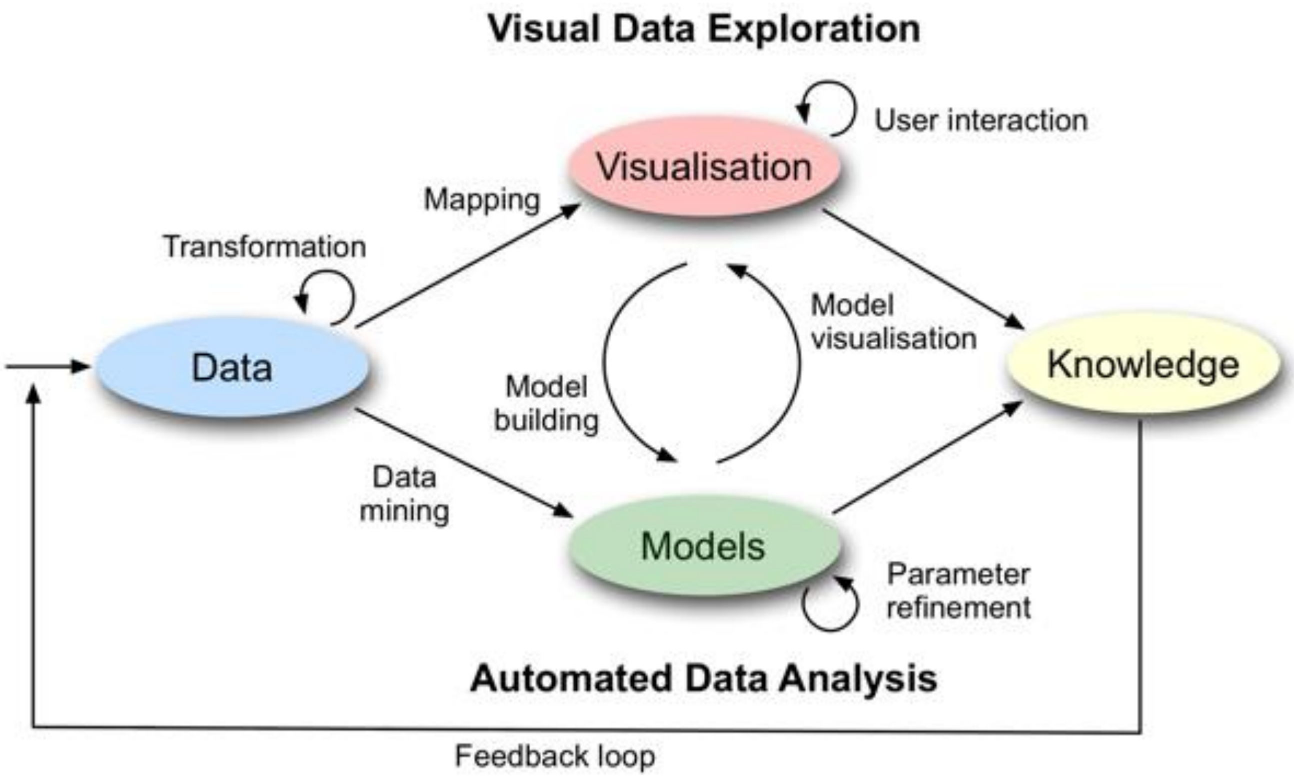
Index	*Rule*
Ball_Hall	*max diff*
Banfeld_Raftery	*min*
C_index	*min*
Calinski_Harabasz	*max*
Davies_Bouldin	*min*
Det_Ratio	*min diff*
Dunn	*max*
GDI	*max*
Gamma	*max*
G_plus	*min*
Ksq_DetW	*max diff*
Log_Det_Ratio	*min diff*
Log_SS_Ratio	*min diff*
McClain_Rao	*min*
PBM	*max*
Point_biserial	*max*
Ratkowsky_Lance	*max*
Ray_Turi	*min*
Scott_Symons	*min*
SD	*min*
S_Dbw	*min*
Silhouette	*max*
Tau	*max*
Trace_W	*max diff*
Trace_WiB	*max diff*
Wemmert_Gancarski	*max*
Xie_Beni	*min*

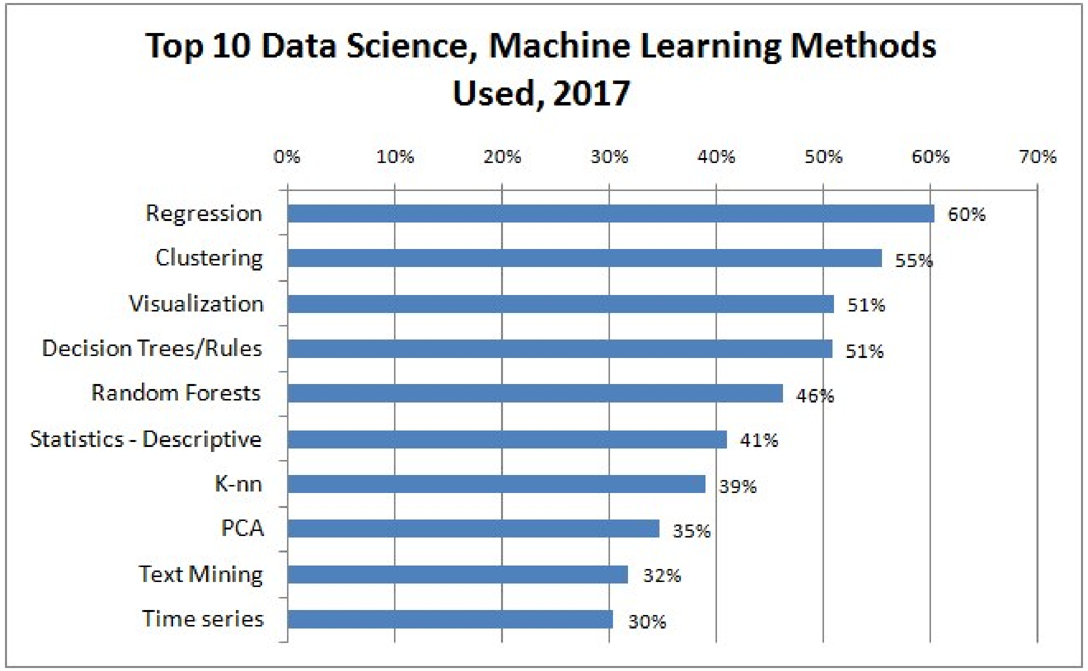
Iris Fischer (150 obs., 4 var., 3 classes)

The Czekanowski-Dice index
The Folkes-Mallows index .
The Hubert $\hat{\Gamma}$ index . . . .
The Jaccard index . . . . .
The Kulczynski index . . .
The McNemar index . . . .
The Phi index . . . .
The Rand index . . . . . .
The Rogers-Tanimoto index
The Russel-Rao index . . .
The Sokal-Sneath indices .



Visual Data Mining





http://www.kdnuggets.com/ (Décembre 2017)
https://www.kdnuggets.com/2017/12/top-data-science-machine-learning-methods.html

Visual Data methods : **38.3%** → 5th position in 2011, **48.7%** → 4th position in 2016
à **51%** → 3rd position in 2017

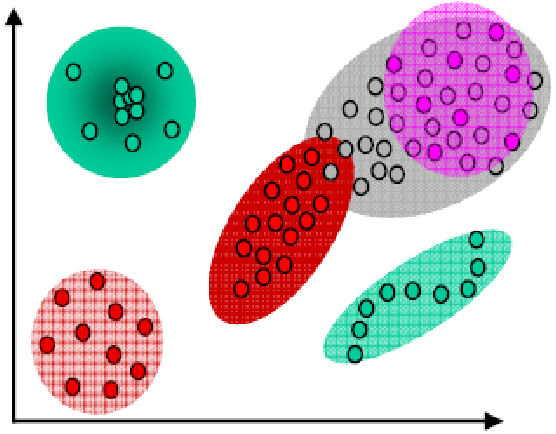
Visual Data Mining

Attributs

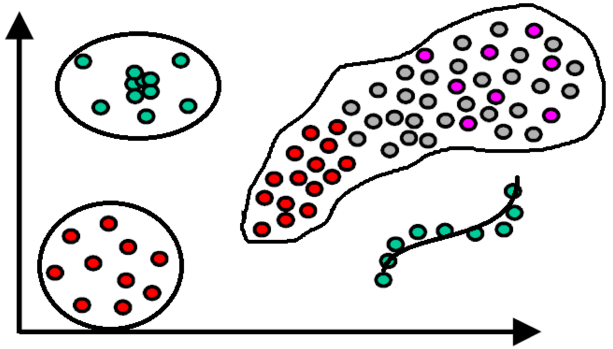
Observations

	V_1	V_2
I_1	2	3.4
I_2	1	-0.7
I_3	0.33	4
...	...	...

Data



Results



Semi-interactive

Semi-visual approach

Goal:
Homogeneous sets of individuals sharing same characteristics

How:
Data representation + risk + evaluation →
automatic methods, visualization, interactivity
The user (the data expert) :

- this suggests that ...
- I improve, I continue ...
- I do not improve, I start again or come back ...
- I valid

Visual Data Mining

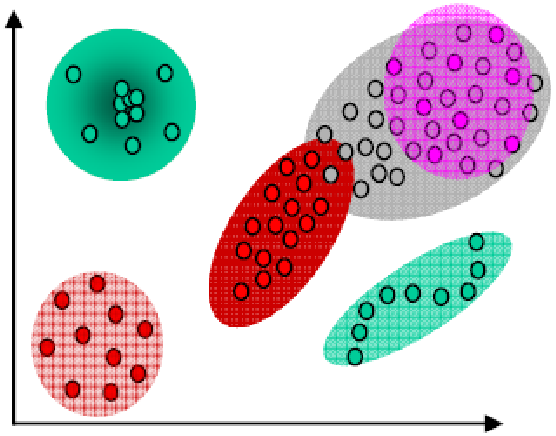
P Variables

Observations,

	V_1	V_2
x_1	2	3.4
x_2	1	-0.7
x_3	0.33	4
...	...	...

Data set E

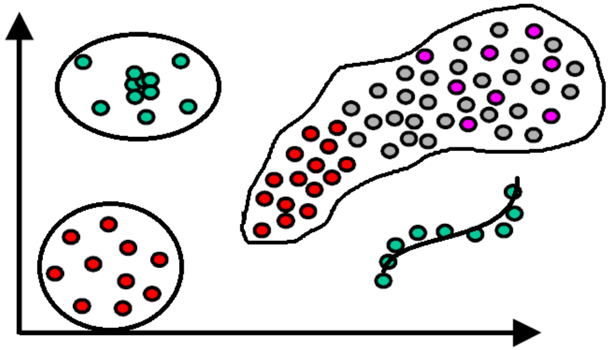
$$f : R^p \to \{1, \dots, k\}$$
$$X \to C$$



Results

$$Op : \{1, \dots, I\} \to \{1, \dots, I\}$$
$$C \to C'$$

$$\boxed{Q(C') > Q(C)}$$



Semi-interactif

Clustering evaluation:

- Minimize intra-classe inertia

$$I_{intra}(C_a) = \sum_{x_i \in C_a} d(x_i, x_a^*)^2$$

- Maximize inter-classe inertia

$$I_{inter}(C) = \sum_{i=2}^{t} \sum_{j<i} d(x_i^*, x_j^*)^2$$

- ....

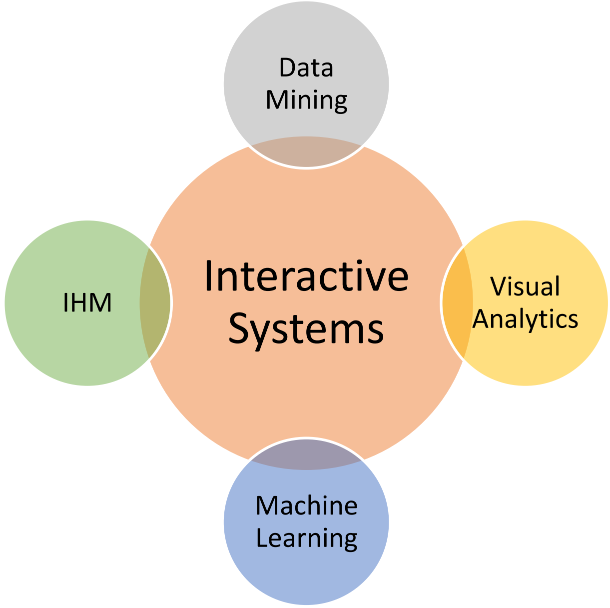
Interactives systems

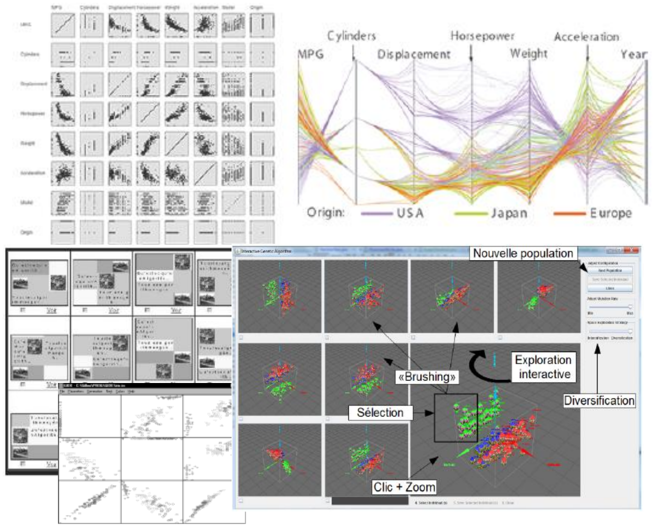
Attributs

Observations

	V_1	V_2
I_1	2	3.4
I_2	1	-0.7
I_3	0.33	4
…	…	…

Data





Differences

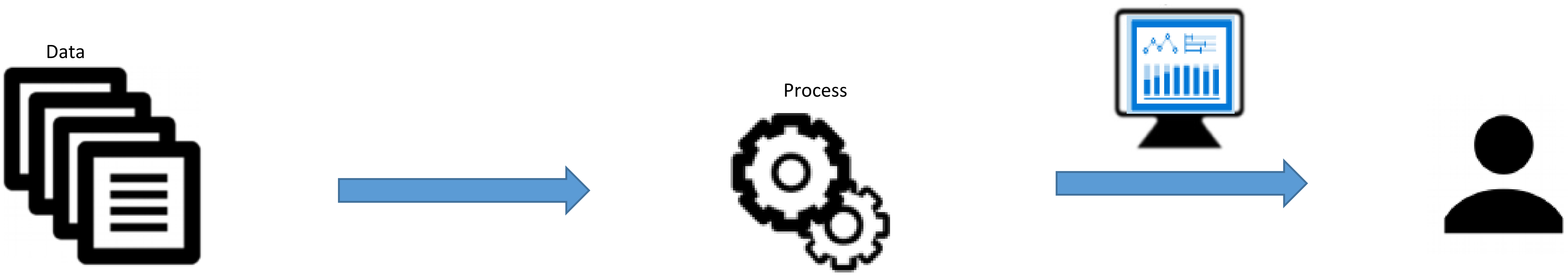
- Interactivity in the processus
- Interactivity steps
- Improve the model

Advantages

- Obtain better results
- Complex tasks to be automated
- Understand data and resluts
- Imply the user in the process

