FACIL: Incremental Rule Learning from Data Streams

Francisco J. Ferrer Troyano

Jesús S. Aguilar, José C. Riquelme

University of Sevilla

ferrer@lsi.us.es

Contents

- Motivation
- FACIL
  - The hybrid knowledge model
  - Three user parameters
  - IIL updating
  - The moderate *growth d*istance
  - Implicit and explicit forgetting heuristics
  - The multi-strategy classification method
- Experiments
- Future work

Motivation

Why Classify?

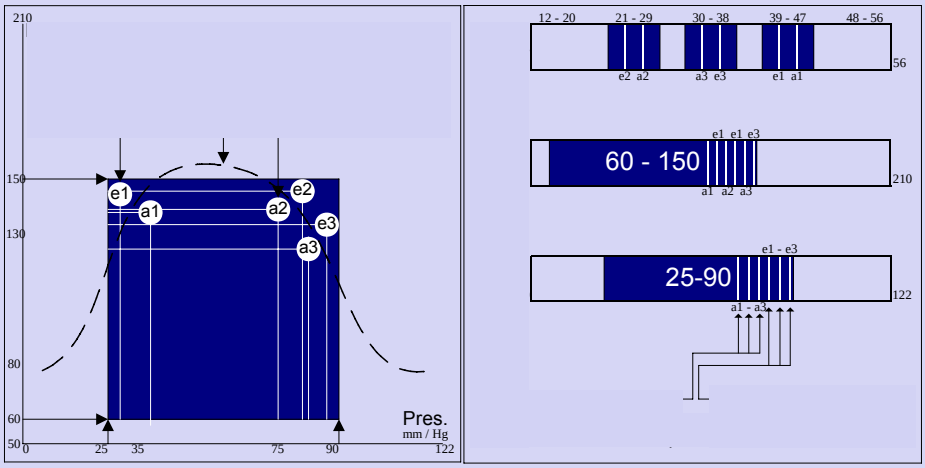
Why Rules?

Heterogeneous Data Sources → Noise, Missing Values, Inconsistency

High-Speed Data→ Fast algorithms, approximate answers

Open-Ended Data → On-line learning… complex models?

FACIL – The hybrid knowledge model



Border examples inside rules

FACIL – The Algorithm

Three user parameters

- Maximum number of rules / label
  → necessary to bound the computational complexity
- Minimum accuracy (purity) / rule: p/(p+n)
  → anytime,approximate answers
- Maximum growth / dimension

FACIL – IIL updating

General IL Approach

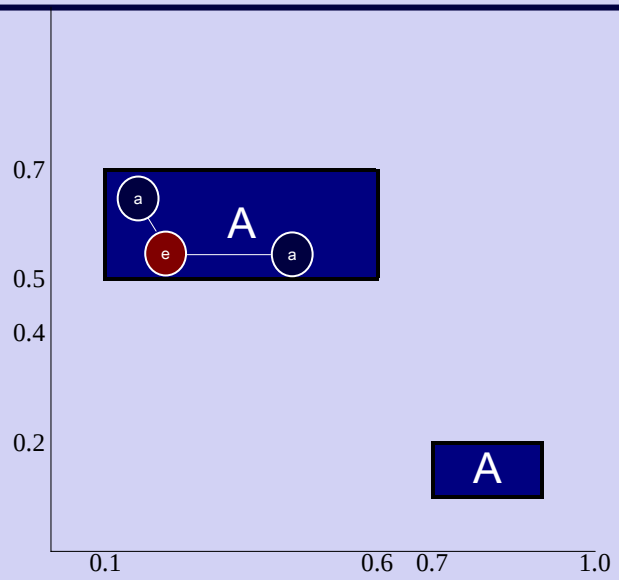
Successive Episodes → Updating: $M_t = f(V_t, M_{t-1})$