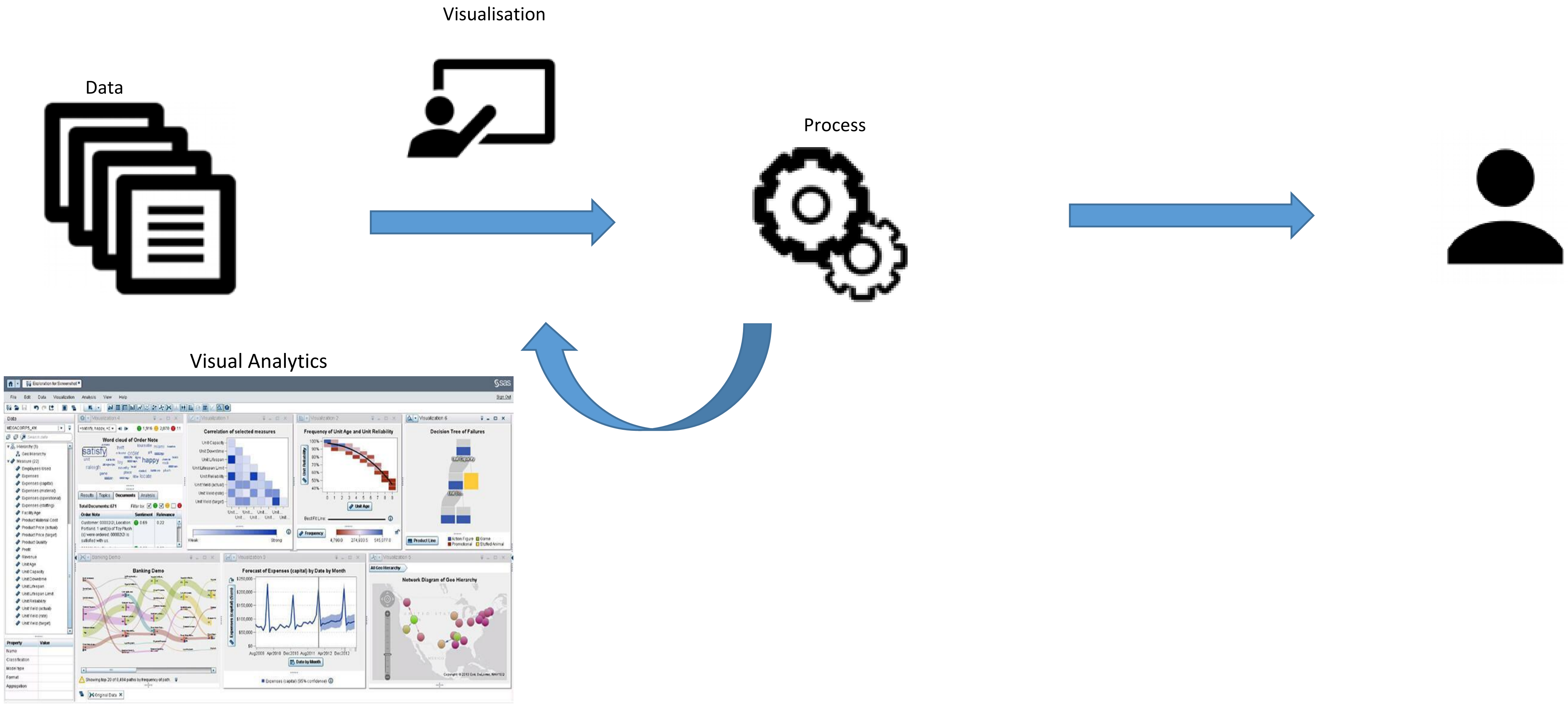
Human-In-The-Loop

Classical approach



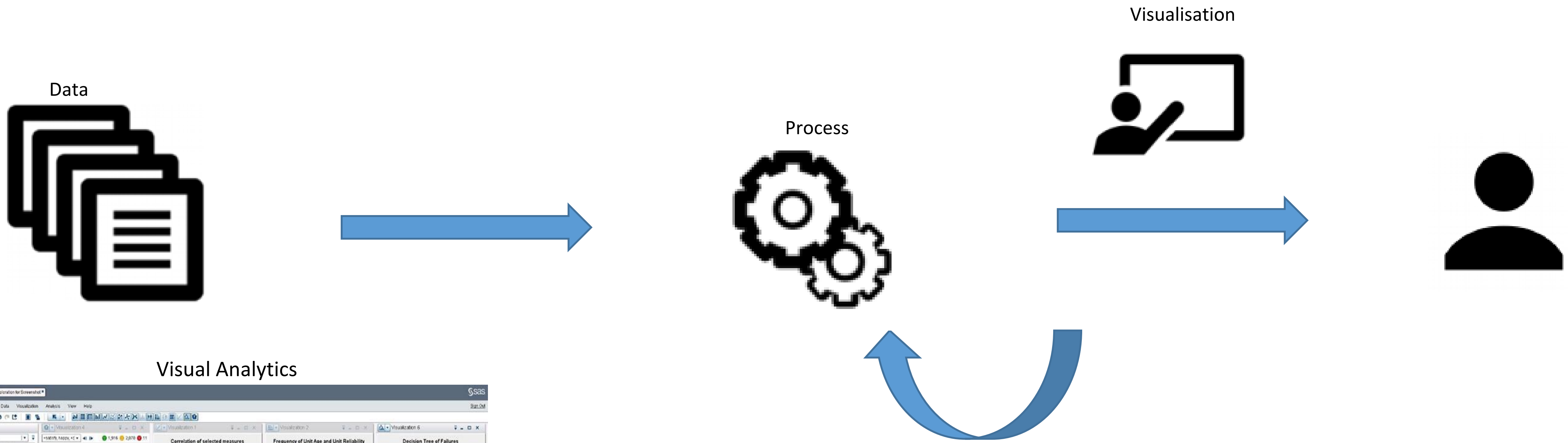
Human-In-The-Loop

Visual Analytics

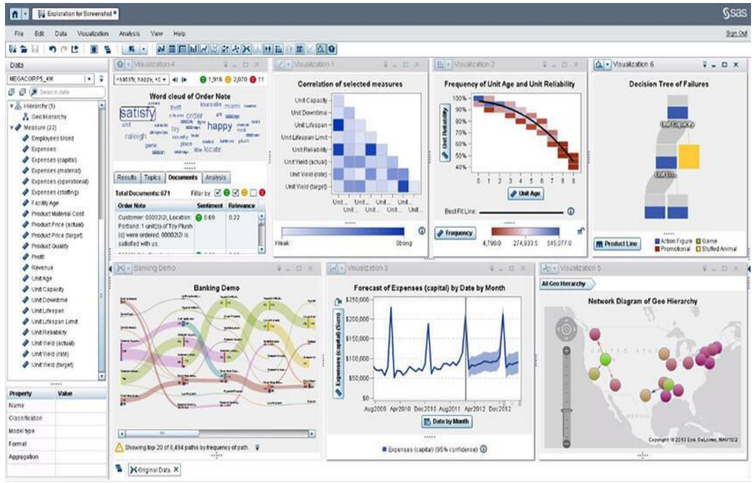


Human-In-The-Loop

Visual Analytics

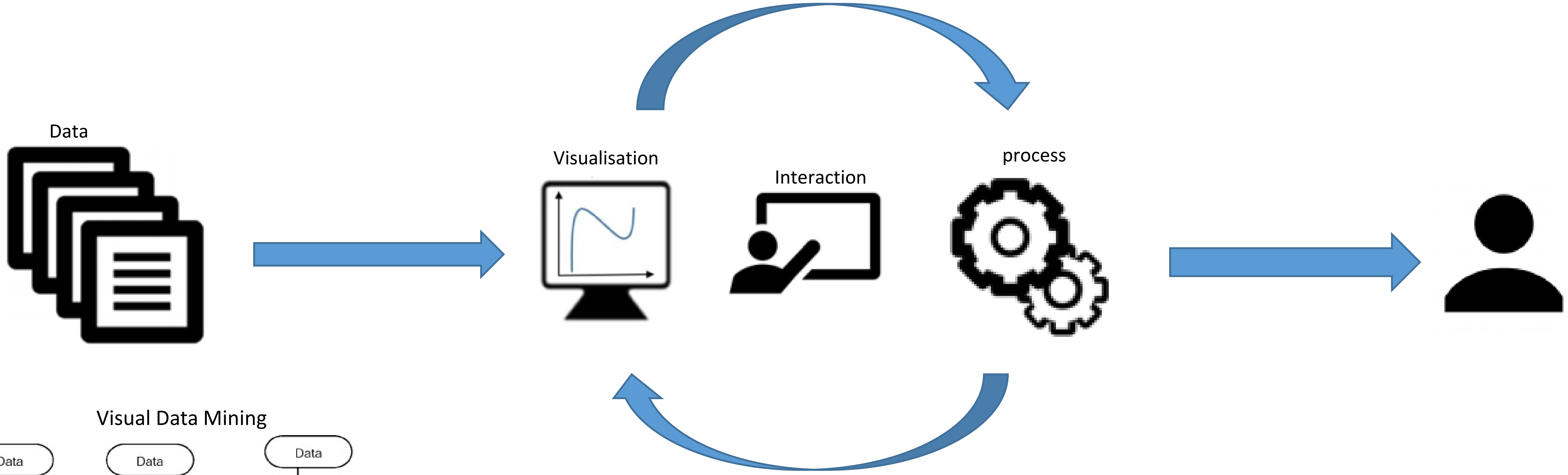


Visual Analytics

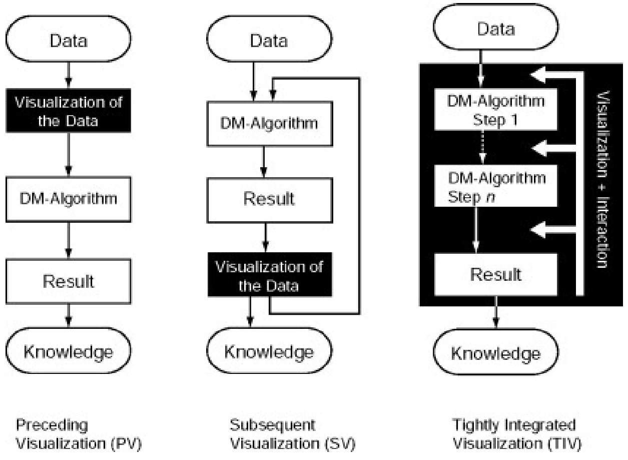


Human-In-The-Loop

Visual Data Mining

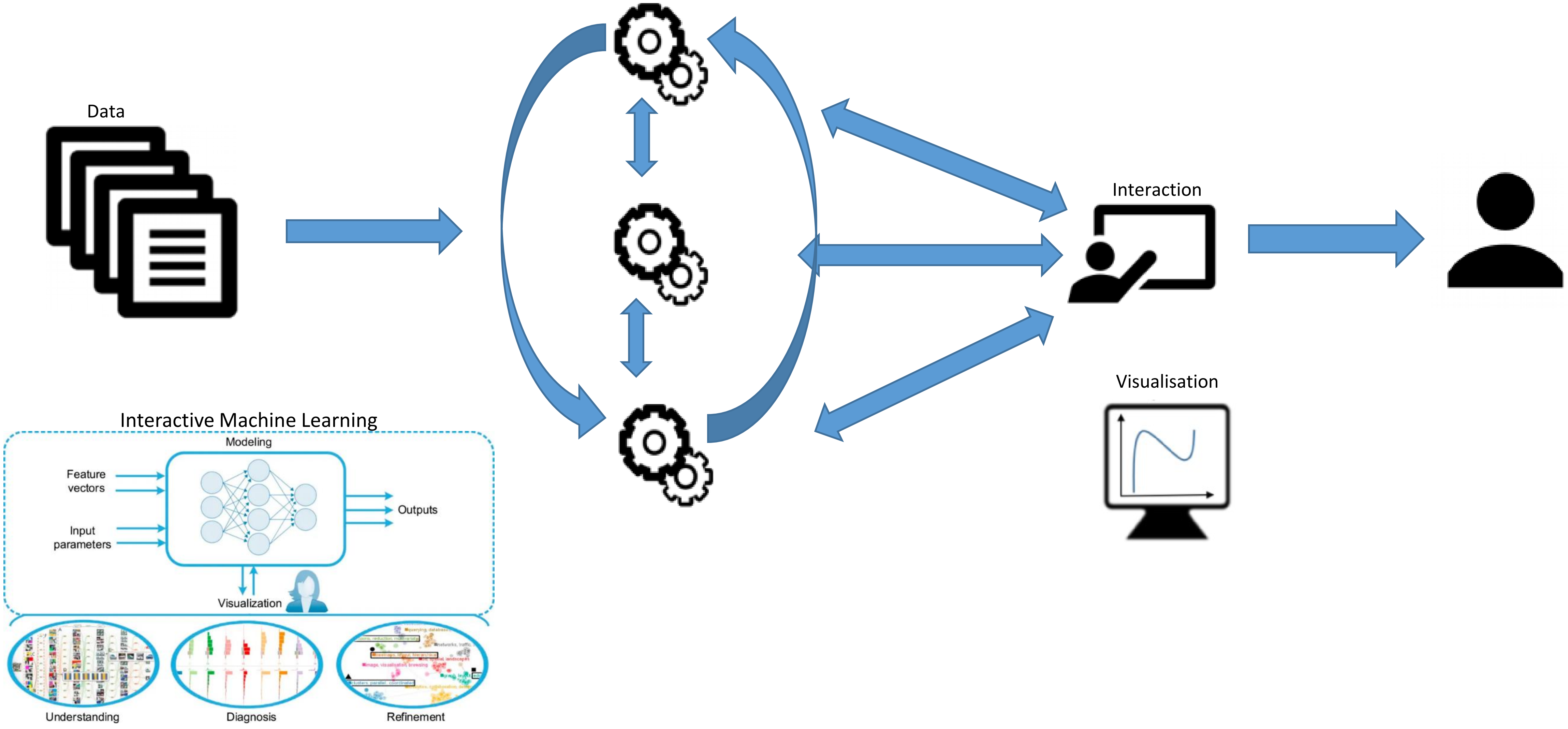


Visual Data Mining



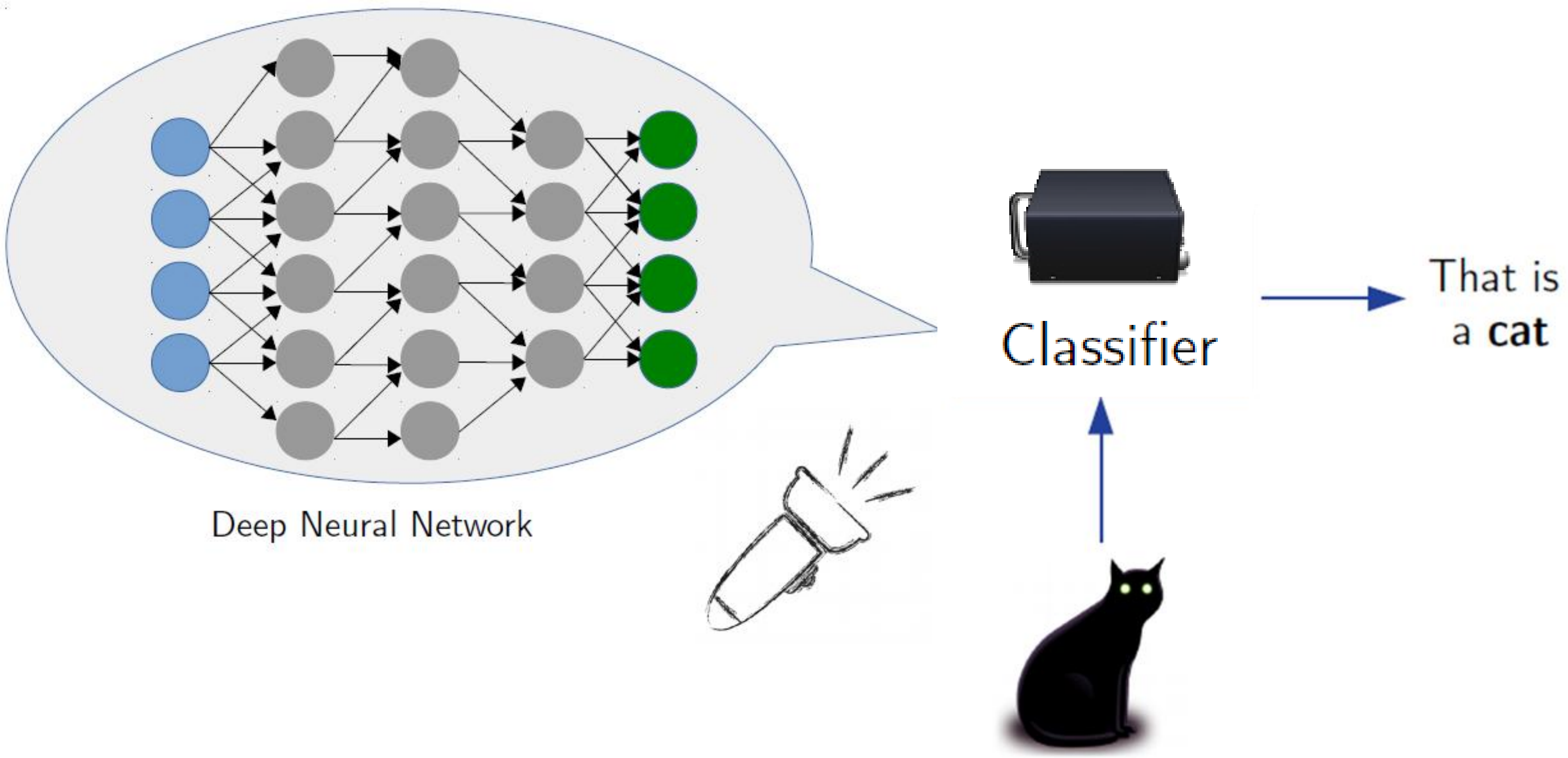
Human-In-The-Loop

Interactive Machine Learning



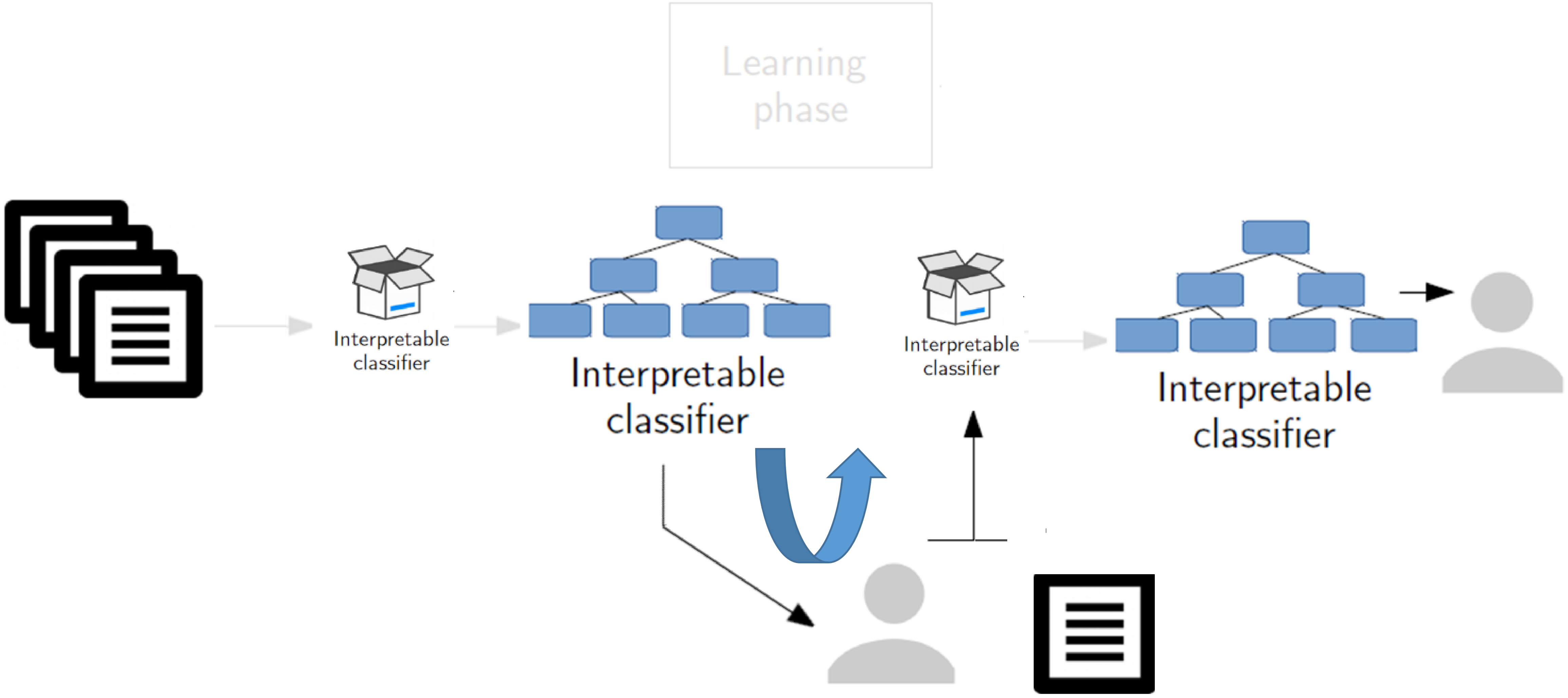
Interactive Machine Learning

Explicability interpretability of decision models and prediction



Interactive Machine Learning

Explicability interpretability of decision models and prediction

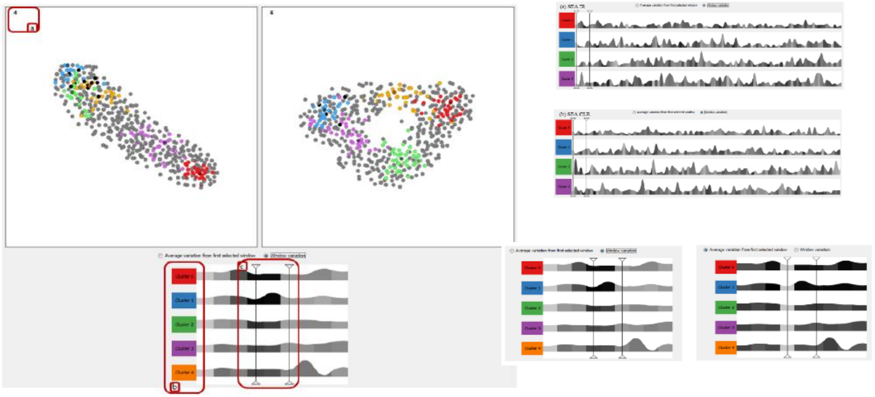


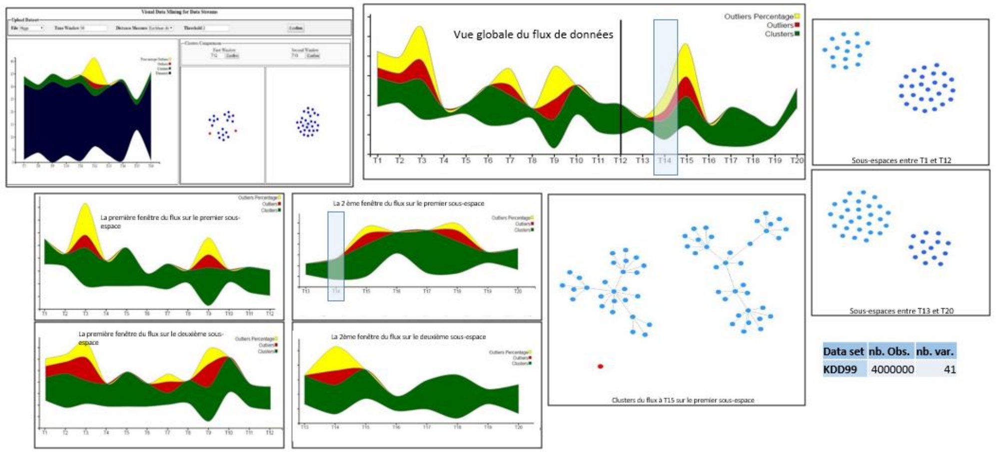
3

Contributions

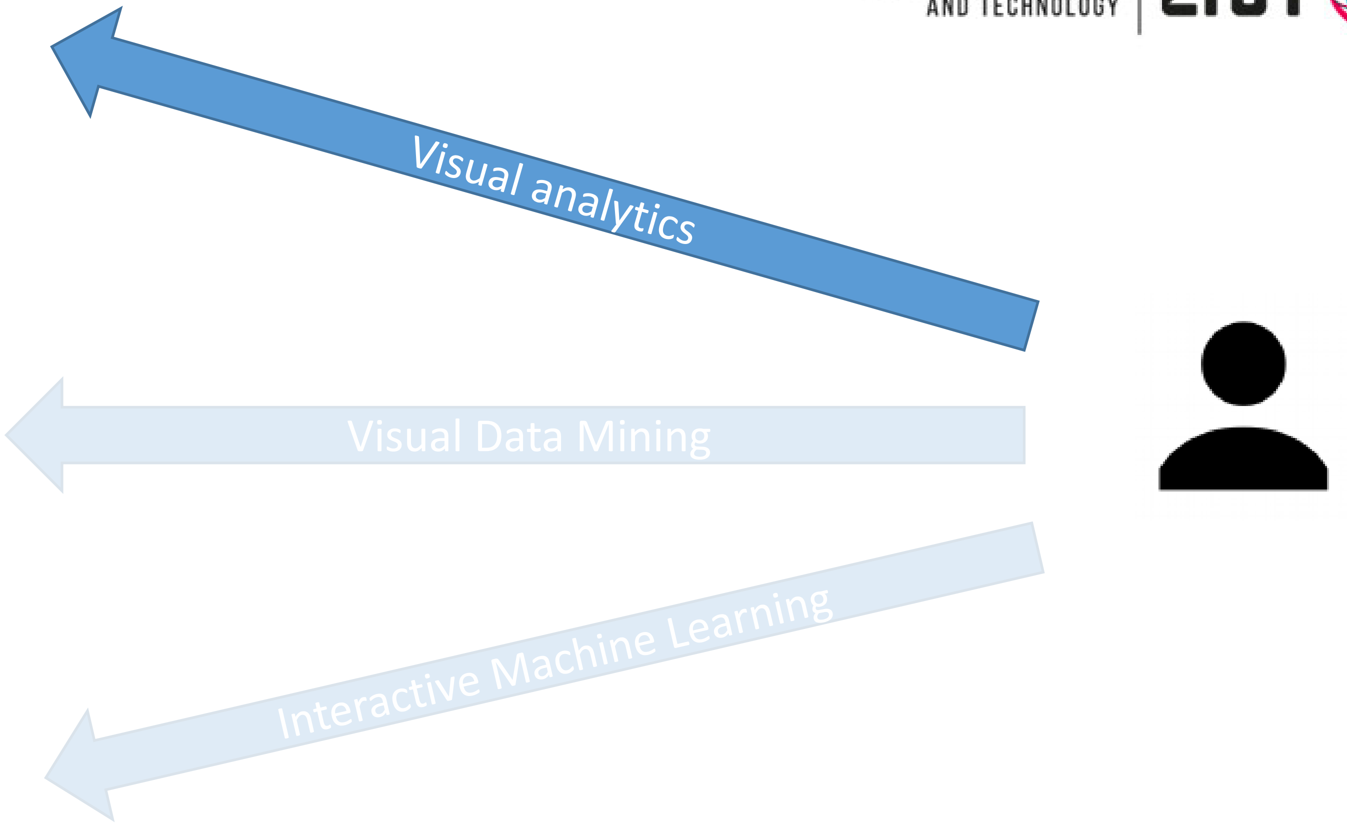
Human-In-The-Loop

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST





Data set	nb. Obs.	nb. var.
KDD99	4000000	41



Data stream clustering

Neighborhood graphs

Temporal windows

Clustering

NNG:
distance < threshold
Connected component=
Clusters

Cluster

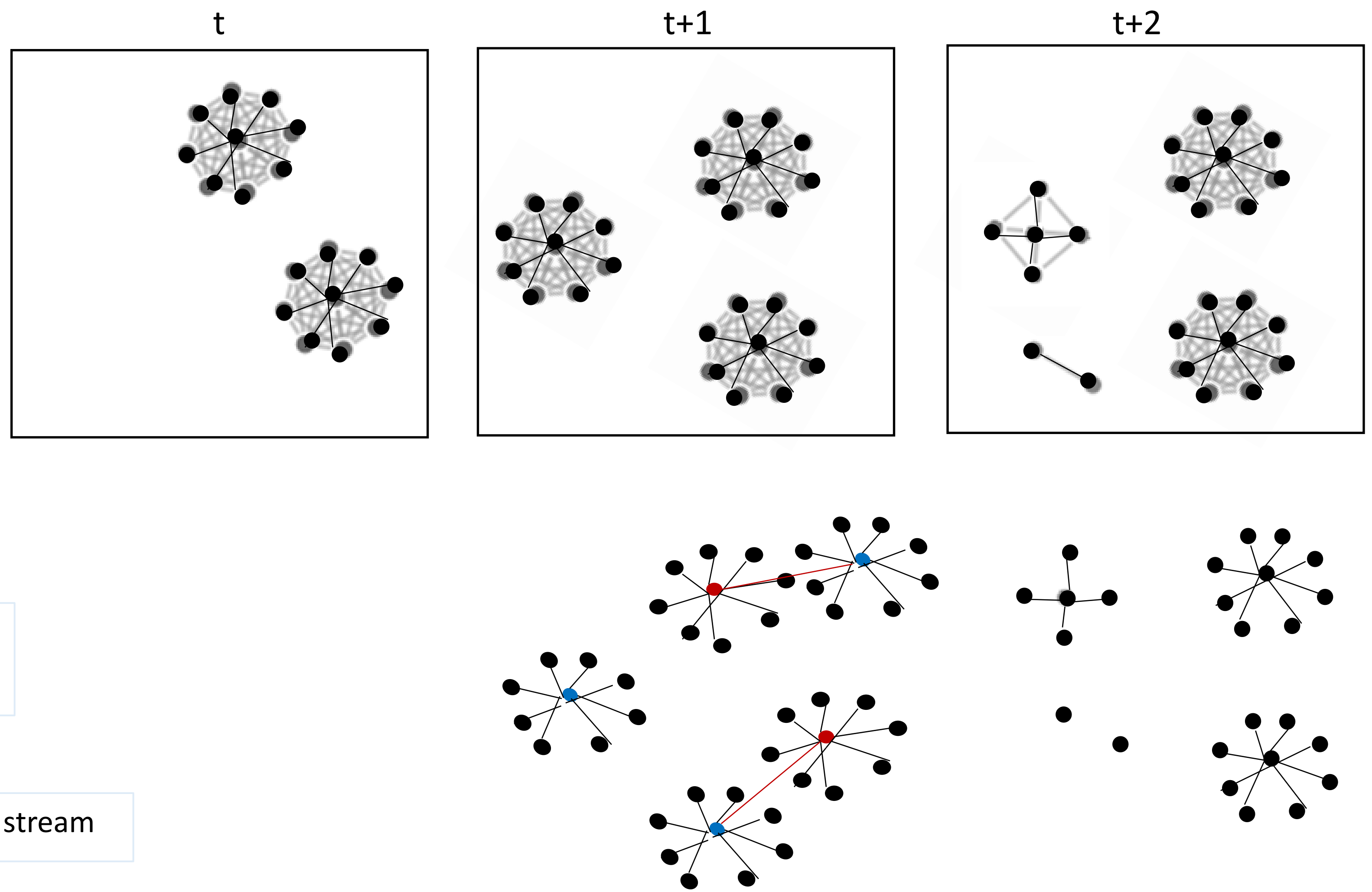
Centroid
neighbors

Stream

Centroids of W(t+1) link with
nearest centroids of W(t)

Visualization

Incremental graph through the stream

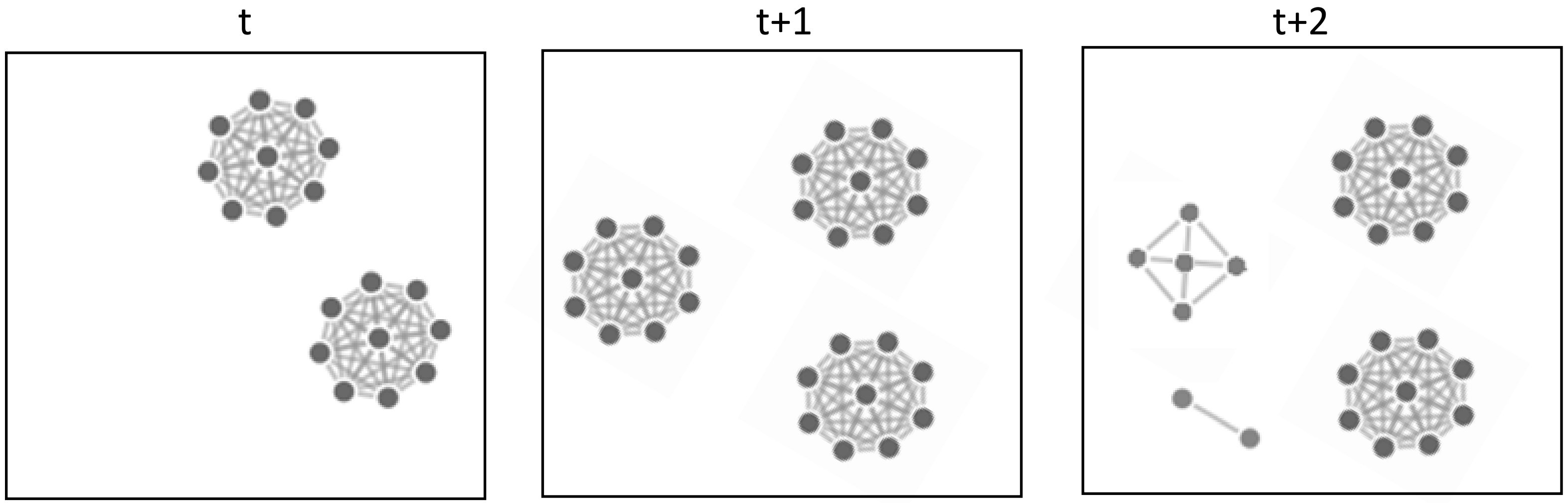


Ibrahim Louhi, Lydia Boudjeloud-Assala, Thomas Tamisier, Incremental nearest neighborhood graph for data stream clustering. IJCNN 2016: 2468-2475

Data stream clustering

Neighborhood graphs

Temporal windows



Clustering

NNG:
distance < threshold
Connected component=
Clusters

Cluster

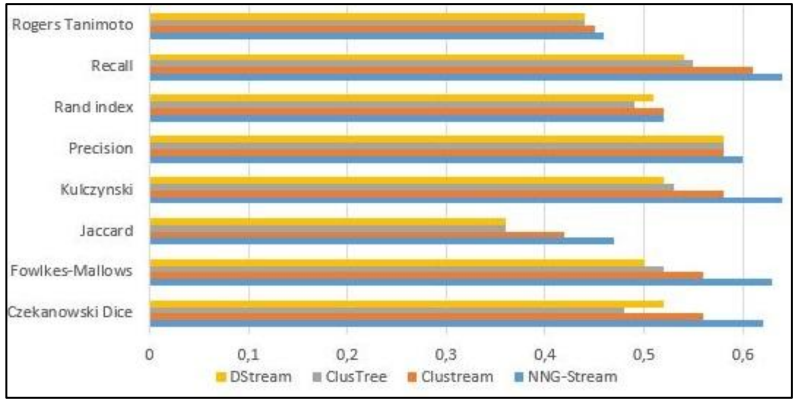
Centroid
neighbors

Stream

Centroids of W(t+1) link with nearest centroids of W(t)

Results:

- Comparison with the state of the art
- External evaluation indices
- Massive data sets



Data set	Type	nb. Obs.	nb. var.
Higgs	Numeric	11000000	28
Hepmass	Numeric	10500000	28
Susy	Numeric	5000000	0
KDD99	Heterogenous	4000000	41

Application:
Outliers detection on log files
Comparison: MCOD algorithm (Kontaki & al., 2011)

Ibrahim Louhi, Lydia Boudjeloud-Assala, Thomas Tamisier, Exploration and Visualization Approach for Outlier Detection on Log Files. New Trends in Intelligent Information and Database Systems 2015: 3-12

Data stream biclustering

Neighborhood graphs

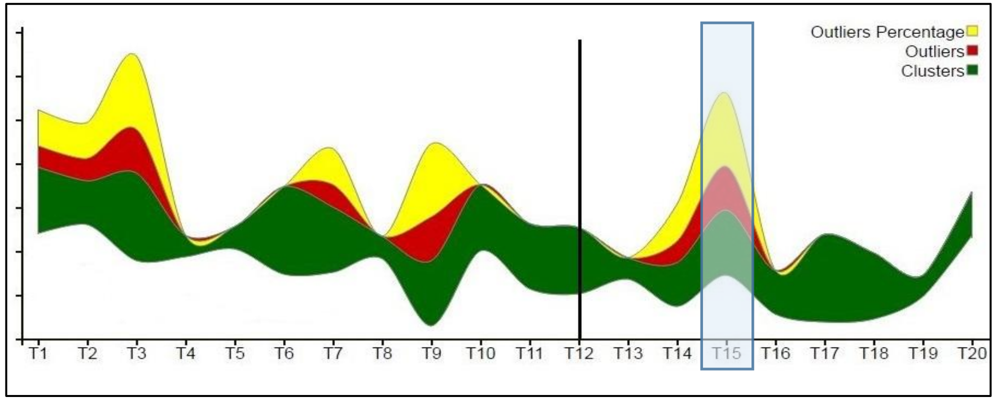
Contr 2

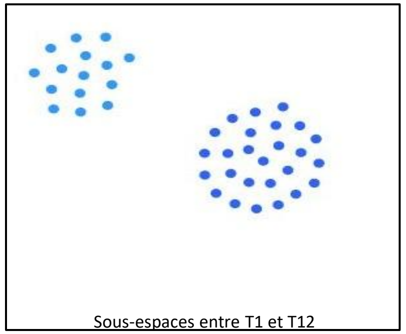
Data set	nb. Obs.	nb. var.
KDD99	4000000	41

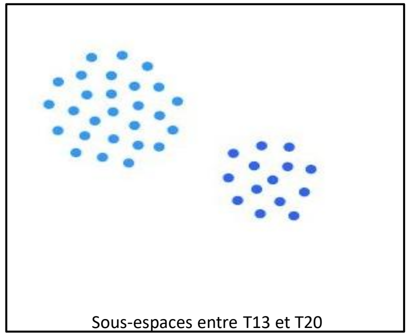
Proposition:
Neighborhood graphs clustering
On attributes then the observations

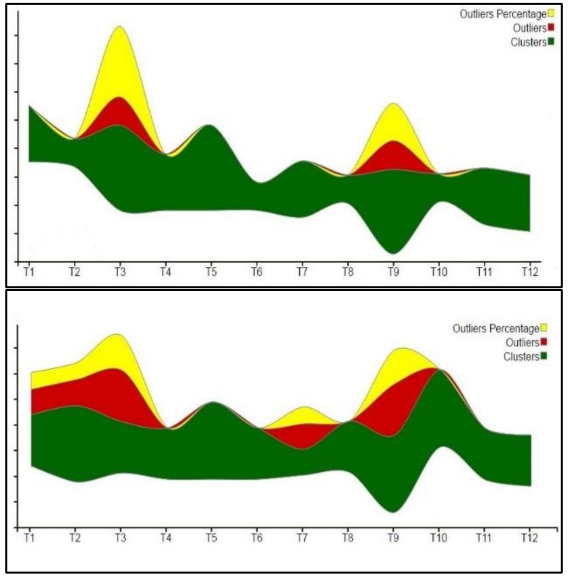
Contr 3

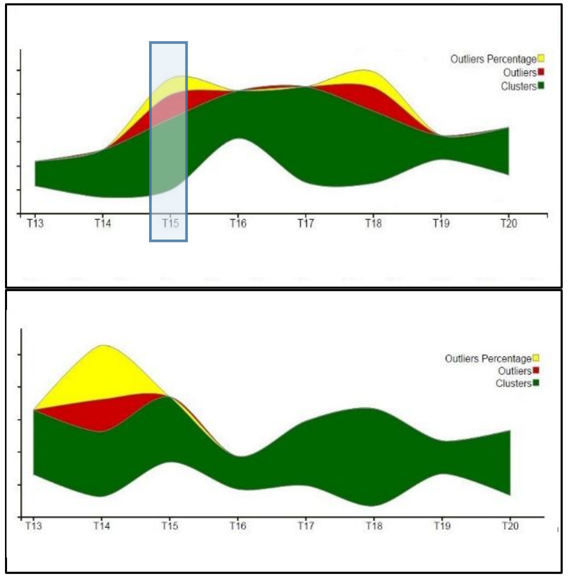
Visual exploration

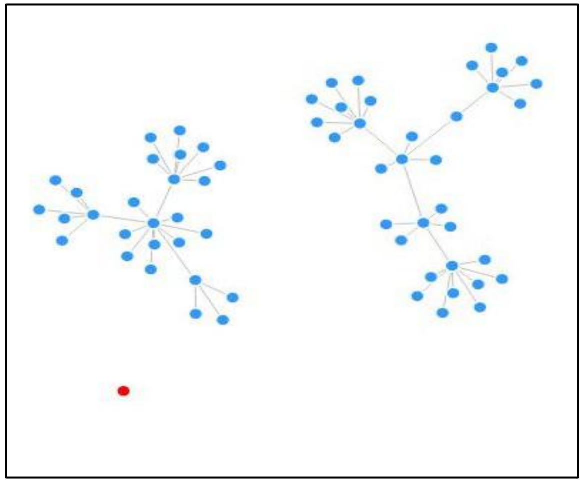










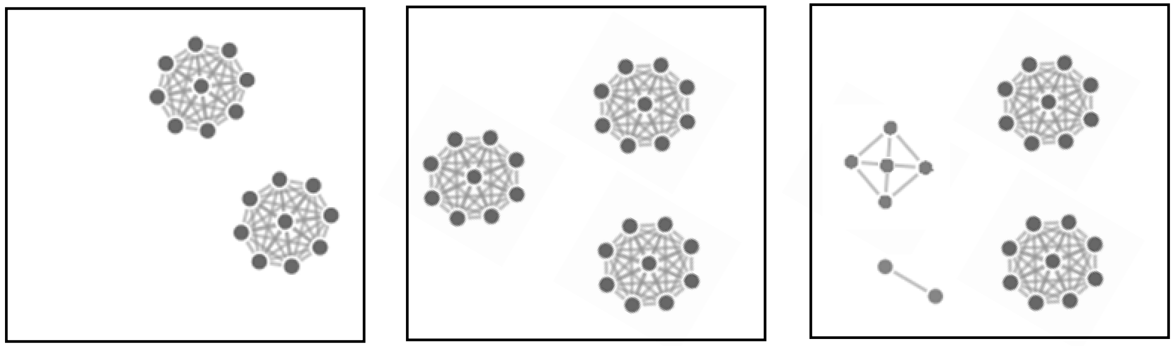


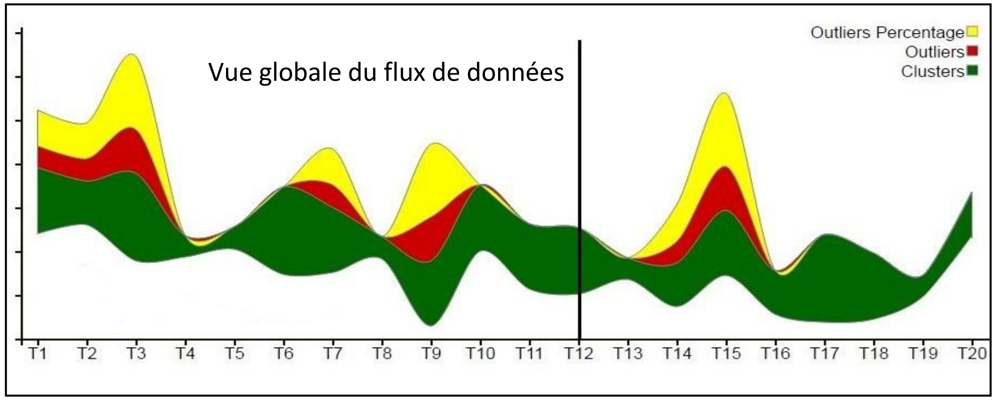
Ibrahim Louhi, Lydia Boudjeloud-Assala, Thomas Tamisier, Subspace Clustering and Visualization of Data Streams. VISIGRAPP (3: IVAPP) 2017: 259-265, 2016

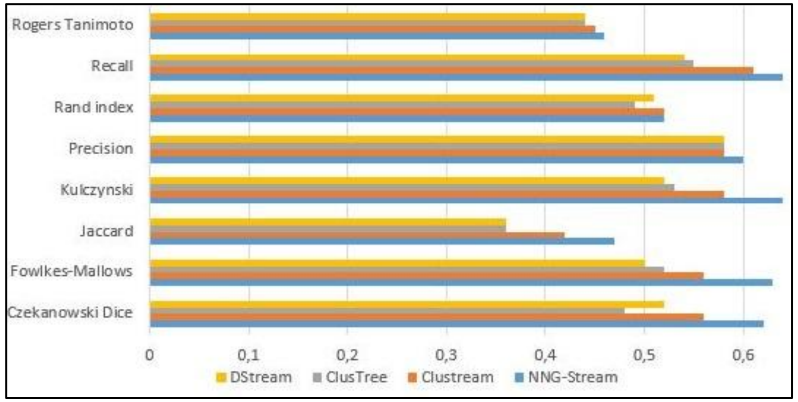
Data stream biclustering

Results and applications

T t+1 t+2





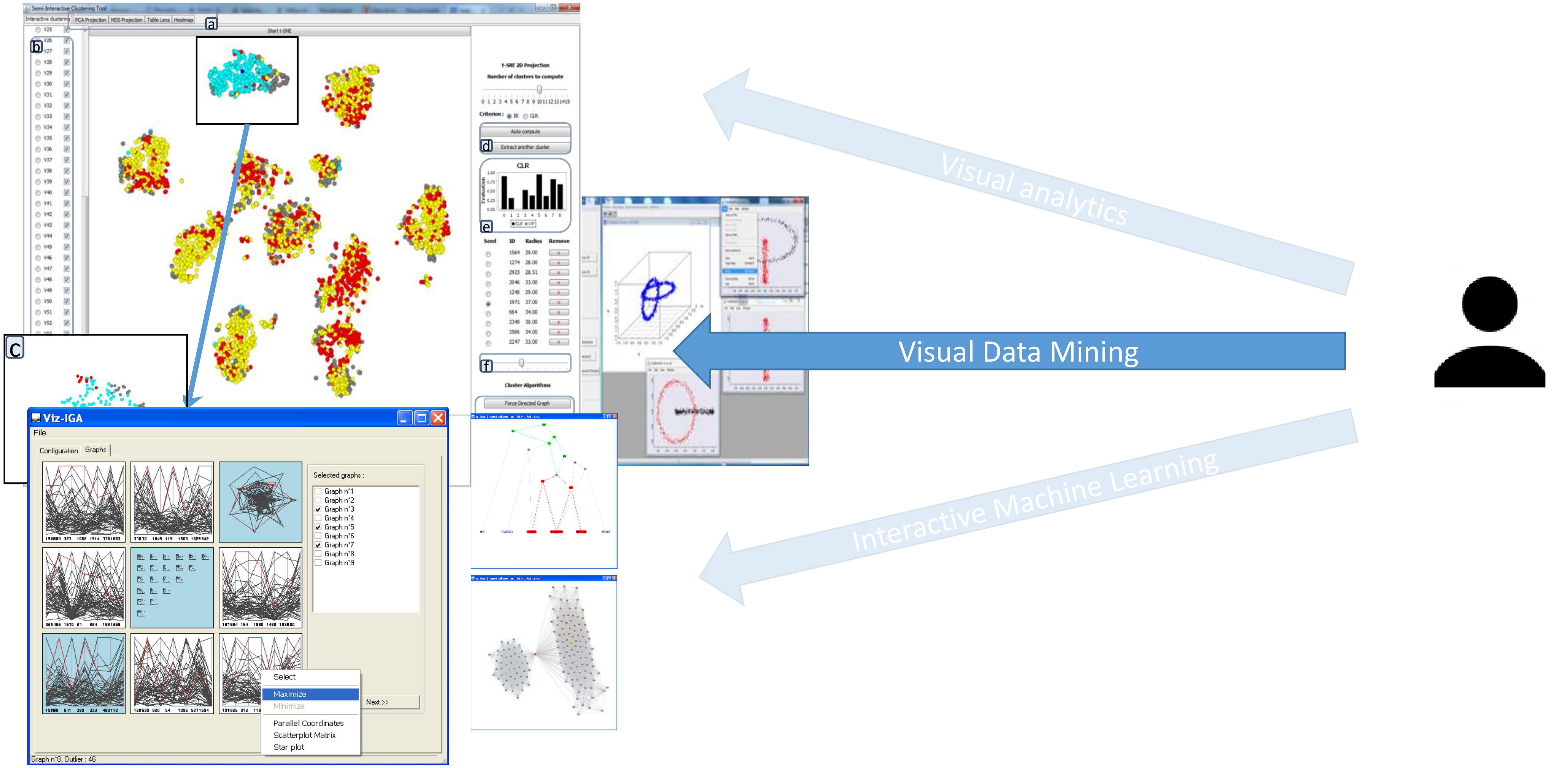


Hepmass data set

Data set	Type	nb. Obs.	nb. var.
Higgs	Numeric	11000000	28
Hepmass	Numeric	10500000	28
Susy	Numeric	5000000	0
KDD99	Heterogenous	4000000	41

Application :
Detection of atypical behavior
Recording of network traffic events
A comparison with MCOD (Kontaki & al., 2011)

Human-In-The-Loop



Clustering

Deterministic approach

Contr 1: cluster extraction

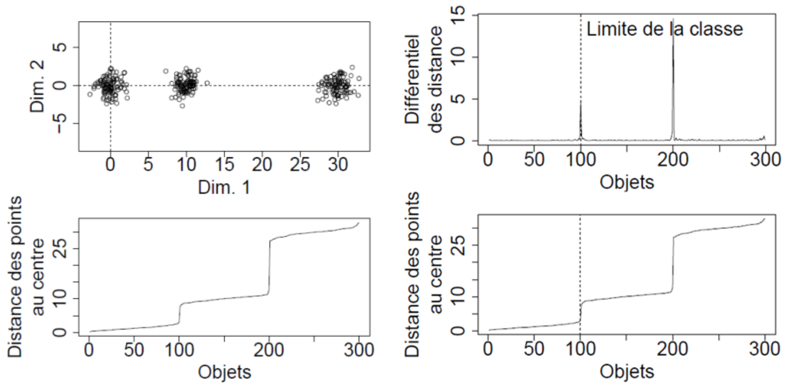
Extraction of a spherical cluster

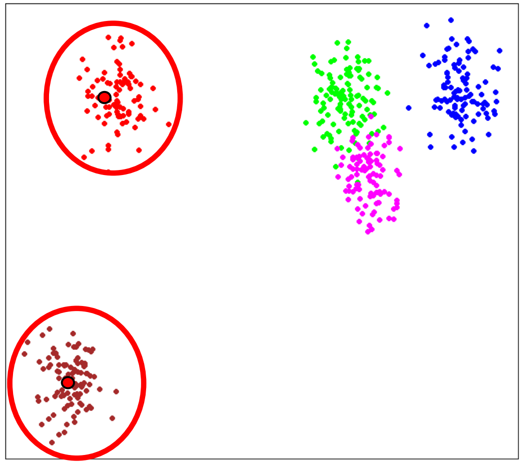
- Cluster defined by a center
- Data distributed around the center
- Cluster radius (distance limit)

Automatic search of the cluster limit

- Calculating the distance of all objects at the given center
- sort the distances
- Finding the first steep increase in distance

→ Peak detection method (Palshikar, 2009): evaluation measure applied to the distance differential





… without knowing k

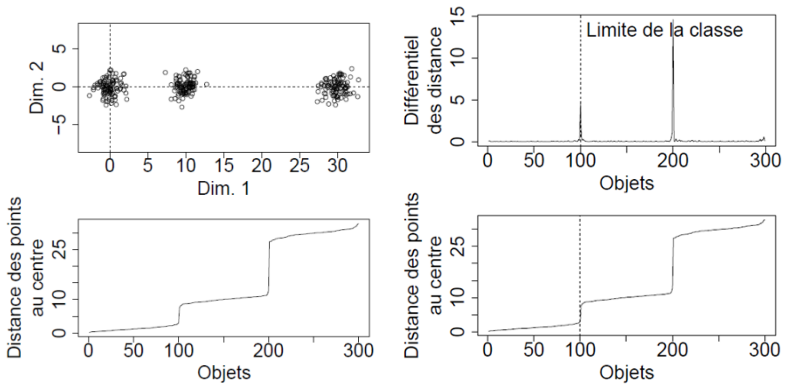
Deterministic approach

Contr 2: cluster evaluation

Compactness criterion of one cluster

- entire partition
- cluster evaluation based on the other clusters (Wemmert-Gançarski, 2000)

Inertia ratio IR: the ratio of the intra-class inertia by the total inertia of the data, normalized by the number of objects in the extracted cluster and the whole data set
Cluster limit ratio CLR: the ratio between the distance of the last object of the extracted cluster and the first object outside the cluster.
OP: penalty for overlapping cluster to make sure that the extracted clusters are different



$$IR(C_k) = \frac{Card(D) \sum_{o \in C_k} d(o, c_k)^2}{Card(C_k) \sum_{o \in D} d(o, g)^2}$$

$$CLR(C_k) = \frac{\max_{o \in C_k} (d(o, c_k))}{\min_{o \notin C_k} (d(o, c_k))}$$