Window of examples → V_t: New **[and]** past examples

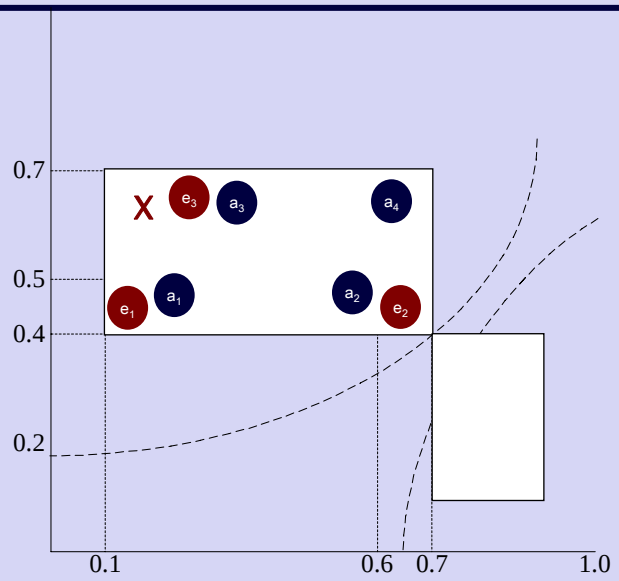
IIL: Every example → 1/3 cases:

1. Success
2. Anomaly
3. Novelty

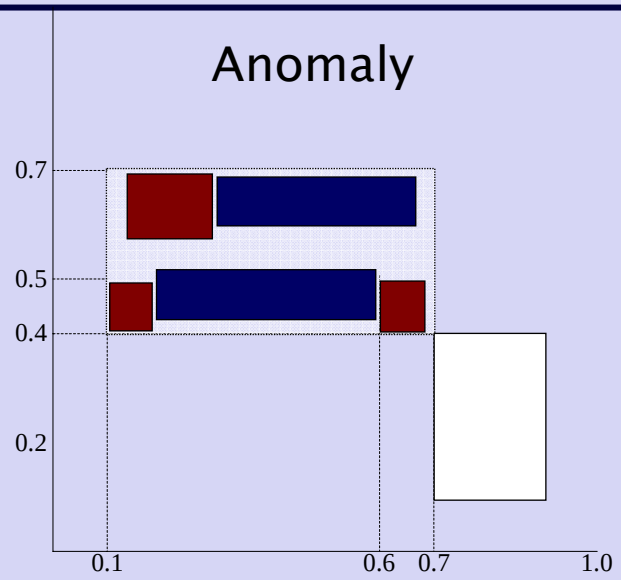
FACIL – IIL / Success



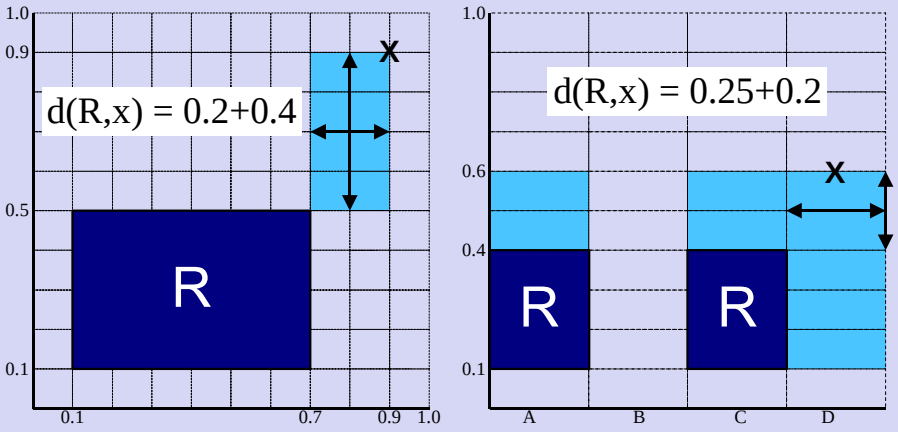
FACIL – IIL / Anomaly



FACIL – Learning Approach / Case 3



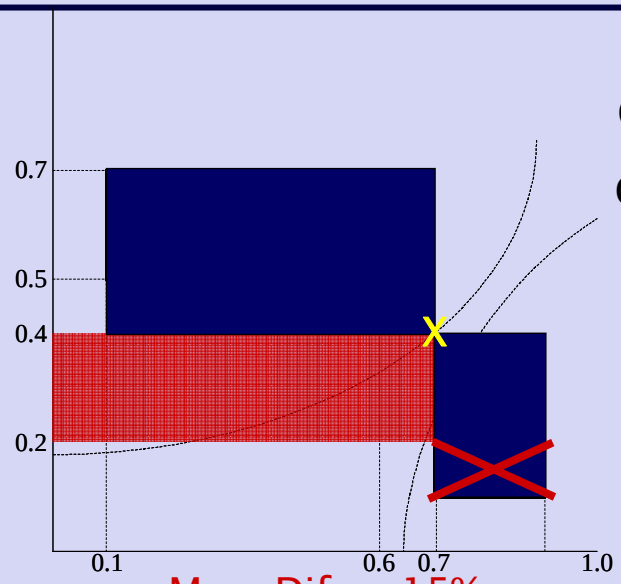
FACIL – IIL / Novelty: The *Growth* Distance