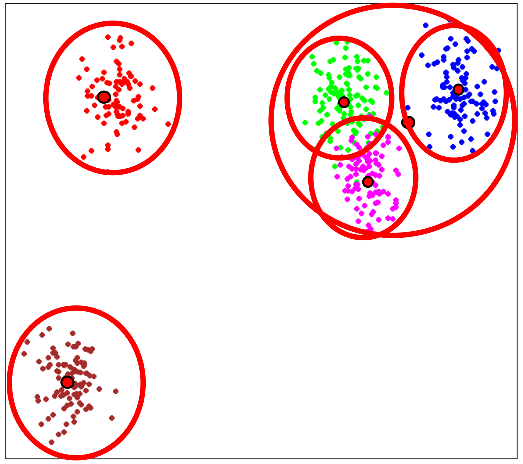
$$OP(C_k) = \lambda \frac{\sum_{o \in C_k} nb_o}{Card(C_k)}$$

Contr 3: deterministic approach

- Combining cluster extractor (contr 1) with evaluation criterion (contr 2)
- First extracted are extracted in beginning (iterative extraction)
- Order is unique according to the criterion
- Allow overlapping

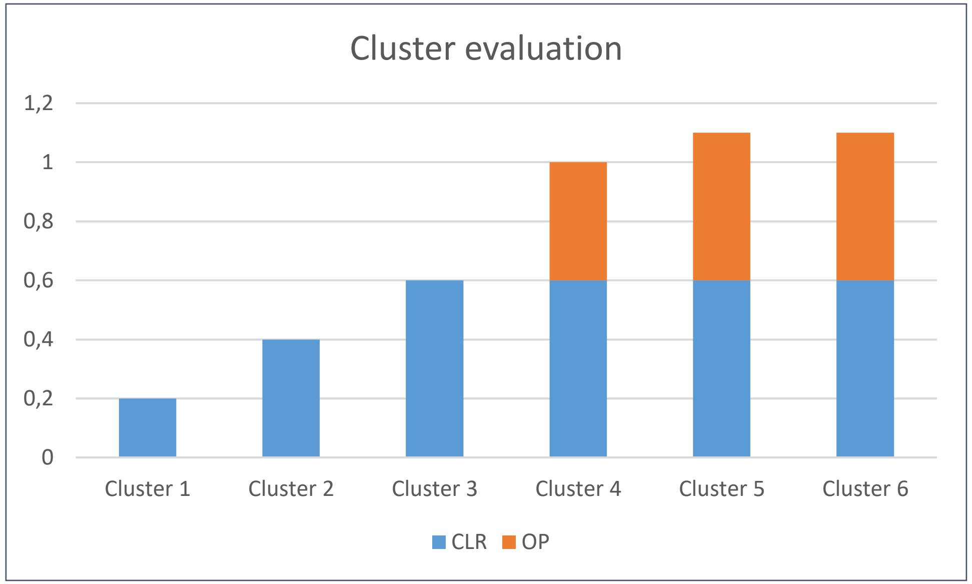
Iterative clustering

Iterative approach



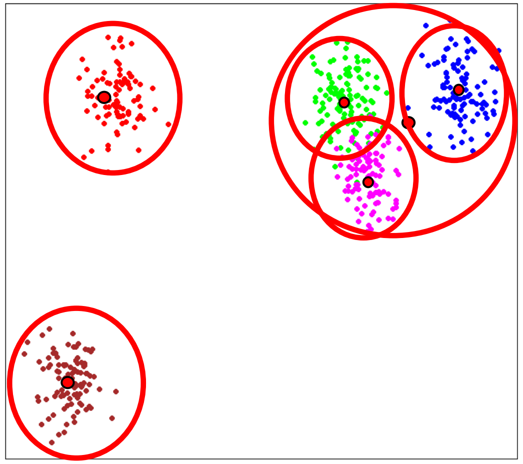
$$CLR\left(C_k\right) = \frac{\max\limits_{o \in C_k} \left(d\left(o, c_k\right)\right)}{\min\limits_{o \notin C_k} \left(d\left(o, c_k\right)\right)}$$

$$OP\left(C_k\right) = \lambda \frac{\sum\limits_{o \in C_k} nb_o}{Card(C_k)}$$



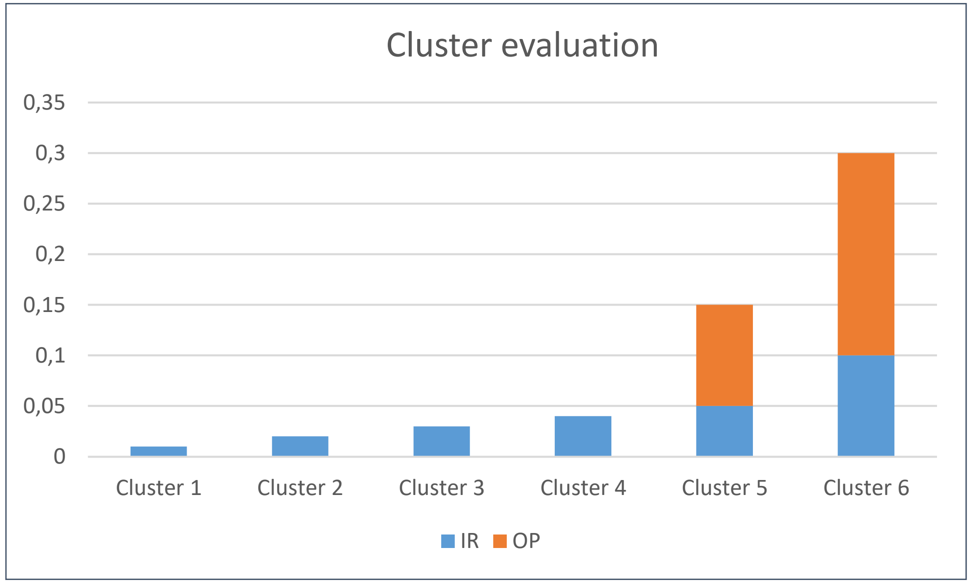
Iterative clustering

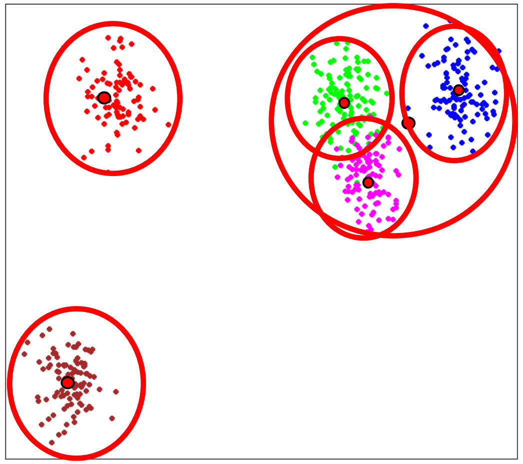
Iterative approach



$$CLR(C_k) = \frac{\max\limits_{o \in C_k}(d(o, c_k))}{\min\limits_{o \notin C_k}(d(o, c_k))}$$

$$OP(C_k) = \lambda \frac{\sum\limits_{o \in C_k} nb_o}{Card(C_k)}$$





$$IR(C_k) = \frac{Card(D) \sum\limits_{o \in C_k} d(o, c_k)^2}{Card(C_k) \sum\limits_{o \in D} d(o, g)^2}$$

$$OP(C_k) = \lambda \frac{\sum\limits_{o \in C_k} nb_o}{Card(C_k)}$$

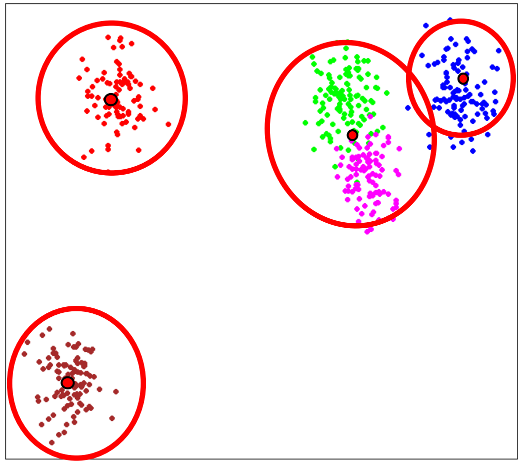
- *Iris de Fisher*, de l'UCI (150 obs., 4 var.) :
  3 clusters
- Accuracy: mean of 100 executions
  - IR: 0.805
  - **CLR: 0.834**
  - K-means: 0.765

Data set	nb. Obs.	nb. var.
WaveFrom	5000	21
Optical Digit	5620	64
Sea Concept	6000	3
Statlog	6435	36
Ovarian	253	15154
Isolet	6238	616
Bicatyeast	419	70

Iterative clustering

Semi-interactive approach

Contr 4



- Active user
- Iterative clustering
- 2D projection

- Modify the automatic proposition
- Propose other centers
- Visual perception

→ Interactive clustering

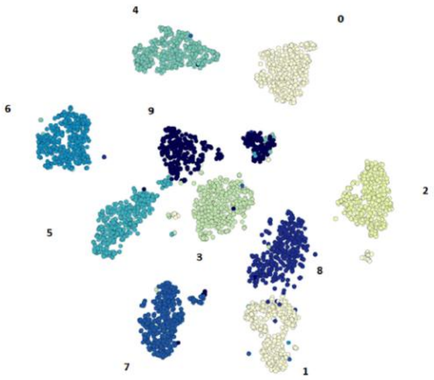
Iterative clustering

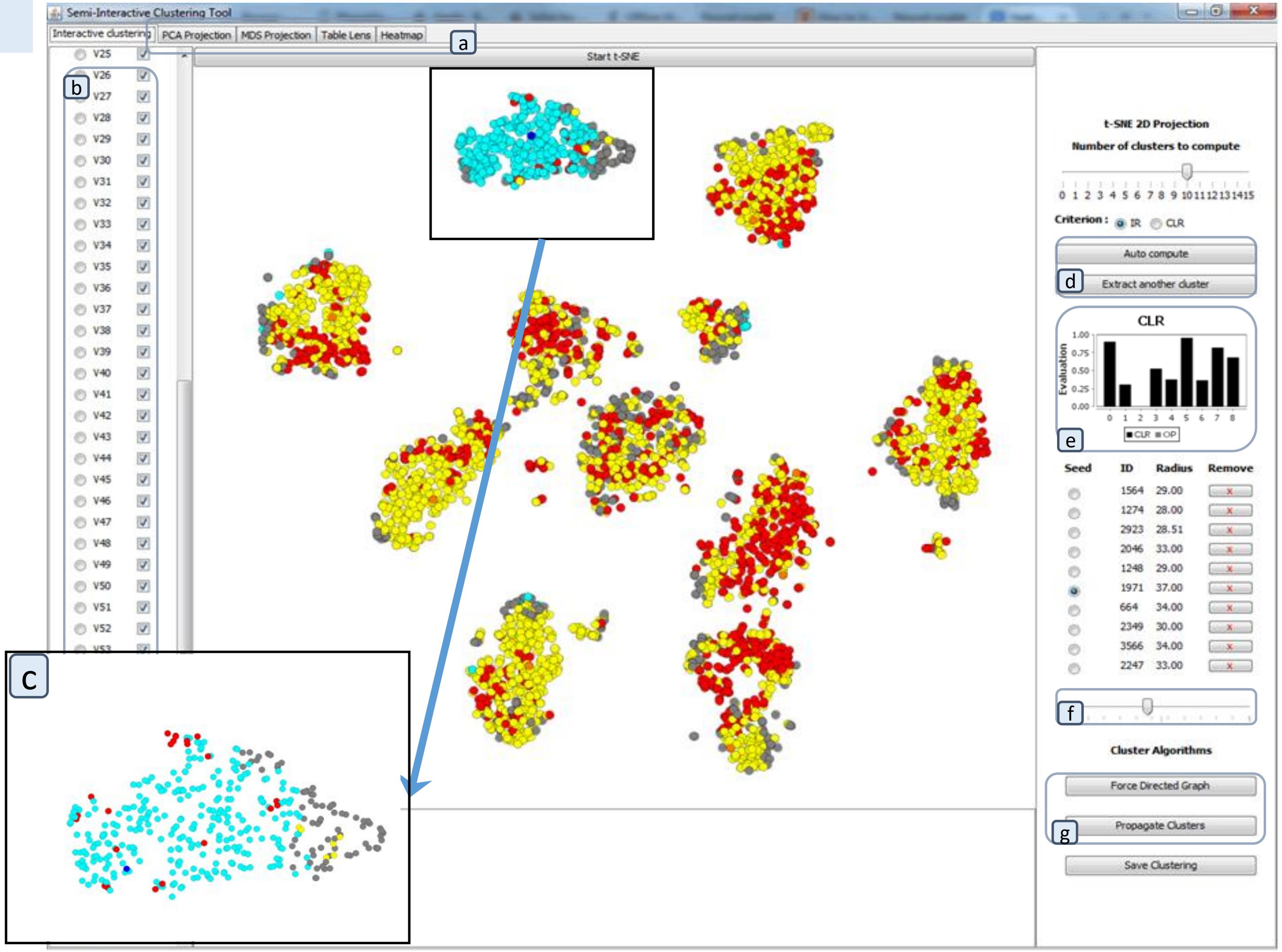
Semi-interactive approach

Contr 4

Data set :
OpticalDigits
5 620 obs.,
64 var.,
10 classes,
UCI ML Repository.

Optical Recognition of Handwritten number of 46 persons





Lydia Boudjeloud-Assala, Philippe Pinheiro, Alexandre Blansché, Thomas Tamisier , Benoît Otjacques, Interactive and iterative visual clustering. Inf. Vis. 15(3): 181-197 (2016)

Iterative clustering

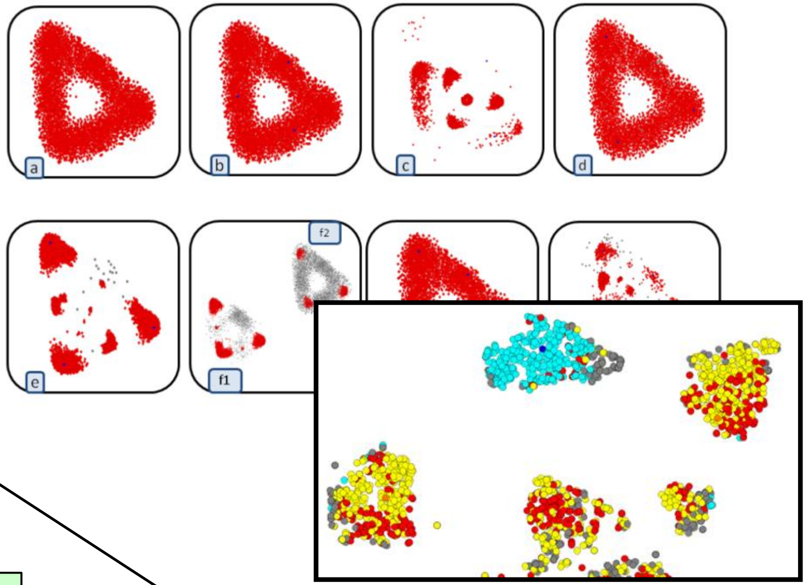
Semi-interactive approach

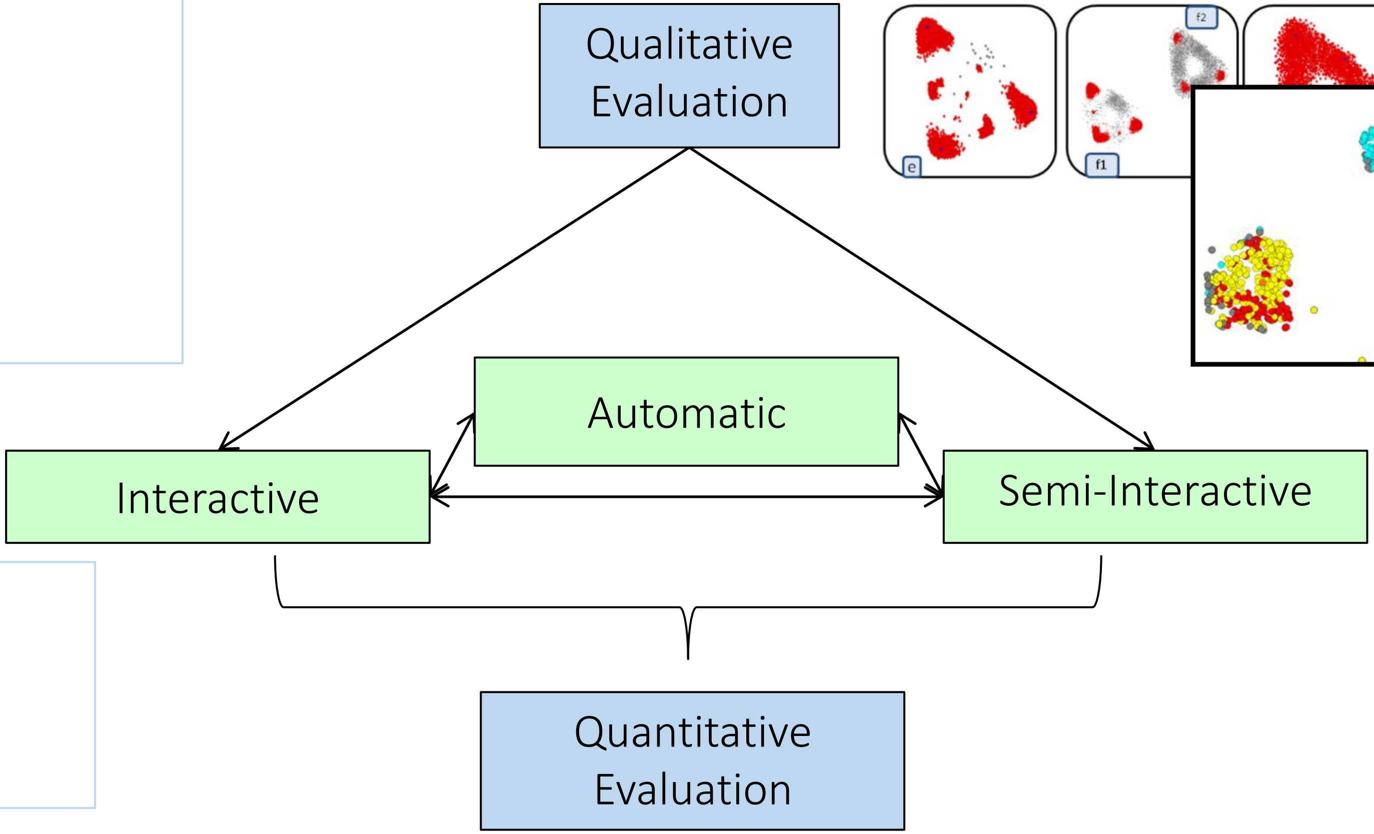
User evaluation

11 participants (Computer science) Age 30-40
- 2 information visualization et interaction,
- 3 data mining

Protocol : 3 scenarios
- purely interactive
- semi-interactive
- purely automatic

clustering / real classes : external criterion





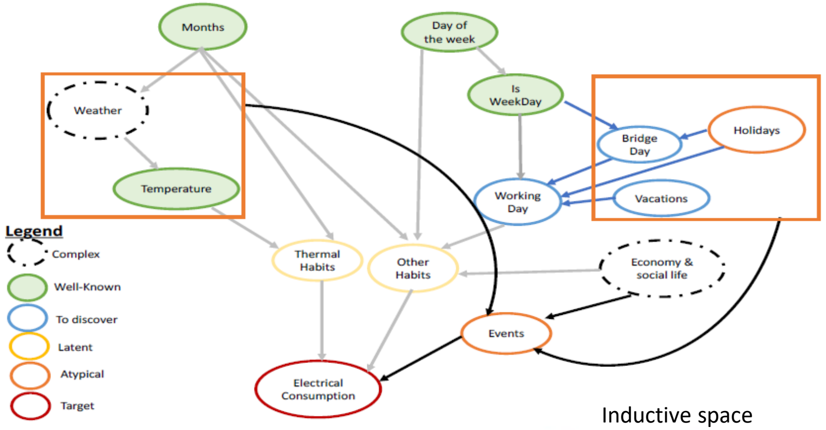
Statistical tests

Shapiro-Wilk test : normality
Student test : means

→ semi-interactive > automatic
semi-interactive > interactive

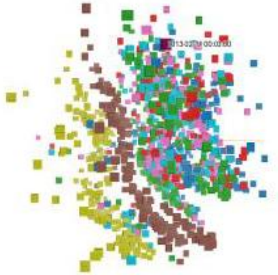
Lydia Boudjeloud-Assala, Philippe Pinheiro, Alexandre Blansché, Thomas Tamisier , Benoît Otjacques, Interactive and iterative visual clustering. Inf. Vis. 15(3): 181-197 (2016)

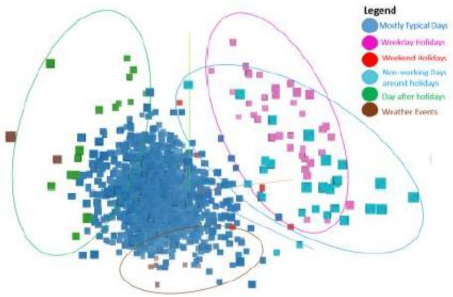
Human-In-The-Loop

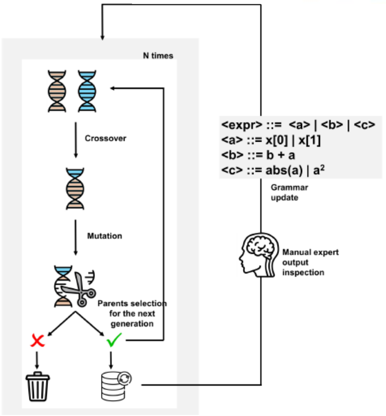


Inductive space

Original space







Réseau de transport d'électricité

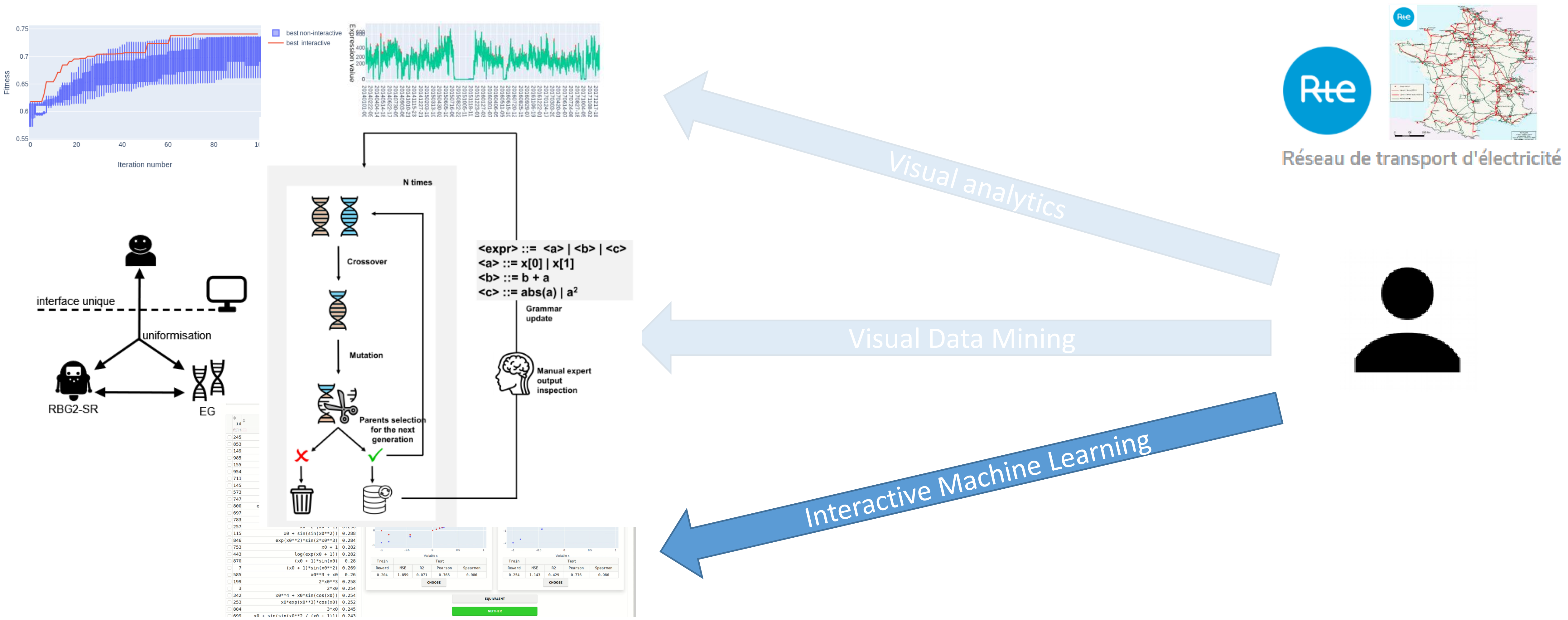
Visual analytics

Visual Data Mining

Interactive Machine Learning

Antoine Marot, Antoine Rosin, Laure Crochepierre, Benjamin Donnot, Pierre Pinson, Lydia Boudjeloud-Assala: Interpreting Atypical Conditions in Systems with Deep Conditional Autoencoders: The Case of Electrical Consumption. ECML/PKDD (3) 2019: 638-654
Laure Crochepierre, Lydia Boudjeloud-Assala, Vincent Barbesant: Interpretable Dimensionally-Consistent Feature Extraction from Electrical Network Sensors. ECML/PKDD (4) 2020: 444-460

Human-In-The-Loop



Laure Crochepierre, Apprentissage automatique interactif pour les opérateurs du réseau électrique, thèse de doctorat de l'Univesité de Lorraine , 27-09-2022

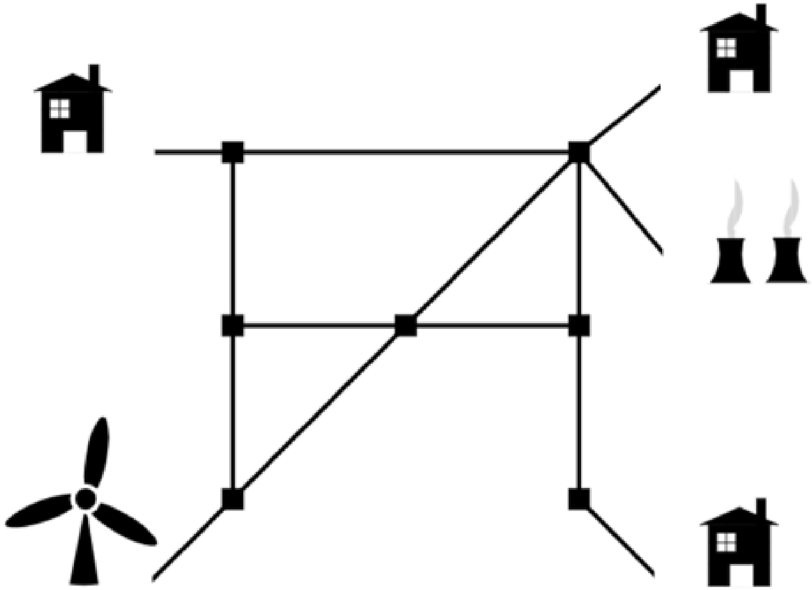
4

French power grid application

Interactive Machine Learning

Explicability interpretability of decision models and prediction

Réseau de transport d'électricité



$$if\left(connected_{or}^{line_i}\right) then\ (watch\ feature_1)$$
$$else\ (watch\ feature_2)$$

$$\sqrt{(p_1+p_2)^2+(q_1+q_2)^2}$$

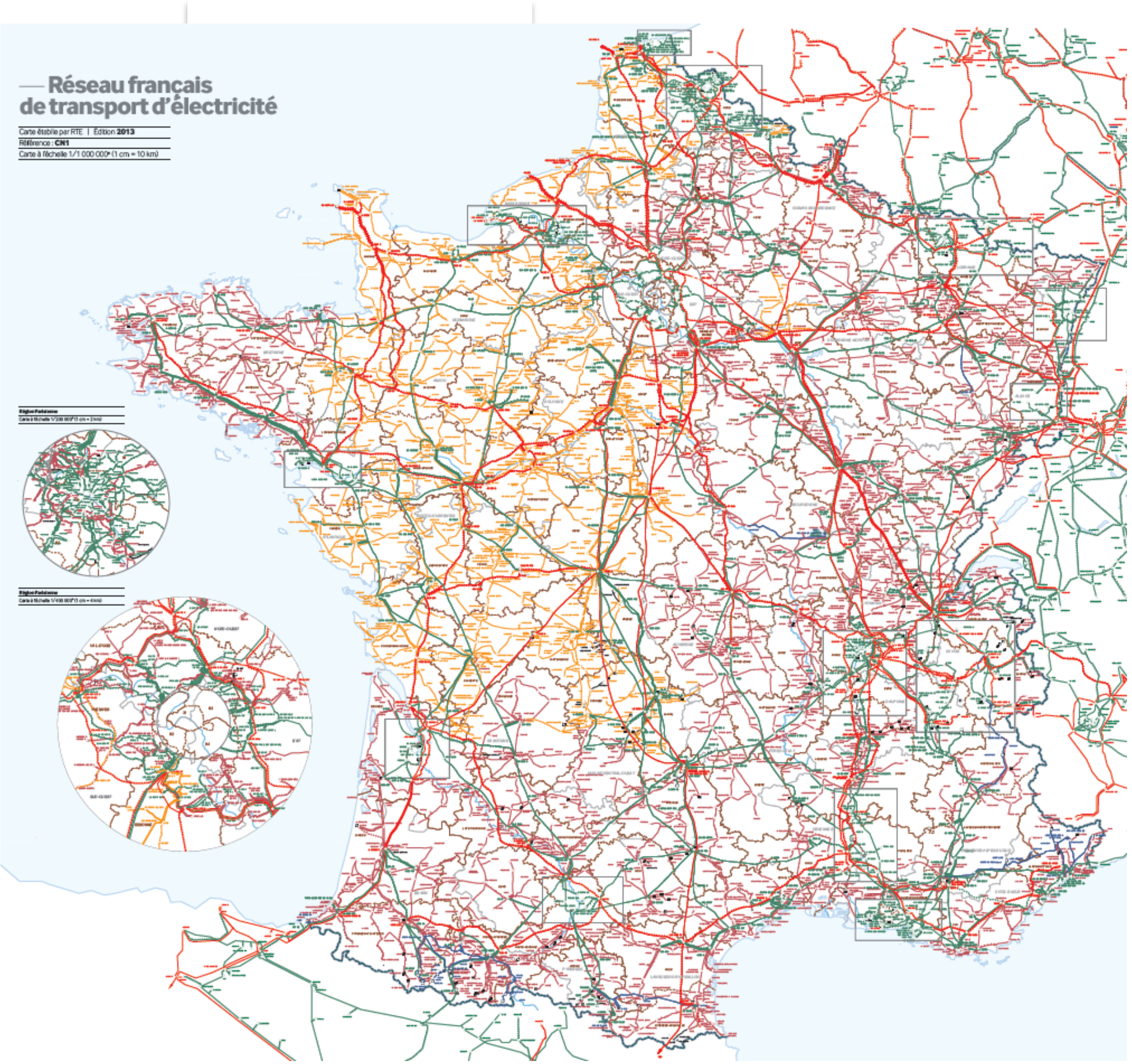
$$\frac{\sqrt{p^2+q^2}}{V} \qquad \sqrt{p^2+q^2}$$

$$i_{frontier\_1} - abs(i_{frontier\_2})$$

- Linear feature extraction is often misinterpreted or overinterpreted
- Non linear methods are even harder to be interpreted for non machine learning experts

L. Crochepierre, L. Boudjeloud-Assala, Vincent Barbesant, Interpretable dimensionally-consistent feature extraction from electrical network sensors, in ECML-PKDD'2020 procedings

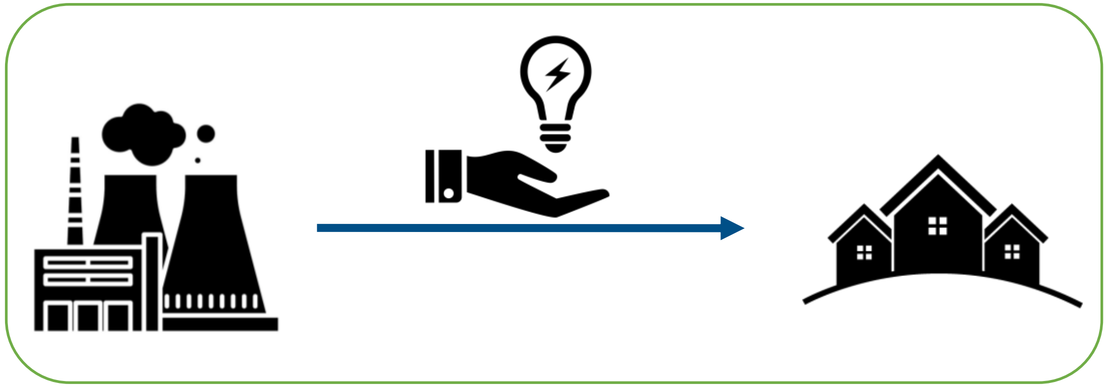
IML for the french power grid



French electrical power grid
(source Rte)

The French high-voltage power grid is made of 100,000 kilometers power lines and interconnected to 6 European countries

Rte (Réseau de transport d'Electricité) responsible for its operation, maintenance and development



Motivation: monitoring the power grid

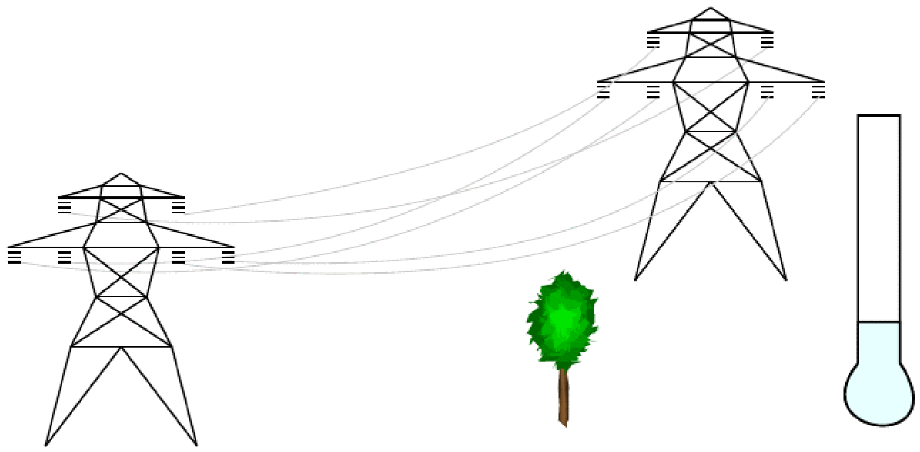
Rte Control room

Anticipated increase in the difficulty to operate the grid (renewable, cross-border exchanges, etc.)

Need for synthesis in the information provided to operators

There will always be operators to *"drive"* the network.

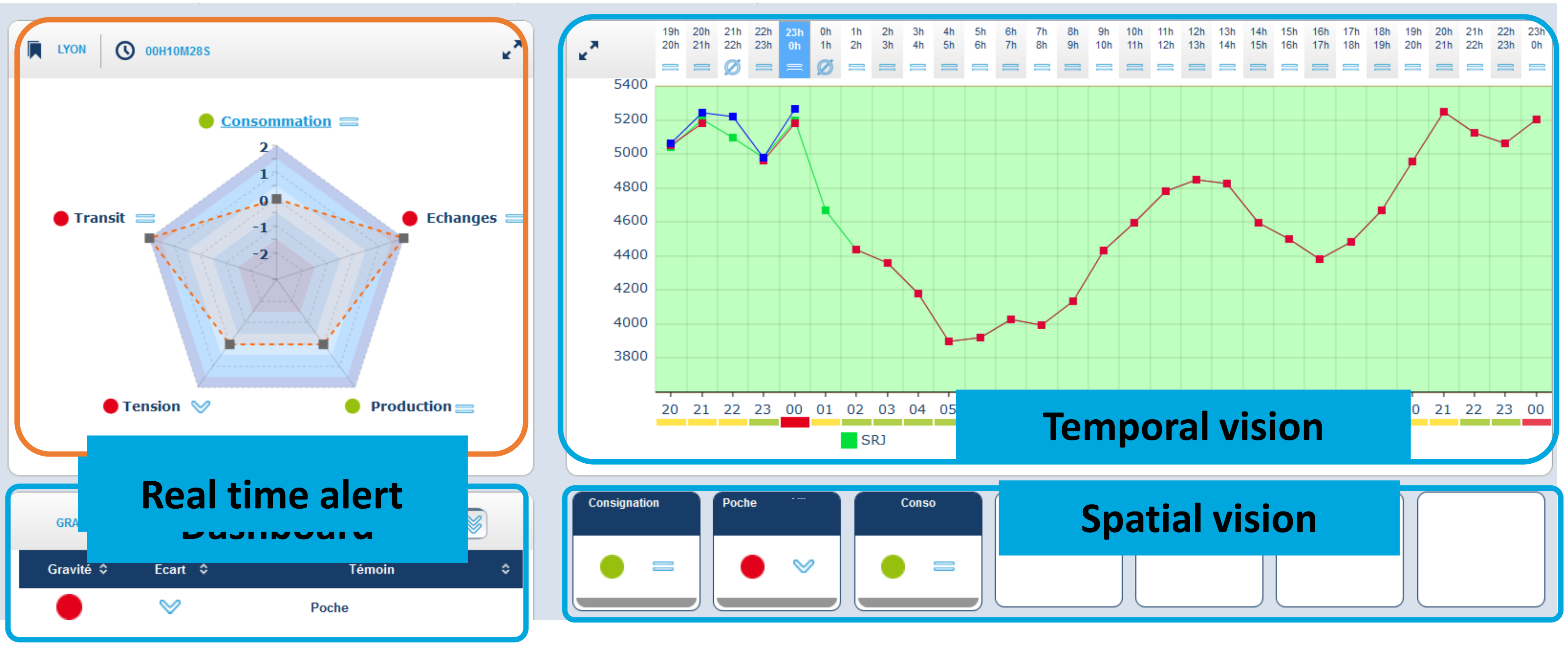
Study of electrical transits



A thermal limit is the physical threshold above which a short circuit can occur, which would endanger people in the vicinity and could damage equipment

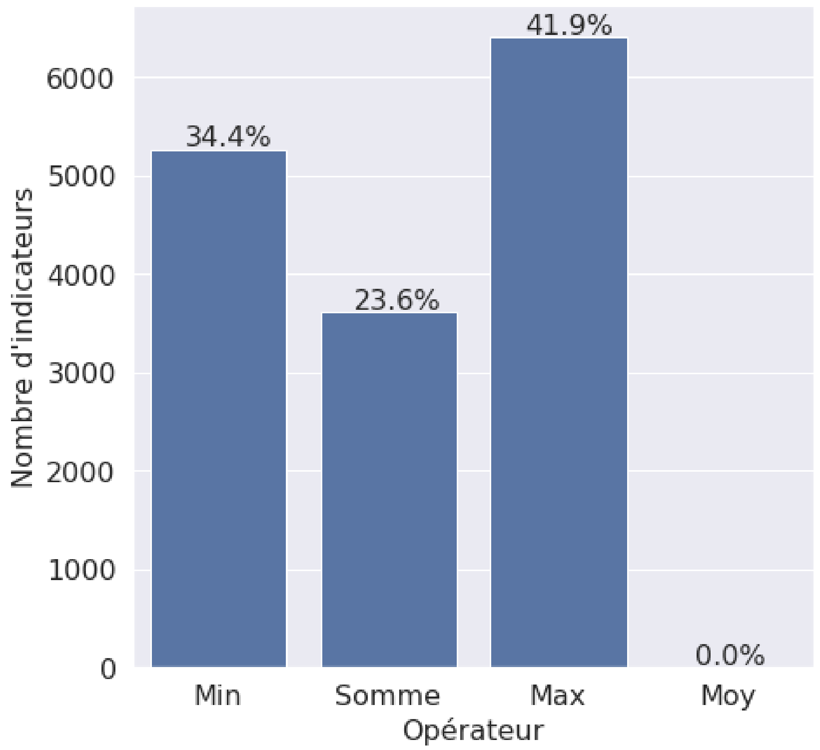
Create indicators with a physical meaning, in order to summarize the state of the network to operators

COMPARE Industrial Software

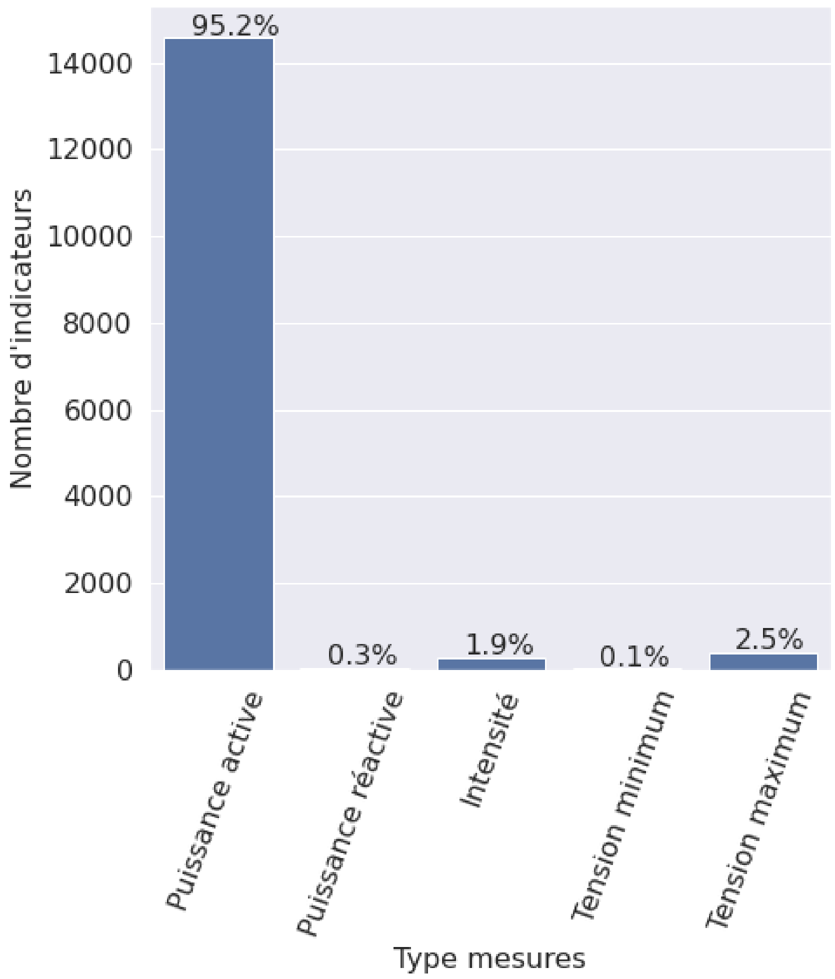


- Created by Rte to summarize the state of a zone to dispatchers and identify the origin of a constraint
- Powered by forecasting and measured data on the grid
- Commissioned in 2014
- Approximately 200 users

Indicators created manually by experts



$$\max(p_{production_1}, p_{production_2})$$



Statistics of existing indicators (October 2019)

Industrial constraints

Create indicators with a physical meaning, in order to summarize the state of the network to operators and going beyond the existing

Propose ***interpretable*** solutions for experts

Take into a count expert ***knowledge*** and physics aspect

Reusability

Minimize ***expert user*** involvement

Objective

Definition:

Find a mathematical expression f between a set of observations $X \in \mathbb{R}^{M \times D}$ with D variables and M observations, and a target variable $y \in \mathbb{R}^{M}$:

$f(x)$=y such that f is *interpretable* by a human

- **Interpretability** is characterized by two elements:
  - The physical (or trade) **sense**
  - The **legibility** (**readability**) of the function

- A symbolic expression created will be used as an indicator

Symbolic regression

Definition:

Find a mathematical expression f between a set of observations $X \in \mathbb{R}^{M \times D}$ with D variables and M observations, and a target variable $y \in \mathbb{R}^{M}$:

$$f(x) = y$$

Multiple methods for regression

- Genetic Programming (Koza, 1990)
- Divide and conquer : AIFeyman (Udrescu et al., 2020)
- Deep neural networks (Biggio et al., 2020)
- Basic expansion: FFX (McConaghy, 2011)

Interpretability criterion: Use of grammatical structure

Biggio, L., Bendinelli, T., Lucchi, A. et Parascandolo, G. (2020). A seq2seq approach to symbolic regression. In the International Workshop on Learning Meets Combinatorial Algorithms (LMCA) co-located with Neurips

Koza, J. R. (1990). Concept Formation and Decision Tree Induction Using the Genetic Programming Paradigm. In Parallel Problem Solving from Nature, 1st Workshop, PPSN I, Proceedings, volume 496, pages 124–128.

McConaghy, T. (2011). FFX : Fast, Scalable, Deterministic Symbolic Regression Technology. In Genetic Programming Theory and Practice IX, pages 235–260. Springer New York.

Udrescu, S.-M. et Tegmark, M. (2020). AI Feynman : A physics-inspired method for symbolic regression. Science Advances, 6(16):eaay2631.

Definition of a grammar

Definition

Non-contextual grammar G (Chomsky, 1959) is a tuple :

$$G = (\sigma_{start}, \sigma_{nt\in NT}, \sigma_{t\in T}, \rho) \text{ where } \rho \text{ is } \sigma_{NT} ::= rule_1 \,|\, .. | rule_K$$

Starting Symbol

Symbols Non-Terminals

Symbols Terminals

Production rules

Language $L(G)$ with symbolic expressions $f \in L$ to represent physical constraints and expert knowledge

Grammar guided symbolic regression

- Grammatical evolution (O'Neill and Ryan, 2001),
- Programmatically Interpretable reinforcement learning (Verma et al., 2018)

Chomsky, N. (1959). On certain formal properties of grammars. Information and control, 2(2):137–167.

O'Neill, M. et Ryan, C. (2001). Grammatical evolution. IEEE Transactions on Evolutionary Computation, 5(4):349–358.