Selection:

The rule implying the minimum growth

FACIL – IIL / Novelty: The *Moderate Growth* Distance

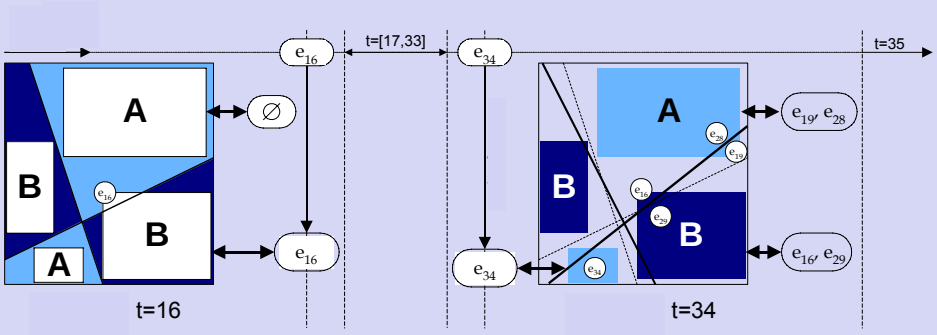


$G(R_a,x) = 0.2$

$G(R_b,x) = 0.2$

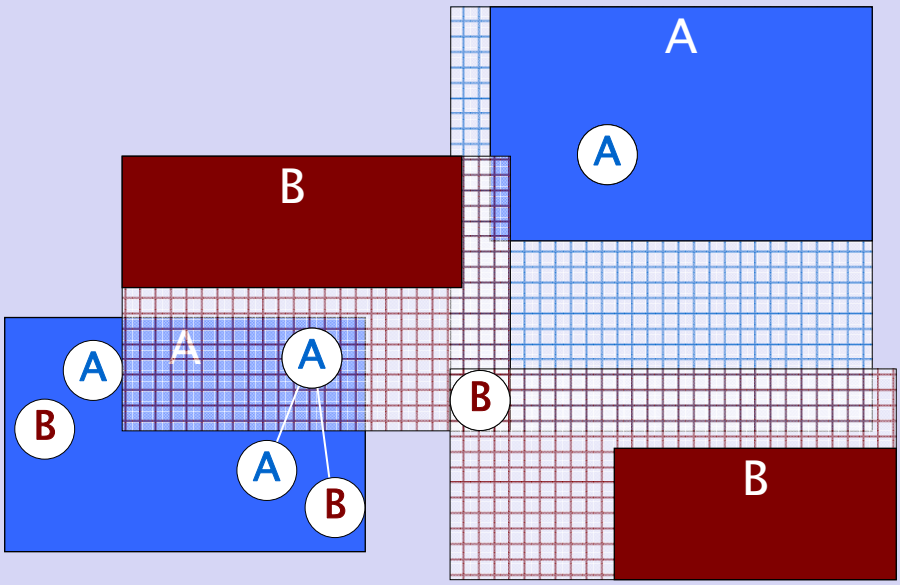
Max. Dif. = 15%

FACIL – Implicit and Explicit Forgetting Heuristics



multiple windows

FACIL – Multi-strategy Classification



Experiments – General Purpose Classifier Evaluation

FACIL as a UCI Repository

10 Folds Cross Validation, 10 times, t-student (0.05)

Implicit Forgetting Heuristic

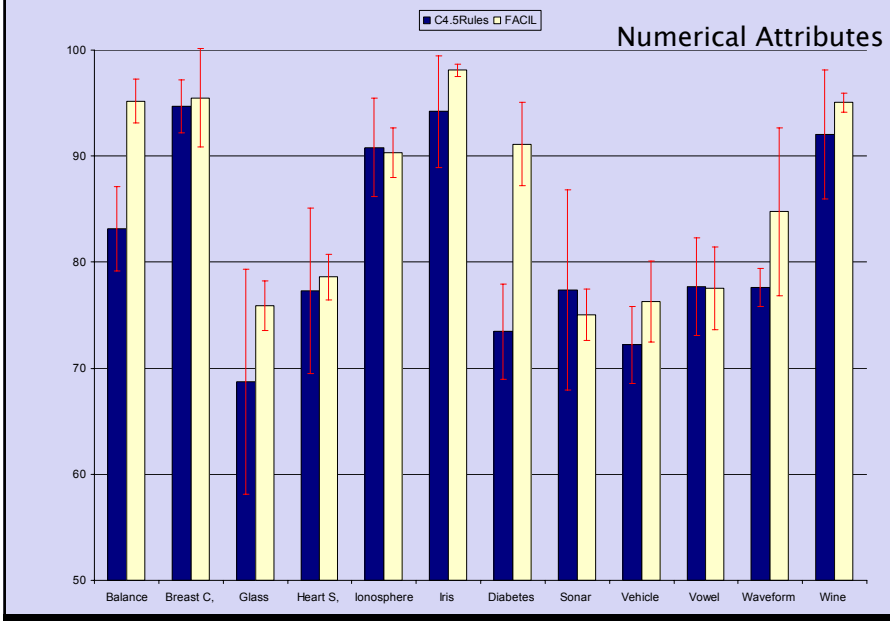
Synthetic Data Streams

Change Magnitude = ±0.1 for 40% attributes every 10^4 examples

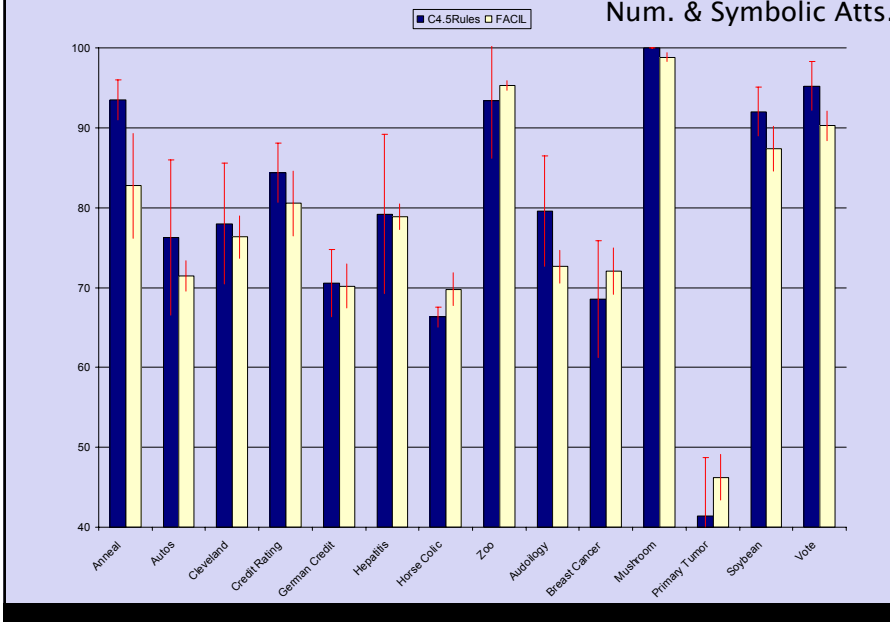
INPUT: 10^5 examples - 5% class noise: 900 training / 100 testing