Verma, A., Murali, V., Singh, R., Kohli, P. et Chaudhuri, S. (2018). Programmatically Interpretable Reinforcement Learning. In Proceedings of the ICML, pages 5045–5054

Generating an expression from a grammar

Grammar Example

$< expr > ::= < a > \mid < b >$

$< a > ::= < c > * < c > \mid < c > / < c >$

$< b > ::= < c > + < c > \mid < c > - < c >$

$< \boldsymbol{c} > ::= \boldsymbol{x_0} \mid \boldsymbol{x_1} \mid \boldsymbol{x_2} \mid \boldsymbol{x_3} \mid \boldsymbol{x_4} \mid \boldsymbol{x_5}$

Generate expression "$\boldsymbol{x_1 + x_4}$"

- Symbole courant :

$$< c >$$

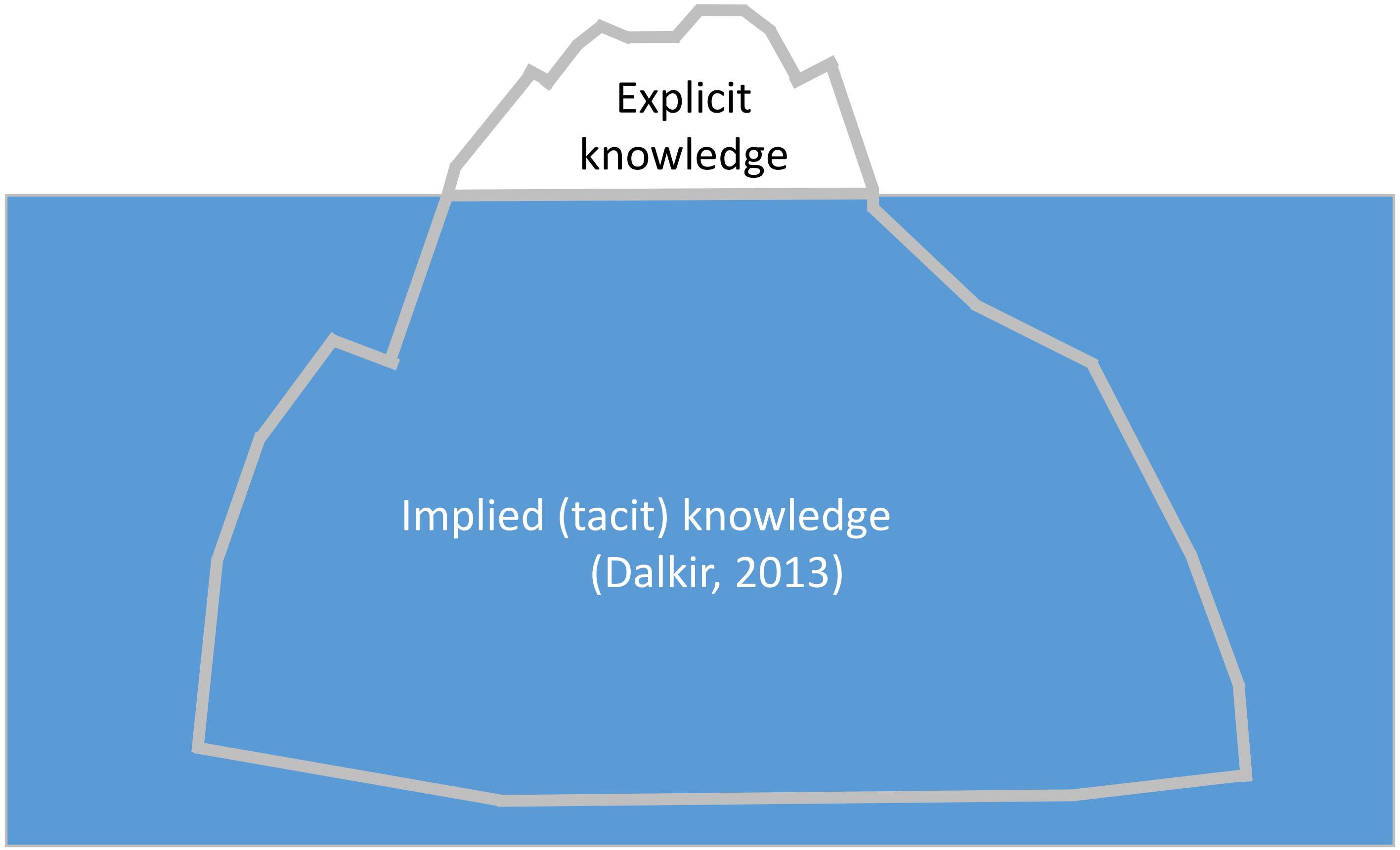
- Règles grammaticales choisies :
  - Itération 1 : $< expr > ::= < b >$
  - Itération 2 : $< b > ::= < c > + < c >$
  - Itération 3 : $< c > ::= x_1$
  - Itération 4 : $< c > ::= x_4$
- Liste de symboles à remplacer :

$$[]$$

- Expression construite :

$$"\, x_1 + x_4 "$$

Knowledge Elicitation



Dalkir, K. (2013). Knowledge management in theory and practice. Routledge.

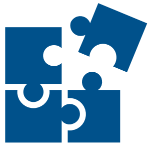
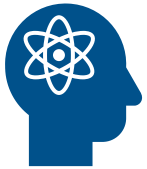
Objectives

Preliminary issues

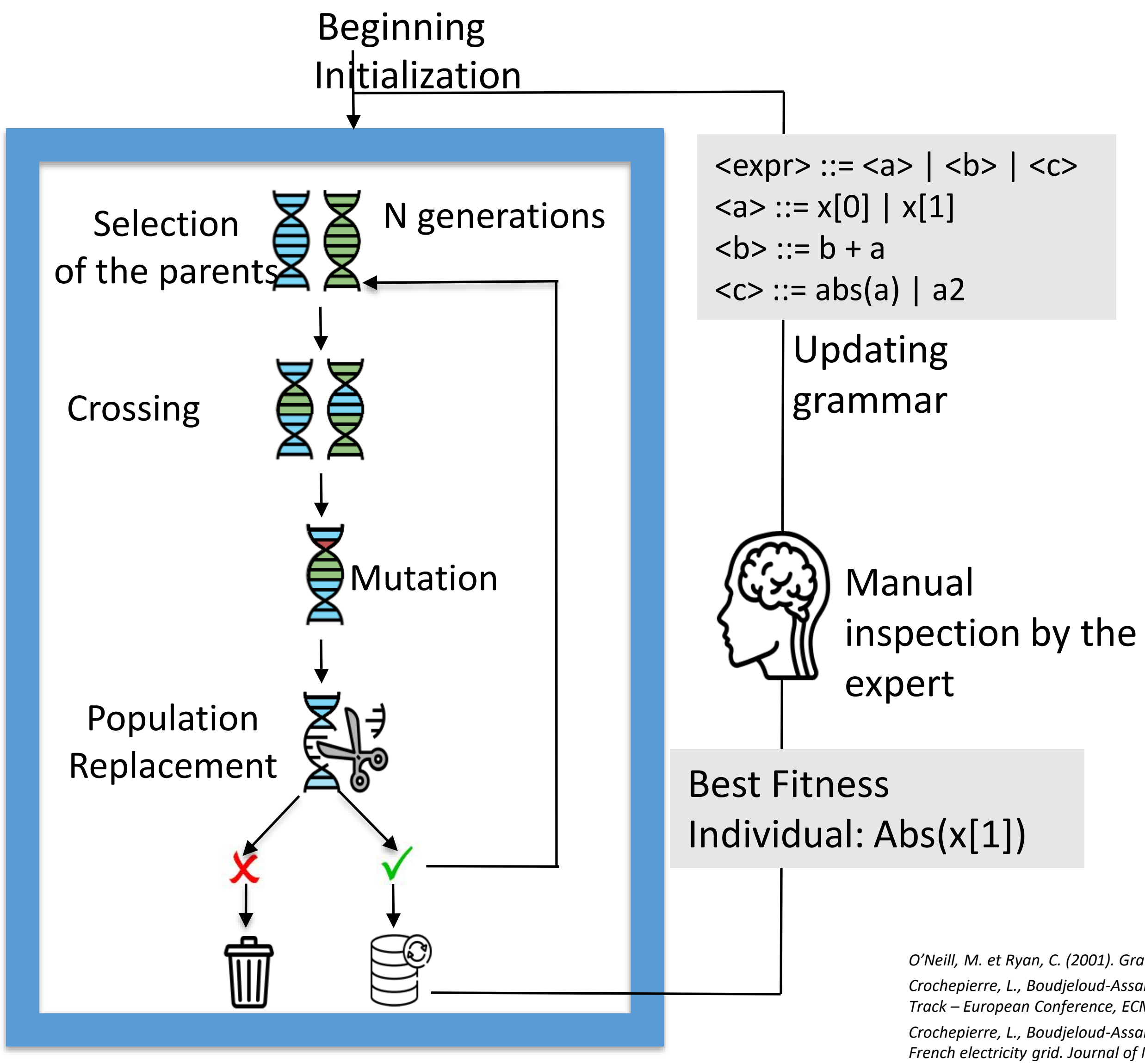
- How can knowledge be represented on the electricity network?
- How do you learn from indicators while using this knowledge?
- What method should be considered to comply with the industrial constraints mentioned?

Selected direction

- Representation of knowledge in the form of grammar
- Learning by evolutionary algorithms
- Compliance with the constraints of interpretability, reusability and knowledge taking into account


Proposed Approach: Interactive Grammatical Evolution



Evolutionary Learning

- Learning based on grammatical evolution(O'Neil, 2001)
  $$fitness(y, y_{hat}) = pearson(y, y_{hat}) + 10^{-8} \times profondeur\_individu$$
- Analysis of best-fit individuals
- Collect expert feedback between learnings
- Incremental grammar improvement between training sessions

O'Neill, M. et Ryan, C. (2001). Grammatical evolution. IEEE Transactions on Evolutionary Computation, 5(4):349–358.

Crochepierre, L., Boudjeloud-Assala, L. et Barbesant, V. (2020). Interpretable Dimensionally-Consistent Feature Extraction from Electrical Network Sensors. In Applied Data Science Track – European Conference, ECML PKDD 2020, Proceedings, pages 444–460.

Crochepierre, L., Boudjeloud-Assala, L. and Barbesant, V. (2021). Interactive approach to extracting interpretable and explanatory variables for the management of constraints in the French electricity grid. Journal of New Information Technologies, Knowledge Extraction and Management, RNTI-E-37:373–380.

Updating grammar

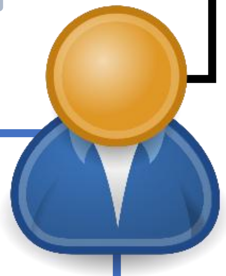
```
<expr> ::= <p> | <q> | <i> | <v>

<p> ::= <p> - <p> | <pop>(<p>, <p>) | <pvar>
<q> ::= <q> - <q> | <pop>(<q>, <q>) | <qvar>
<v> ::= <v> - <v> | <pop>(<v>, <v>) | <vvar>
<i> ::= <i> - <i> | <pop>(<i>,<i>) | <i_frontiervar>

<pop>::= sum | minimum | maximum
```

Initial grammar

Update



- New Units
- New Operations
- Physical Relationships
- Additional knowledge (e.g. output units)

User-Suggested Enhancements

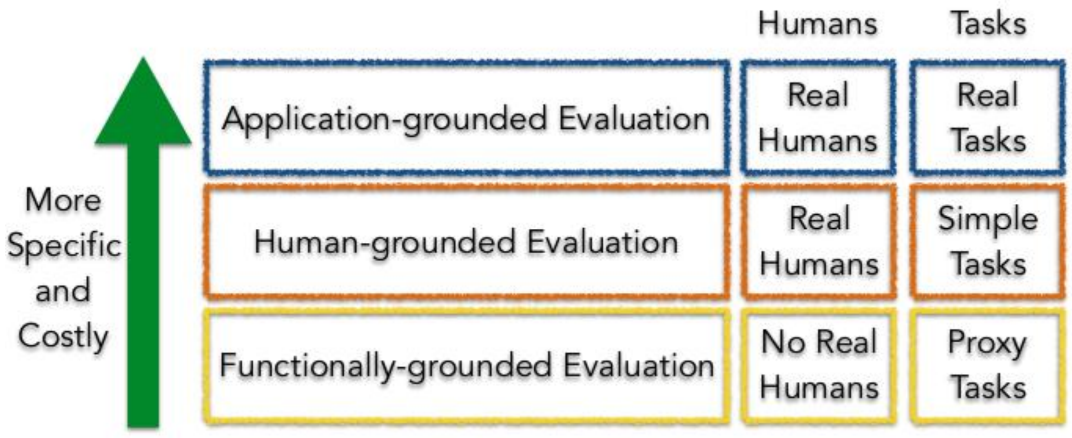
```
<expr> ::= <p> | <s> | <i> | <f> * <expr> | <p>/<v>

<p> ::= <p> - <p> | <pop>(<p>, <p>) | <sop>(<p>) | <p_var>
<q> ::= <q> - <q> | <pop>(<q>, <q>) | <sop>(<q>) | <q_var>
<v> ::= <v> - <v> | <pop>(<v>, <v>) | <sop>(<v>) | <v_var>
<i> ::= <s>/<v> | <pop>(<i>,<i>) | <sop>(<i>) | <i_frontier_var>
<p2> ::= <p>*<p> | square(<p>)
<q2> ::= <q>*<q> | square(<q>)
<v2> ::= <v>*<v> | square(<v>)
<s> ::= sqrt(<p2> + <q2>) |<v> * <i>
<f> ::= <f>*<f> | <p>/<p> | <q>/<q> | <v>/<v>

<pop>::= sum | minimum | maximum
<sop>::= abs | neg | pos
```

Final grammar

Interpretability assessment



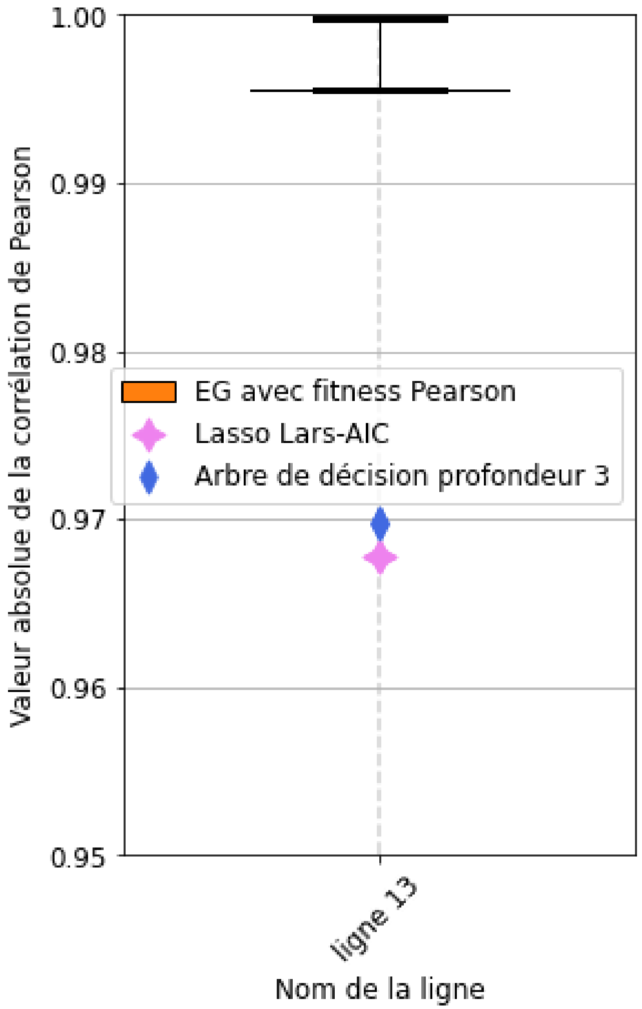
Taxonomy of Method Interpretability Evaluation
(Doshi-velez and Kim, 2017)

2-step interpretability assessment

- "Functionally-grounded": comparison of scores with other interpretable methods: LASSO (Zou et al., 2007), Decision tree (Breiman et al., 1984)
- "Human-grounded": analysis of best fitness expressions in output

Doshi-Velez, F. et Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv :1702.08608.
Zou, H., Hastie, T. et Tibshirani, R. (2007). On the "degrees of freedom" of the lasso. The Annals of Statistics, 35(5):2173–2192.
Breiman, L., Friedman, J. H., Olshen, R. A. et Stone, C. J. (1984). Classification and Regression Trees. Chapman and Hall/CRC.

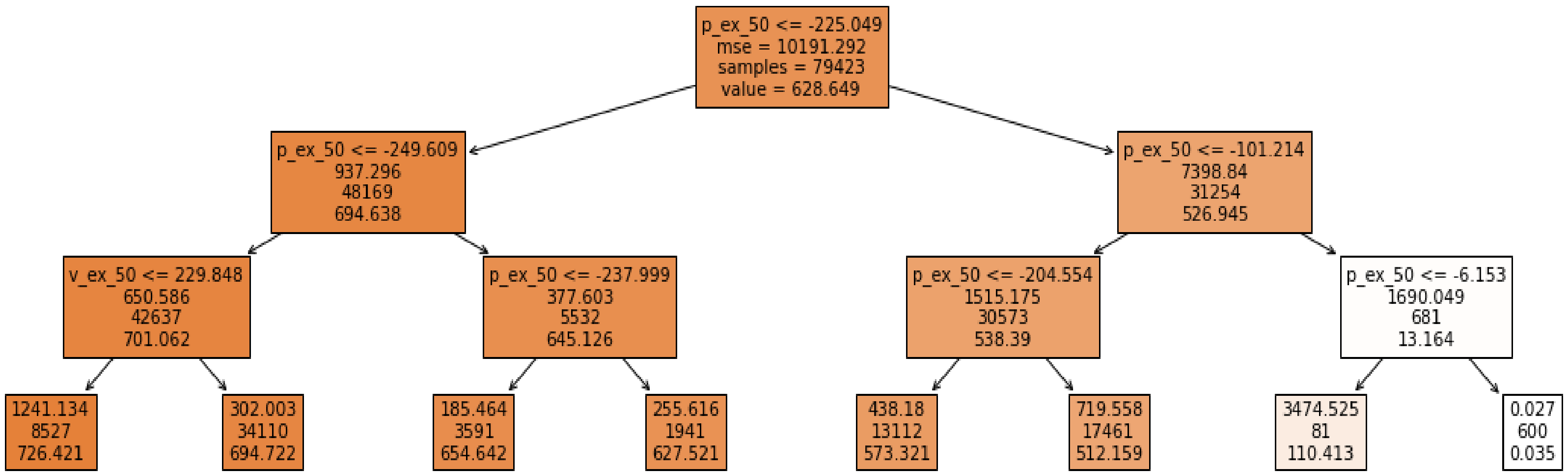
Focus on line 13



- Best expression EG :

$$\frac{\sqrt{p\_ex\_50^2 + q\_ex\_50^2}}{v\_ex\_50}$$

- Best score for decision tree



- LASSO

$$0.68\ \ p\_ex\_10 + 0.01\ \ \ p\_ex\_20 + 0.0009\ \ p\_ex\_61 + 0.14\ \ \ p\_or\_74$$

"Human-grounded" evaluation

Expressions identification

- User-proposed unit validation

$$variations\ autour\ de\ \sqrt{p^2+q^2}\ \ tel\ que\ \sqrt{(p_1+p_2)^2+(q_1+q_2)^2}$$

- Aggregation of variables with a unit useful for explaining transits

$$sum, min, \max d'\ une\ liste\ de\ variables\ de\ puissance\ active\ p$$

- Identifying a global phenomenon: border transit

$$i_{frontière_1} - i_{frontière_2}$$

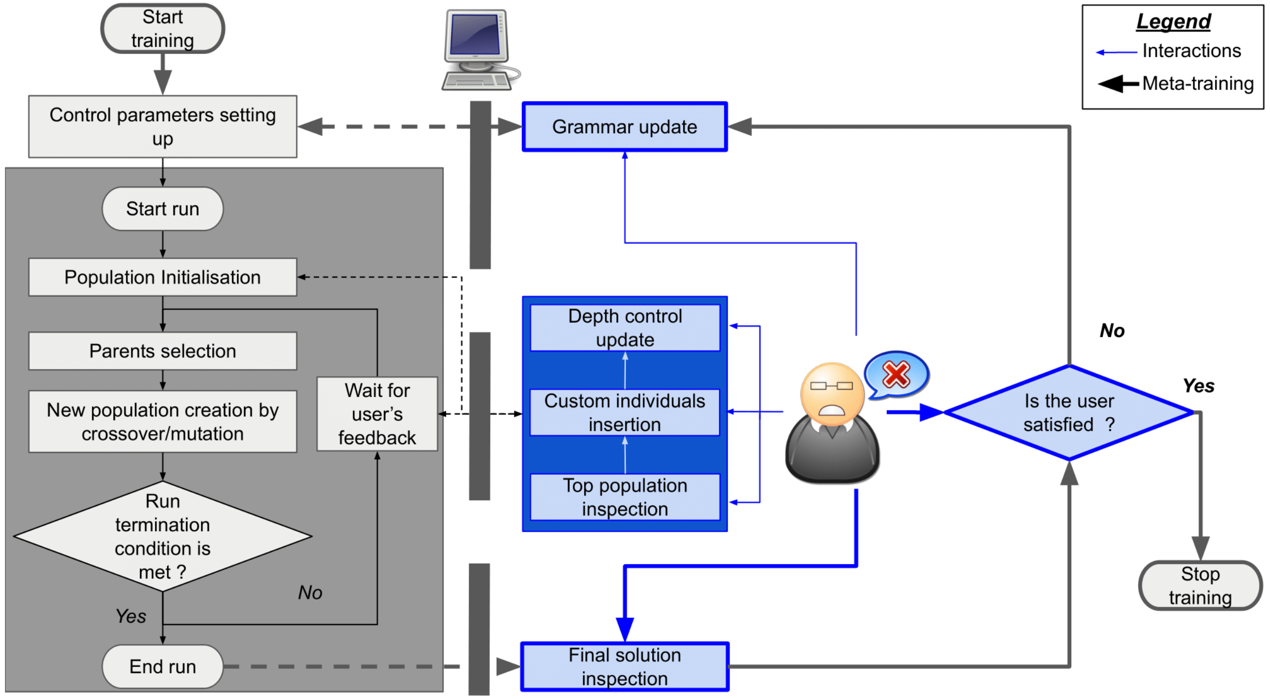
- Introducing topological variables

$$if\,(ligne_1\ \ est\ connectée)\ \ then\ \{sous\_expression_1\}\ \ else\ \{sous\_expression_2\}$$

Expert comment

- Relevant correlation above 0.8 (human acceptance threshold)
- Threshold of 0.9 to use expressions as they are
- Interest in intervening during learning

Multi-level interaction platform



Multi-level interactivity (Crochepierre et al., 2022)

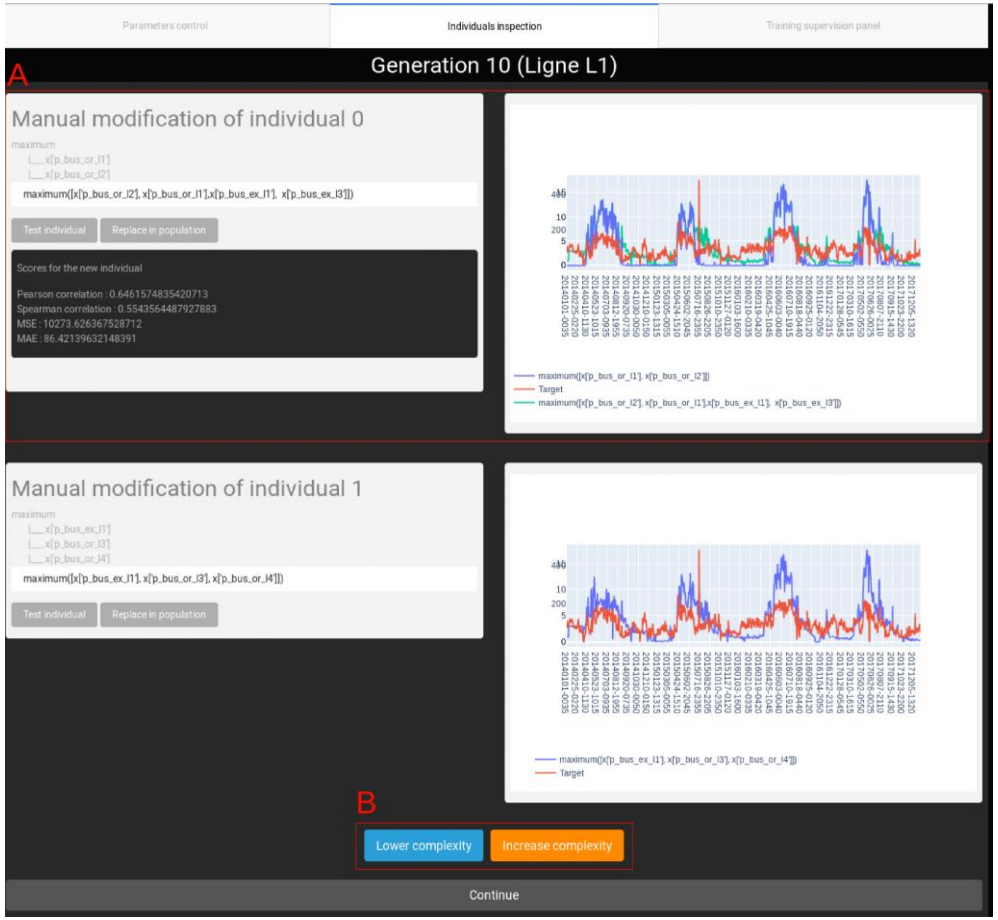
- During training: inspection by the expert and addition of individuals
- Between two learnings: integrating expert comments into the grammar

Implementation de l'interface :

- Dash Plotly Library in Python

Laure Crochepierre, Lydia Boudjeloud-Assala, Vincent Barbesant "Interactive feature extraction using implicit knowledge elicitation : Application to power system expertise." In Hawaii International Conference on System Sciences (HICSS 2022), pages 1–10. ScholarSpace.