Both Implicit and Explicit Forgetting Heuristics

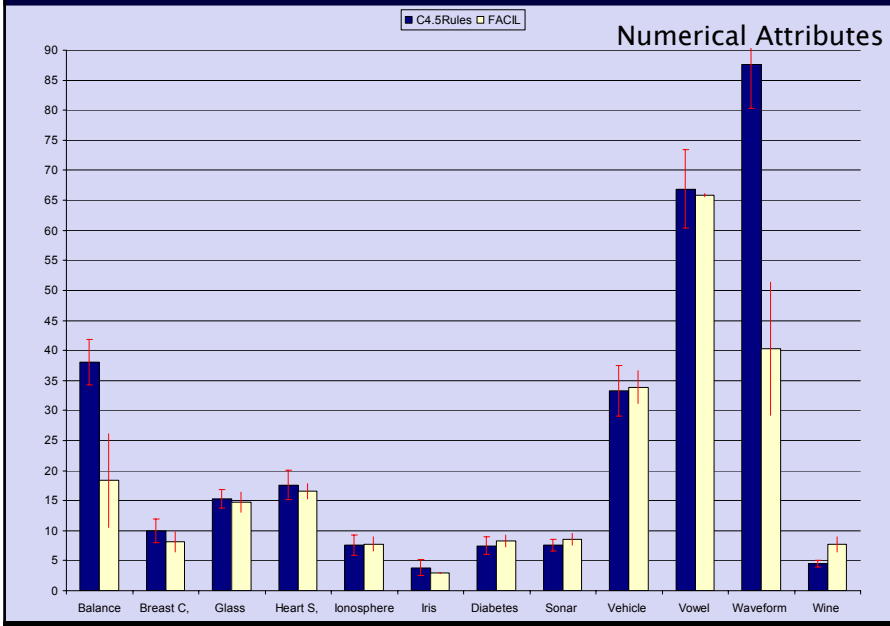
UCI Repository – Accuracy



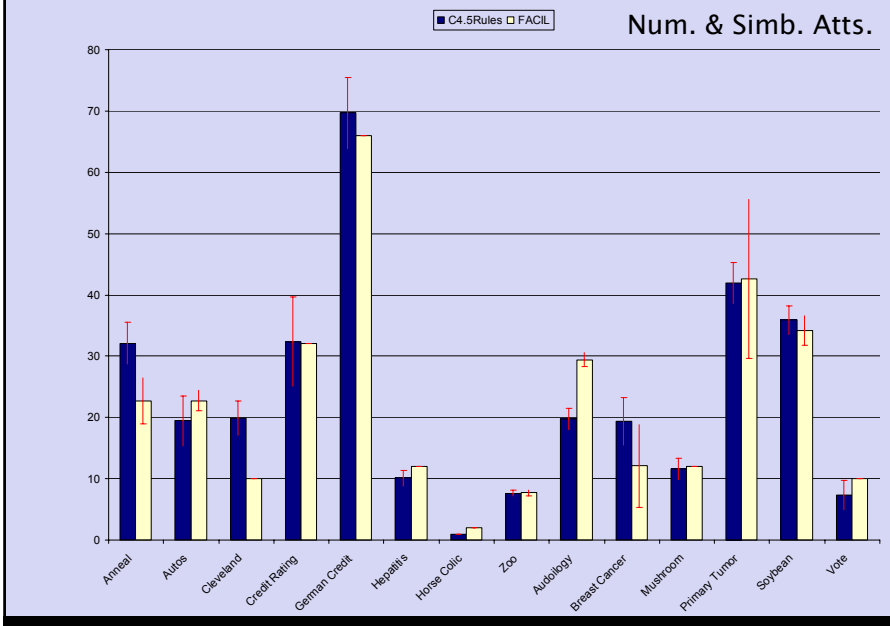
UCI Repository – Accuracy



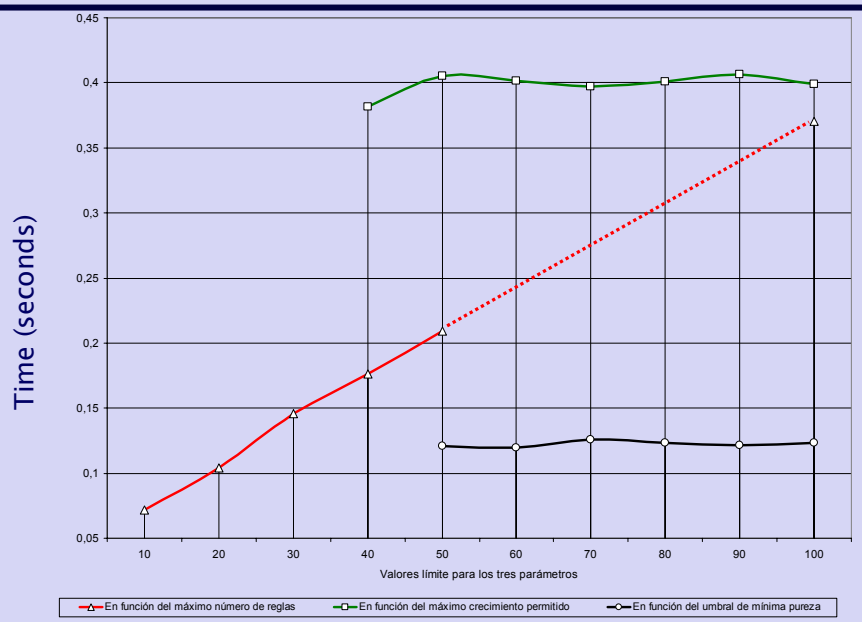
UCI Repository – Number of rules



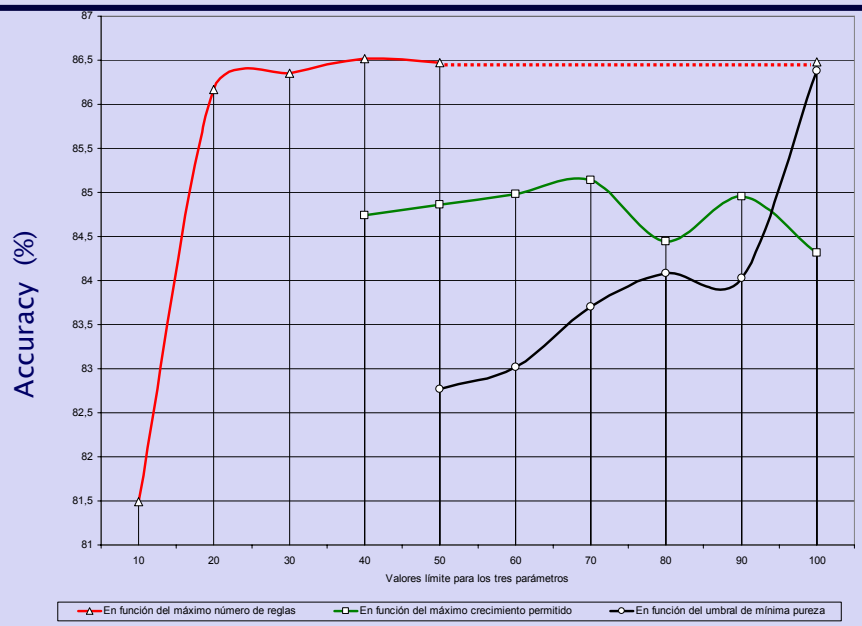
UCI Repository – Number of rules



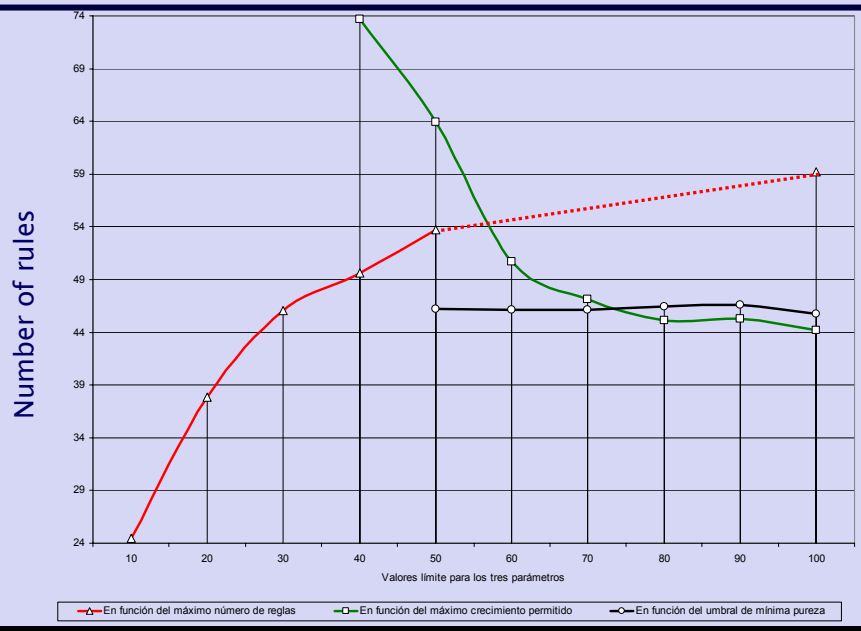
UCI Repository – Parametric Sensitivity



UCI Repository – Parametric Sensitivity



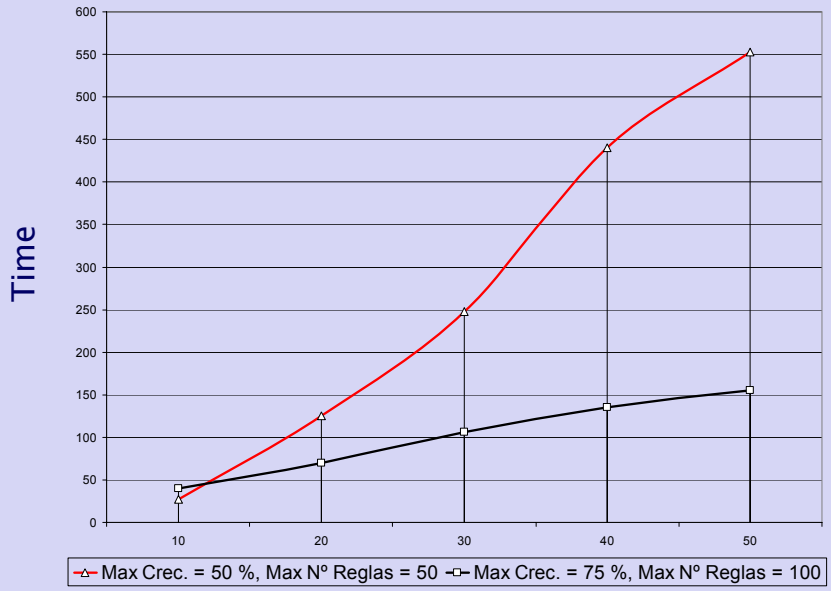
UCI Repository – Parametric Sensitivity



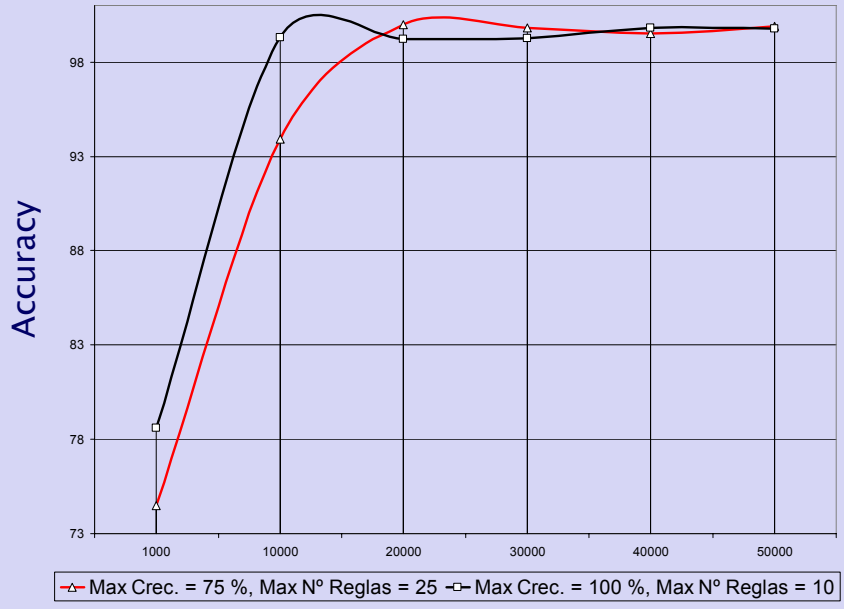
Moving Hyperplane – Accuracy, time, and number of rules

	Máx. Growth = 50%			Máx. Growth = 75%		
Nº Attributes	Accuracy	Time	Nº Rules ≤ 25	Accuracy	Time	Nº Rules ≤ 25
10	➔ >3500 examples / second			➔ >2500 examples / second		
20	93,25	63	9	94.57	35	12
30	91,04	124	3	89.26	53	10
40	89,7	221	2	89.19	67	7
50	➔ >180 examples / second			➔ >650 examples / second		

Moving Hyperplane – Sensitivity to the number of attributes



Moving Hyperplane – Sensitivity to the number of examples



Future Work

Sensitivity to attribute

- Removing / Recovering attributes
- To adapt the Growth distance:
  - Changing attribute weights → improve the selection of rules
  - Reordering attributes according influence → speed-up the updating process

Processing streams with a variable number of attributes

- To evaluate alternative distances between a rule and an example

PreProcessing streams? → Multiple Classification

GRAZIE MILLE

Francisco Ferrer

 LSI - University of Seville 

ferrer@lsi.us.es