Interactions with genetic individuals



Inspection of individuals

- Tree representation
- Error metrics (correlations, EQMs, ...)
- Time series representation

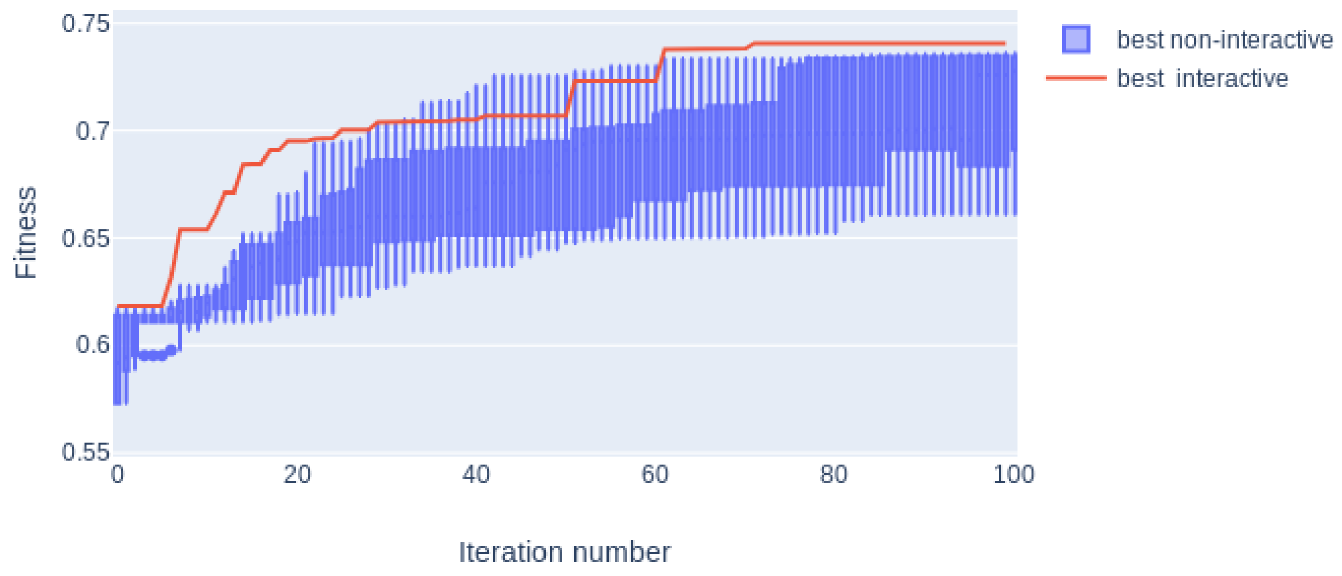
Individual Update

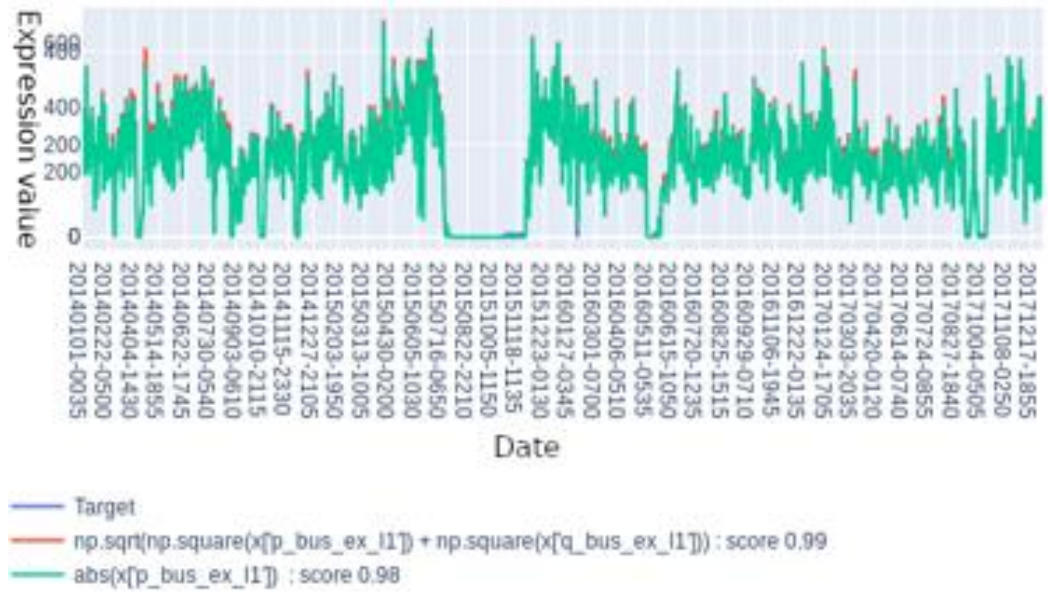
- Text representation
- Comparative test with the initial individual
- Replacement if relevant

Comparison between individuals

- By score and time series

Usage scenarios





Improved performance

- Achieve the best results in fewer iterations
- Discover new expressions

Reduced complexity

- Set preferences between visually similar individuals with identical scores
- Choosing the best balance between error and complexity
- Some more complex lines are more difficult to model
- All the rules of grammar are considered equiprobable

Objectives

Preliminary issues

- How to orient research in the grammatical space?
- How do you learn from indicators while using this knowledge?
- What method should be considered to comply with the industrial constraints mentioned?

Selected direction

- Reinforcement Learning
- Definition of a Markov decision-making process
- Compliance with the constraints of interpretability, reusability and knowledge taking into account

Markov Decision Process definition

Definition (Crochepierre et al, 2022)

▷ State : information on expression under construction

▷ Action : selection of a grammatical rule in G

▷ Reward : $R(f) = \frac{1}{1+EQM\left(y, f(x)\right)}$

Learning distribution of grammar rules $\boldsymbol{G}$

- Neuronal agent π_θ with REINFORCE (Williams, 1992)
- Risk-seeking policy gradient (Petersen et al., 2021)

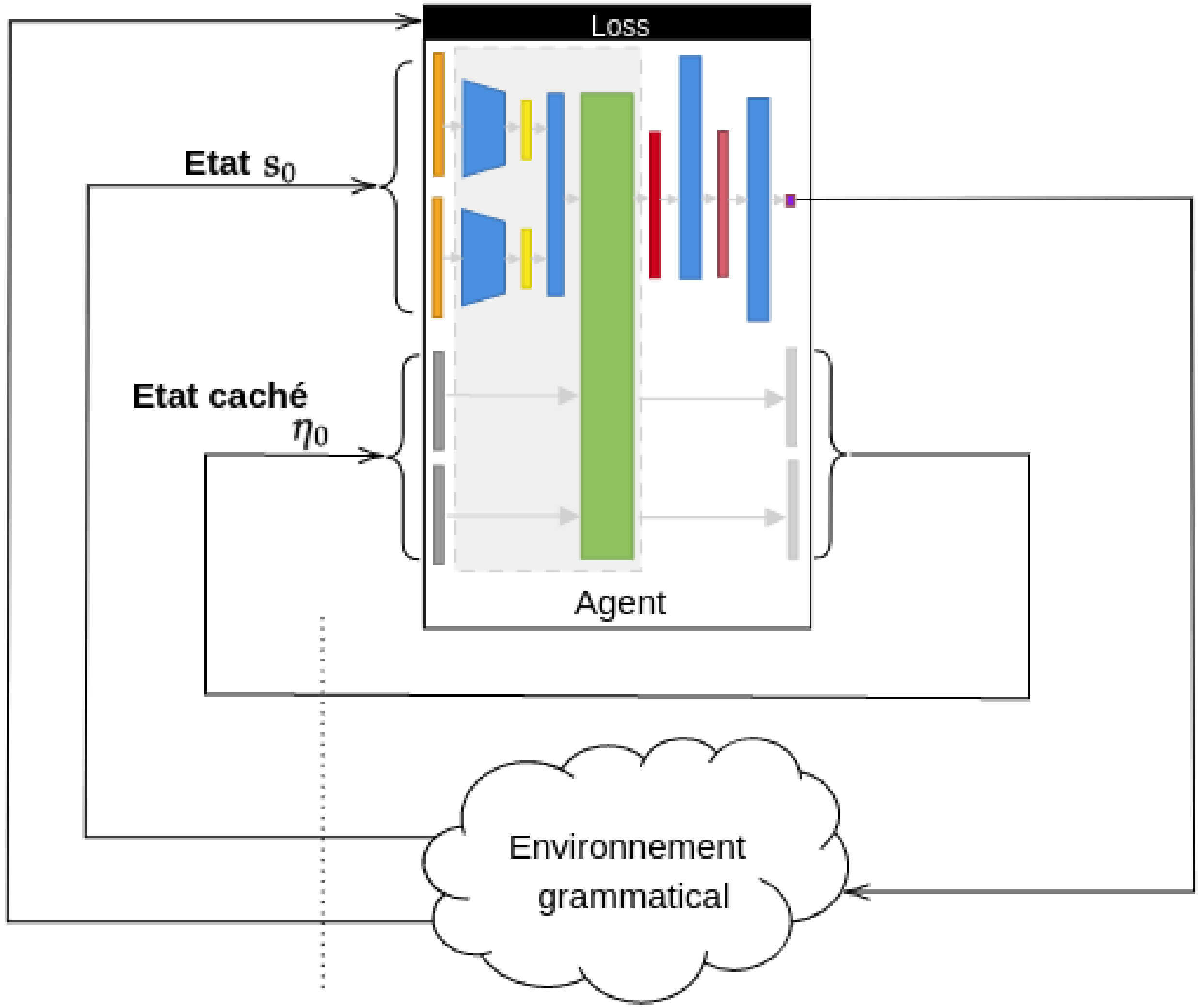
$$J_\theta = \mathbb{E}_{f\sim\pi_\theta}\left[log\left(P_\theta(f)\right)R(f)\middle|R(f) > R_\varepsilon\right]$$

Crochepierre, L., Boudjeloud-Assala, L. et Barbesant, V. (2022a). A Reinforcement Learning Approach to Domain-Knowledge Inclusion Using Grammar Guided Symbolic Regression. arXiv preprint arXiv :2202.04367

Petersen, B. K., Larma, M. L., Mundhenk, T. N., Santiago, C. P., Kim, S. et Kim, J. T.(2021). Deep symbolic regression : Recovering mathematical expressions from data via risk-seeking policy gradients. In 9th International Conference on Learning Representations, (ICLR 2021).

Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine learning, 8(3-4):229–256.

Architecture neuronale



Agent neuronal et environnement de renforcement

Evaluation of the method

Symbolic Regression Benchmarks

- Nguyen : 10 fonctions (Uy et al., 2011)
- Keijzer: 15 functions (Keijzer, 2003)
- Pagie: 1 function (Pagie, 1997)
- Vladislavleva : 8 fonctions (Vladislavleva, 2009)

Comparison with state-of-the-art methods

- Grammatical Evolution (GE): without updating the grammar
- Probabilistic Model Building Genetic Programming (GB-LGP): updating the grammar according to the proportion of rules present in the best individuals

Uy, N. Q., Hoai, N. X., O'Neill, M., McKay, R. I. et Galván-López, E. (2011). Semantically-based crossover in genetic programming : application to real-valued symbolic regression. Genetic Programming and Evolvable Machines, 12(2):91–119.

Keijzer, M. (2003). Improving symbolic regression with interval arithmetic and linear scaling. In European Conference on Genetic Programming, pages 70–82. Springer

Pagie, L. et Hogeweg, P. (1997). Evolutionary Consequences of Coevolving Targets. Evolutionary computation, 5:401–18

Vladislavleva, E. J., Smits, G. F. et den Hertog, D. (2009). Order of Nonlinearity as a Complexity Measure for Models Generated by Symbolic Regression via Pareto Genetic Programming. IEEE Transactions on Evolutionary Computation, 13(2):333–349.

Sotto, L. F. D. P. et de Melo, V. V. (2017). A probabilistic linear genetic programming with stochastic context-free grammar for solving symbolic regression problems. In Proceedings of the Genetic and Evolutionary Computation Conference, pages 1017–1024. ACM

Results

- A benchmark discarded
- Bilateral Mann-Whitney U-test: best score on 22 of 33 benchmarks
- Lower average MSEs
- More Expressions Discovered Accurately

Méthode	U test	EQM moyenne	# solutions exactes découvertes
EG	$+3/\sim5/-25$	1,7	0
GB-LGP	$+1/\sim3/-29$	11,4	0 à 7
RBG2-SR	$\mathbf{+22/\sim7/-4}$	$\mathbf{6.09\times10^{-1}}$	**4 à 14**

Comparison of EQM scores, U-test and number of exact functions discovered on 33 benchmarks (30 learnings)

Autres évaluations

- Définition de l'état étudié par ablation

$$s_h = (a_h^{past}, a_h^{parent}, a_h^{siblings}, d_h, \sigma_h, m_h, \eta_h)$$

- Meilleur compromis erreur-complexité sur le jeu de données Airfoil

Objectives

Preliminary issues

- How can implicit knowledge be integrated into learning?
- Is it possible to reduce the number of user interactions?
- What method should be used to comply with the four industrial constraints?

Selected direction

- Learning on user preferences
- Switching between the REINFORCE algorithm and the use of preferences
- Compliance with the four constraints

Algorithm description

Learning from human preferences on pairs of symbolic expressions

$$f_1 > f_2 \, , (f_1, f_2) \sim \pi_\theta$$

3 strategies

- Preferences sorted into 3 categories (Kuhlman *et al.*, 2018)
- Preferences on pairs (Christiano *et al.*, 2017)
- Direct suggestion for solutions

Objective function

$$J_\theta = \mathbb{E}_{\substack{(f_1, f_2) \sim \pi_\theta \\ R(f_1) > R(f2)}} \left[log\big(P_\theta(f_1)\big)\big(R(f_1)\big) - \; log\big(P_\theta(f_2)\big)\big(R(f_2)\big) \big| R\big(f_{i, i \in \{1,2\}}\big) > R_\varepsilon \right]$$

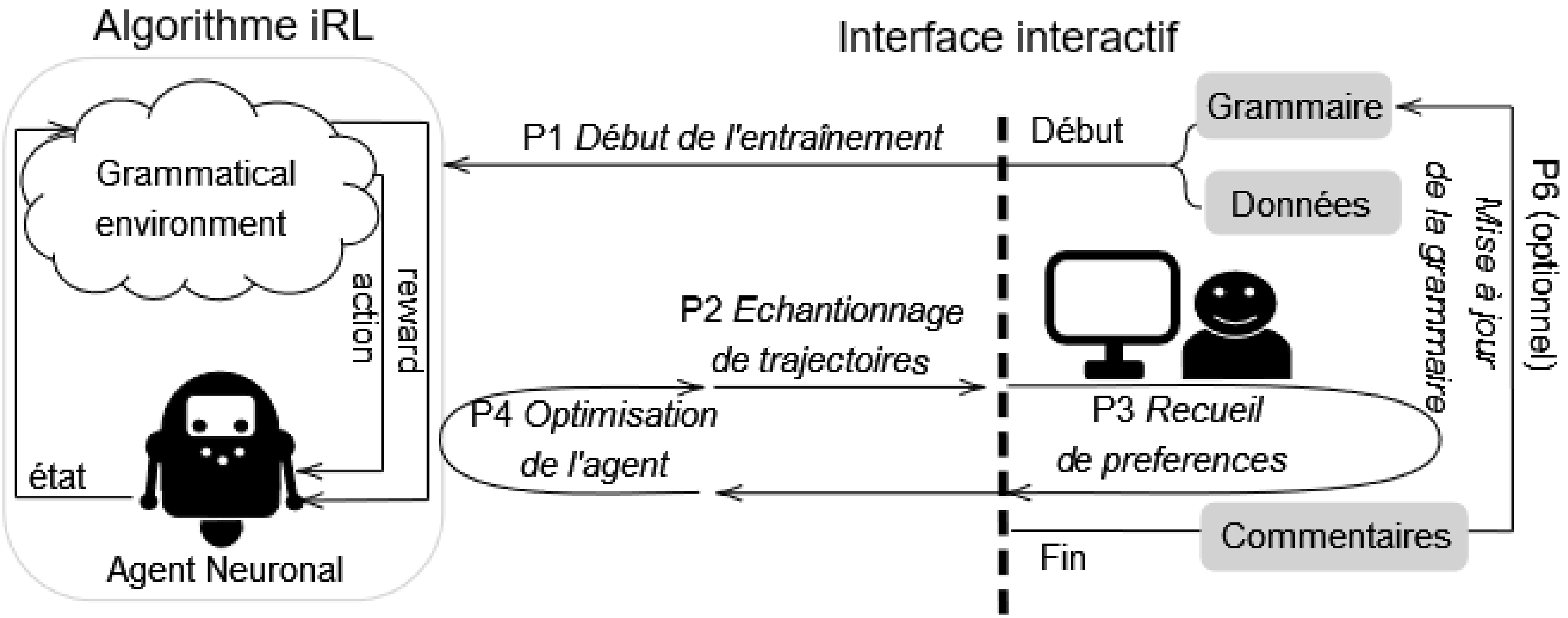
Alternate use of objectives with/without interaction

Wirth, C., Akrour, R., Neumann, G., Fürnkranz, J. et al. (2017). A survey of preference-based reinforcement learning methods. Journal of Machine Learning Research, 18(136):1–46.

Christiano, P. F., Leike, J., Brown, T. B., Martic, M., Legg, S. et Amodei, D. (2017). Deep Reinforcement Learning from Human Preferences. In Advances in Neural Information Processing Systems, volume 30, pages 4299–4307

Kuhlman, C., Valkenburg, M. V., Doherty, D., Nurbekova, M., Deva, G., Phyo, Z., Rundensteiner, E. A. et Harrison, L. (2018). Preference-driven Interactive Ranking System for Personalized Decision Support. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, (CIKM 2018), pages 1931–1934.

Platform Overview



Interactive preference-based reinforcement learning

Laure Crochepierre, Lydia Boudjeloud-Assala, Vincent Barbesant "Interactive reinforcement learning for symbolic regression from multi-format human-preference feedbacks.", In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI 2022), pages 5900–5903. International Joint Conferences on Artificial Intelligence Organization. Demo Track.

Applications

Symbolic regression benchmark (Uy et al., 2011):

- The target symbolic expression is known to the research team but not to the user
- Visual and score-based exploration
- A priori knowledge (or not) of the solution

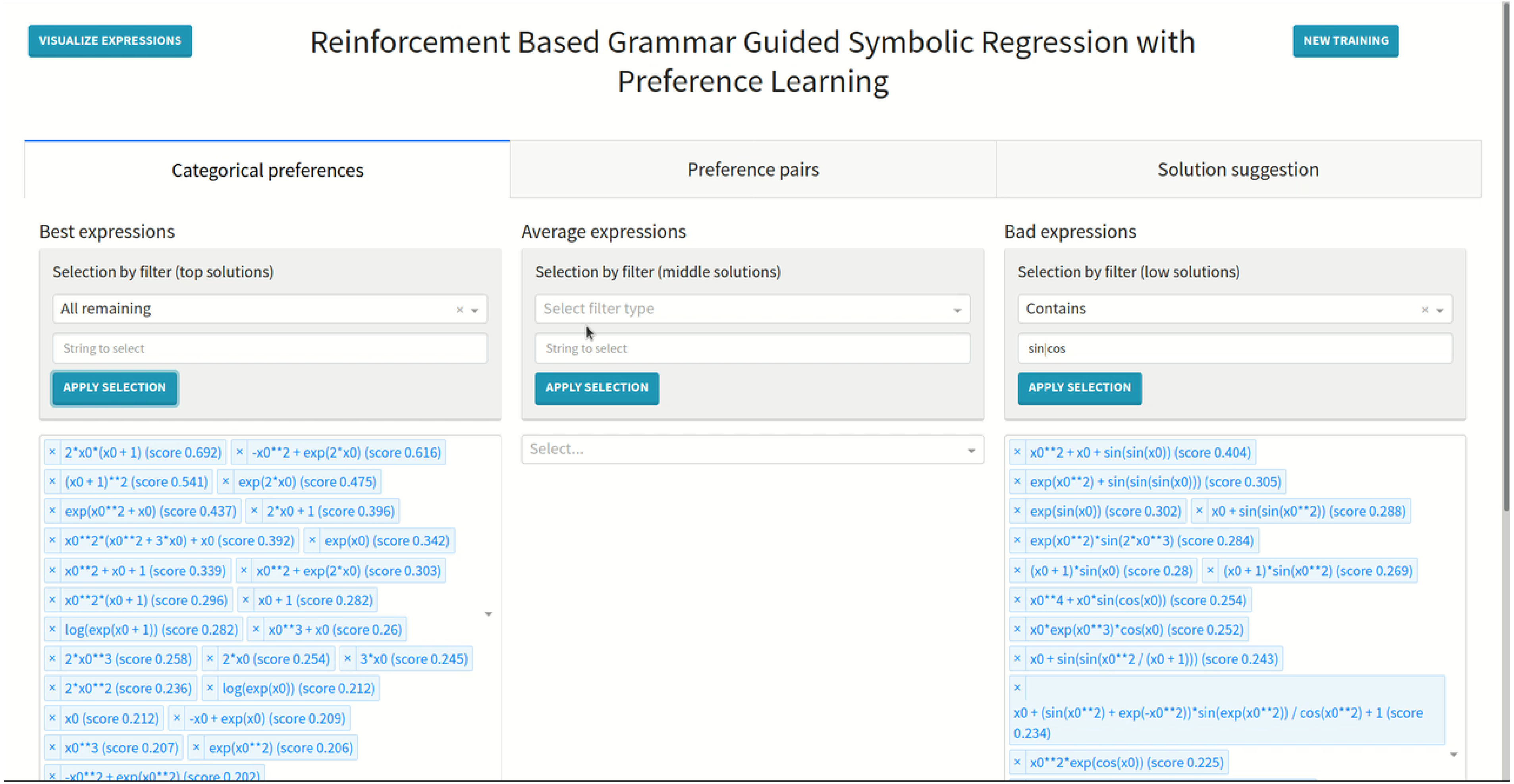
Application to power system :

- Symbolic expression exact target unknown
- Exploratory task, representative of the industrial objective
- Data simulated with Grid2Op * library on an IEEE14 network

Uy, N. Q., Hoai, N. X., O'Neill, M., McKay, R. I. et Galván-López, E. (2011). Semantically-based crossover in genetic programming : application to real-valued symbolic regression. Genetic Programming and Evolvable Machines, 12(2):91–119

* Github repo : https://GitHub.com/rte-france/grid2op

Tab 1 – Categorical Preferences



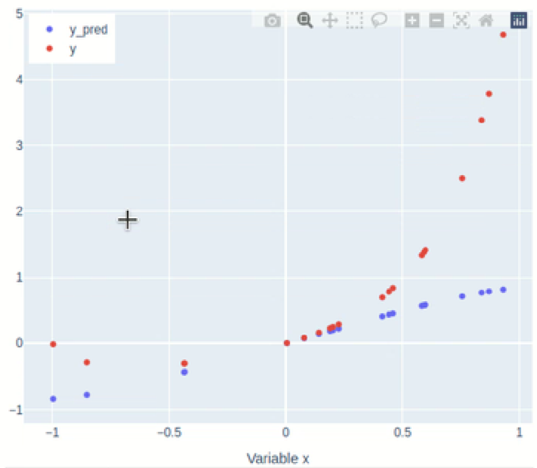
Tab 2 – Preference Pairs

id	Expression	Reward
245	2*x0*(x0 + 1)	0.692
853	-x0**2 + exp(2*x0)	0.616
149	(x0 + 1)**2	0.541
985	exp(2*x0)	0.475
155	exp(x0**2 + x0)	0.437
954	x0**2 + x0 + sin(sin(x0))	0.404
711	2*x0 + 1	0.396
145	x0**2*(x0**2 + 3*x0) + x0	0.392
573	exp(x0)	0.342
747	x0**2 + x0 + 1	0.339
800	exp(x0**2) + sin(sin(sin(x0)))	0.305
697	x0**2 + exp(2*x0)	0.303
783	exp(sin(x0))	0.302
257	x0**2*(x0 + 1)	0.296
115	x0 + sin(sin(x0**2))	0.288
846	exp(x0**2)*sin(2*x0**3)	0.284
753	x0 + 1	0.282
443	log(exp(x0 + 1))	0.282
870	(x0 + 1)*sin(x0)	0.28
7	(x0 + 1)*sin(x0**2)	0.269
585	x0**3 + x0	0.26
199	2*x0**3	0.258
3	2*x0	0.254
342	x0**4 + x0*sin(cos(x0))	0.254
253	x0*exp(x0**3)*cos(x0)	0.252
884	3*x0	0.245
699	x0 + sin(sin(x0**2 / (x0 + 1)))	0.243

Iteration n°0

Pair [365, 958]

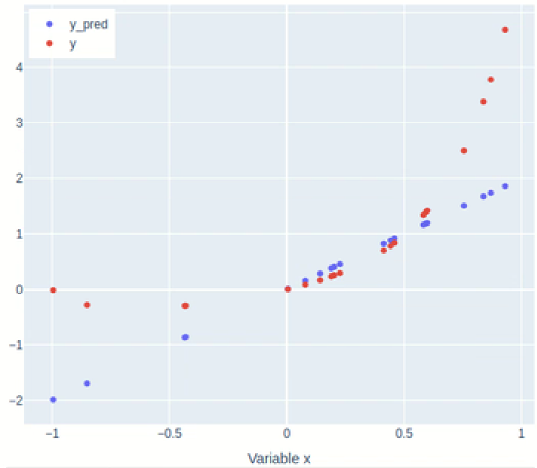
sin(x0**2) / x0



Train	Test			
Reward	MSE	R2	Pearson	Spearman
0.204	1.859	0.071	0.765	0.986

CHOOSE

2*x0



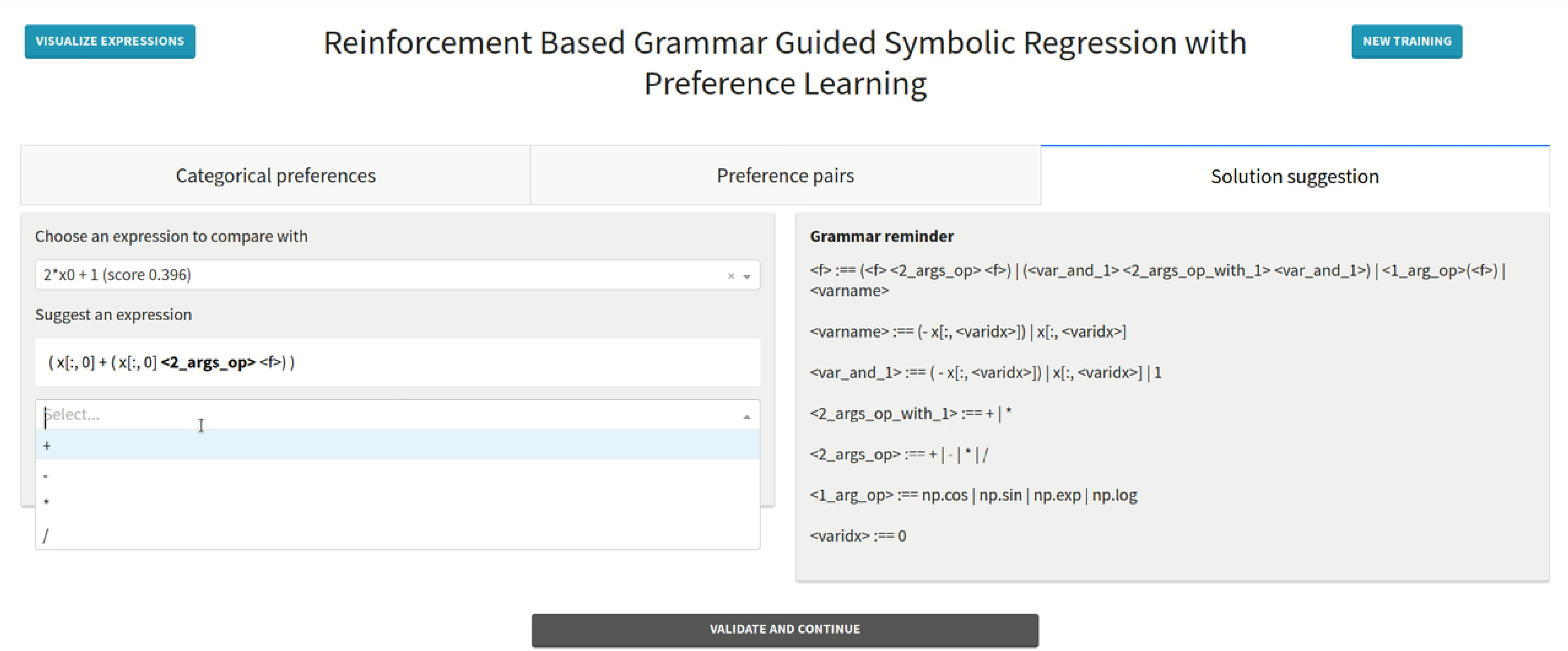
Train	Test			
Reward	MSE	R2	Pearson	Spearman
0.254	1.143	0.429	0.776	0.986

CHOOSE

EQUIVALENT

NEITHER

Tab 3 – Suggestion of expressions



Preliminary interaction

User Profile

- User U1: No knowledge of the form of the function
- User U2: Knowledge of the polynomial form of the function
- User U3: Knowledge of the exact form of the function

Utilisateur	Nombre d'interactions	Retrouve la solution exacte
U1	15	✗
U2	21	✗
U3	16 et 26	✓

4 common strategies used by users

- Prefer expressions with the highest score
- Prefer the best performer most of the time (with a random probability $p \in \{0.2, 0.5, 0.8\}$)
- Avoid cos/sin operations
- Avoid cos/sin/exp operations

Results of the simulations

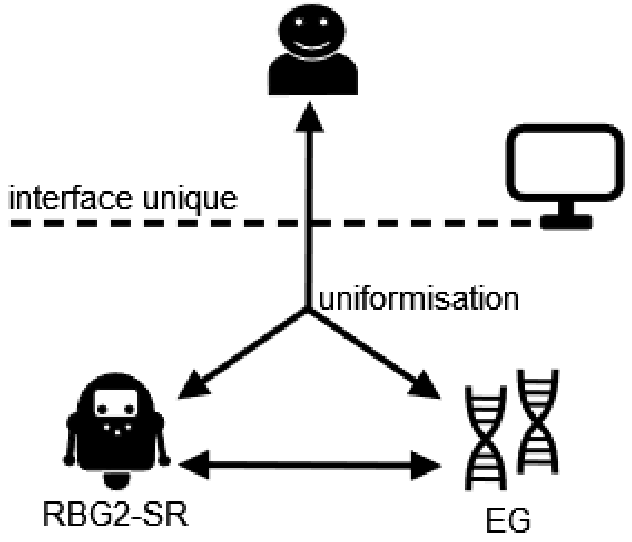
- Error reduced by an order of magnitude with QR1 interactivity: it is possible to reduce the error by preferences

- Optimal interaction frequency of 5 QR2: it is possible to solicit the user only at certain time steps

Fréquence d'interaction	EQM moyenne ($\pm$ écart-type)	R2 moyenne ($\pm$ écart-type)
Référence	$1{,}62 \times 10^{-2}$ ($\pm 1{,}6 \times 10^{-2}$)	$9{,}92 \times 10^{-1}$ ($\pm 8{,}2 \times 10^{-3}$)
1	$6{,}57 \times 10^{-2}$ ($\pm 8{,}91 \times 10^{-2}$)	$9{,}67 \times 10^{-1}$ ($\pm 4{,}45 \times 10^{-2}$)
2	$6{,}14 \times 10^{-2}$ ($\pm 2{,}09 \times 10^{-2}$)	$9{,}69 \times 10^{-1}$ ($\pm 1{,}04 \times 10^{-1}$)
5	$\mathbf{7{,}21 \times 10^{-3}}$ ($\pm \mathbf{8{,}93 \times 10^{-2}}$)	$\mathbf{9{,}96 \times 10^{-1}}$ ($\pm \mathbf{4{,}46 \times 10^{-3}}$)
10	$1{,}14 \times 10^{-2}$ ($\pm 1{,}79 \times 10^{-2}$)	$9{,}94 \times 10^{-1}$ ($\pm 8{,}92 \times 10^{-3}$)
15	$8{,}56 \times 10^{-3}$ ($\pm 1{,}03 \times 10^{-2}$)	$\mathbf{9{,}96 \times 10^{-1}}$ ($\pm \mathbf{5{,}12 \times 10^{-3}}$)
20	$1{,}13 \times 10^{-2}$ ($\pm 1{,}71 \times 10^{-2}$)	$9{,}94 \times 10^{-1}$ ($\pm 8{,}57 \times 10^{-3}$)

Comparaison des scores d'EQM et R2 avec et sans (baseline) interactivité sur les données de la fonction f_{nguyen_4} (10 apprentissages)

Interactivity, RBG2-SR and EG



Evaluation of function f_{nguyen_4}

Algorithme	Fréquence d'interaction	EQM moyenne (± écart-type)	R2 moyenne (± écart-type)
RBG2-SR	-	1.62×10^{-2} $(\pm 1.6 \times 10^{-2})$	9.92×10^{-1} $(\pm 8.2 \times 10^{-3})$
EG	-	1.48×10^{-2} $(\pm 1.6 \times 10^{-2})$	9.93×10^{-1} $(\pm 6.5 \times 10^{-3})$
iRBG2-SR	**5**	$\mathbf{7.21 \times 10^{-3}}$ $(\pm \mathbf{8.93 \times 10^{-3}})$	$\mathbf{9.96 \times 10^{-1}}$ $(\pm \mathbf{4.46 \times 10^{-3}})$
	15	8.56×10^{-3} $(\pm 1.03 \times 10^{-2})$	$\mathbf{9.96 \times 10^{-1}}$ $(\pm \mathbf{5.12 \times 10^{-3}})$
RBG2-SR + EG	-	1.24×10^{-2} $(\pm 1.56 \times 10^{-2})$	9.94×10^{-1} $(\pm 7.79 \times 10^{-3})$
i(RBG2-SR + EG)	5	1.98×10^{-2} $(\pm 4.75 \times 10^{-2})$	9.90×10^{-1} $(\pm 2.37 \times 10^{-2})$
	15	1.11×10^{-2} $(\pm 1.47 \times 10^{-2})$	9.94×10^{-1} $(\pm 6.98 \times 10^{-3})$
	25	1.12×10^{-2} $(\pm 1.78 \times 10^{-2})$	9.94×10^{-1} $(\pm 8.90 \times 10^{-3})$

Comparaison scores d'EQM et R2 (10 learning)

Mundhenk, T., Landajuela, M., Glatt, R., Santiago, C. P., Petersen, B. K. et al. (2021). Symbolic Regression via Deep Reinforcement Learning Enhanced Genetic Programming Seeding. Advances in Neural Information Processing Systems, 34:24912–24923.

Assessment of this experience

Improved search in grammatical space through interactivity

- Consideration of implicit knowledge
- Decrease in error compared to the RBG2-SR approach
- Reduction in the number of user requests

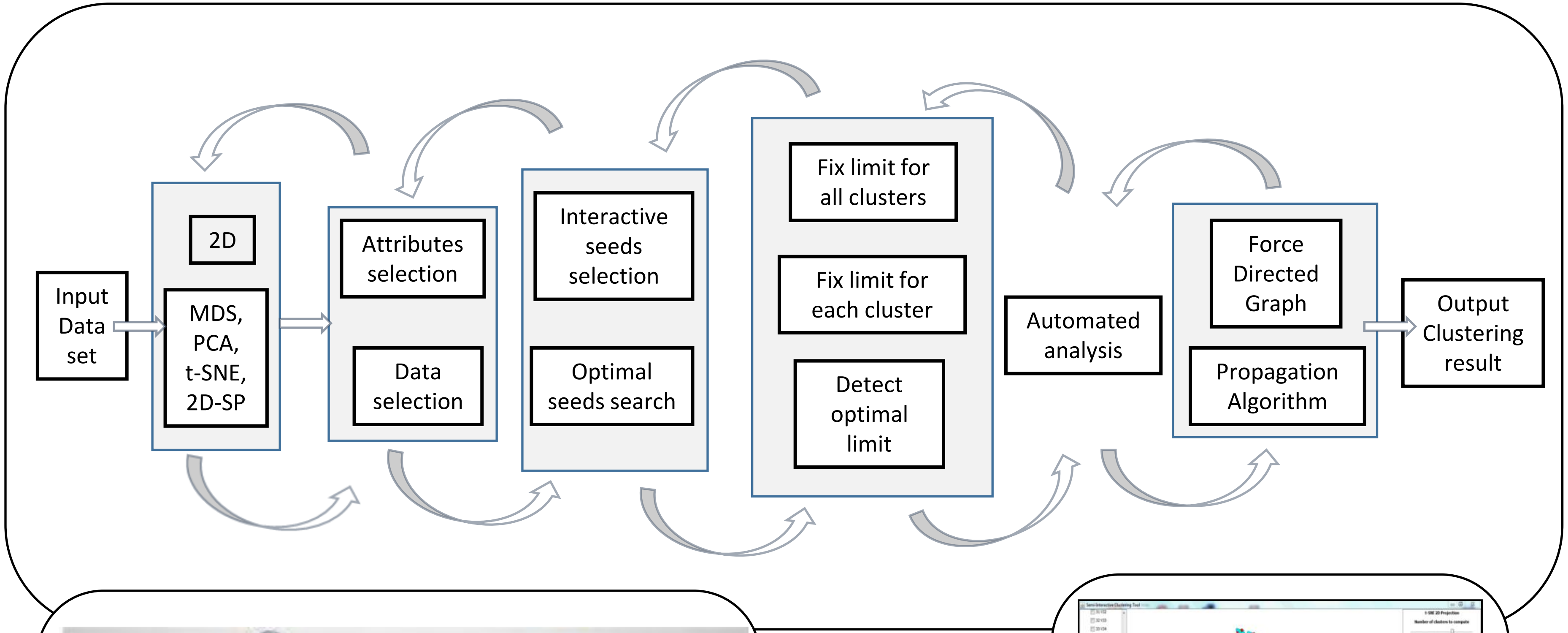
Preliminary results on a single function

- To be replicated on new data
  - Other Functions
  - Power grid data
  - Scaling up (larger network, real data, ...)
- Test new user behaviors
- Evaluate on a larger panel of users

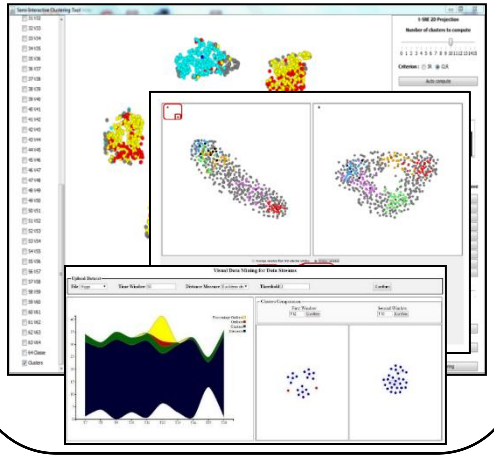
5

Conclusion and Challenges

Conclusion



Input Data set
2D
MDS, PCA, t-SNE, 2D-SP
Attributes selection
Data selection
Interactive seeds selection
Optimal seeds search
Fix limit for all clusters
Fix limit for each cluster
Detect optimal limit
Automated analysis
Force Directed Graph
Propagation Algorithm
Output Clustering result



Answers

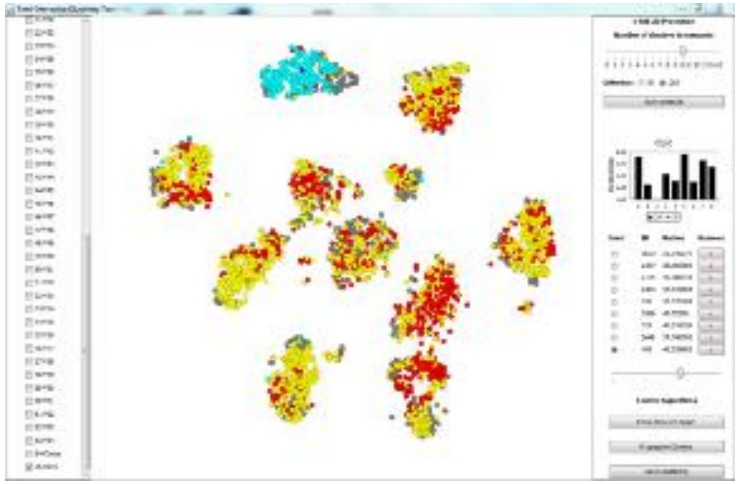
How to integrate the user constraints in the process?

Visual Data Mining

- Evaluation update through the clusters’ modification

Interactive Machine Learning

- Grammar evaluation and updating



How to better guide the user in the process ?

Visual analytics

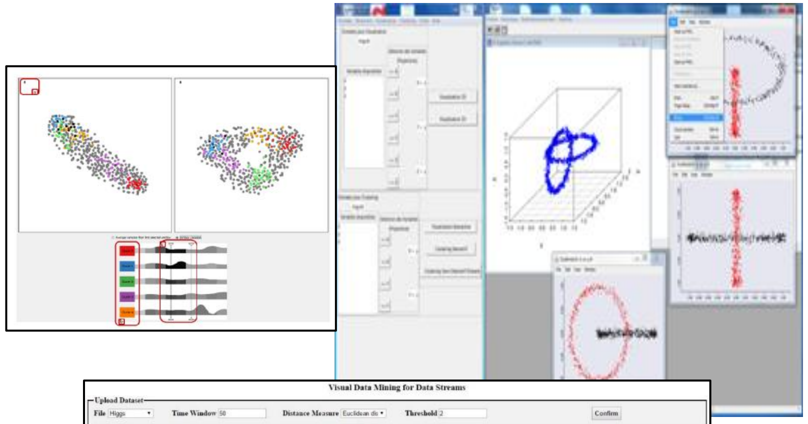
- Data stream visualization:
  - Change detection
  - Selected window exploration
  - Optimal subspaces to describe the data stream
- Select subspaces and visualize the associated themeriver

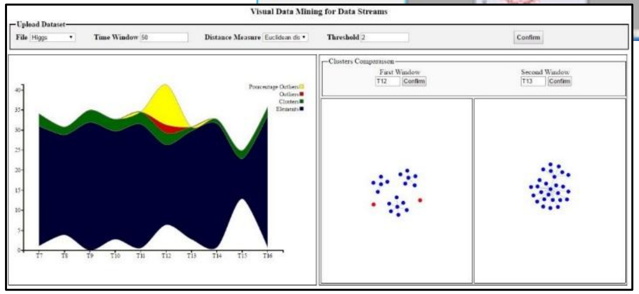
Visual Data Mining

- Evaluation update a each cluster modification
- Evaluation criterion of each extracted cluster

Interactive Machine Learning

- Extract interpretable explanation of a network constraint





How to avoid the visualization bias?

Cooperative approaches: automatic, visual & interactives

Challenges

Track 1: How to integrate the user in the process: exploration & processing

Challenge 1

Evaluation in visual data mining

- How to detect artefacts /How to evaluate the visualizations : subspace projections

Challenge 2

Interactive machine learning

- User interactions learning (IA) : model the user behavior
- How avoid the visual perception bias

Challenge 3

Semi-interactive approaches

- Cooperatives approaches automatics & visualizations: explanations, interpretability

Track 2: Massive and temporal data

Challenge 1

Visual data mining

- How to use all semi-interactive approaches advantages

Challenge 2

Data and associated difficulties

- times series, spatio-temporal data, symbolic data, complex data
- Different data: images, Graphs, omics, bioscience