

Exploring antimicrobial resistance prediction using post-hoc interpretable methods

B. Cánovas-Segura¹, A. Morales¹, A. López Martínez-Carrasco¹,
M. Campos¹, **J. M. Juarez**¹, L. López Rodríguez², F. Palacios¹

¹ AIKE research group, University of Murcia

² Hospital University of Getafe

UNIVERSIDAD DE
MURCIA

Jun 29 TEAAM workshop, Poznan



Funded by Spanish Ministry of Science, Innovation and Universities under the SITSUS project (Ref: RTI2018-094832-B-I00), and by the European Fund for Regional Development (EFRD, FEDER).

- **MEDICAL PROBLEM**

- **Reduce** spread of resistant bacteria in hospitals
- AST: Antimicrobial suscep. test

- Effect of AB on bacteria strains
- Days to be completed



- **Predict** if infected by resistant bacteria before AST
- **Focus on VRE:** enterococcus species resistant to vancomycin

Enterococci / ,ɛntərəʊ'kɒkəi/

- **OUTLINE:**

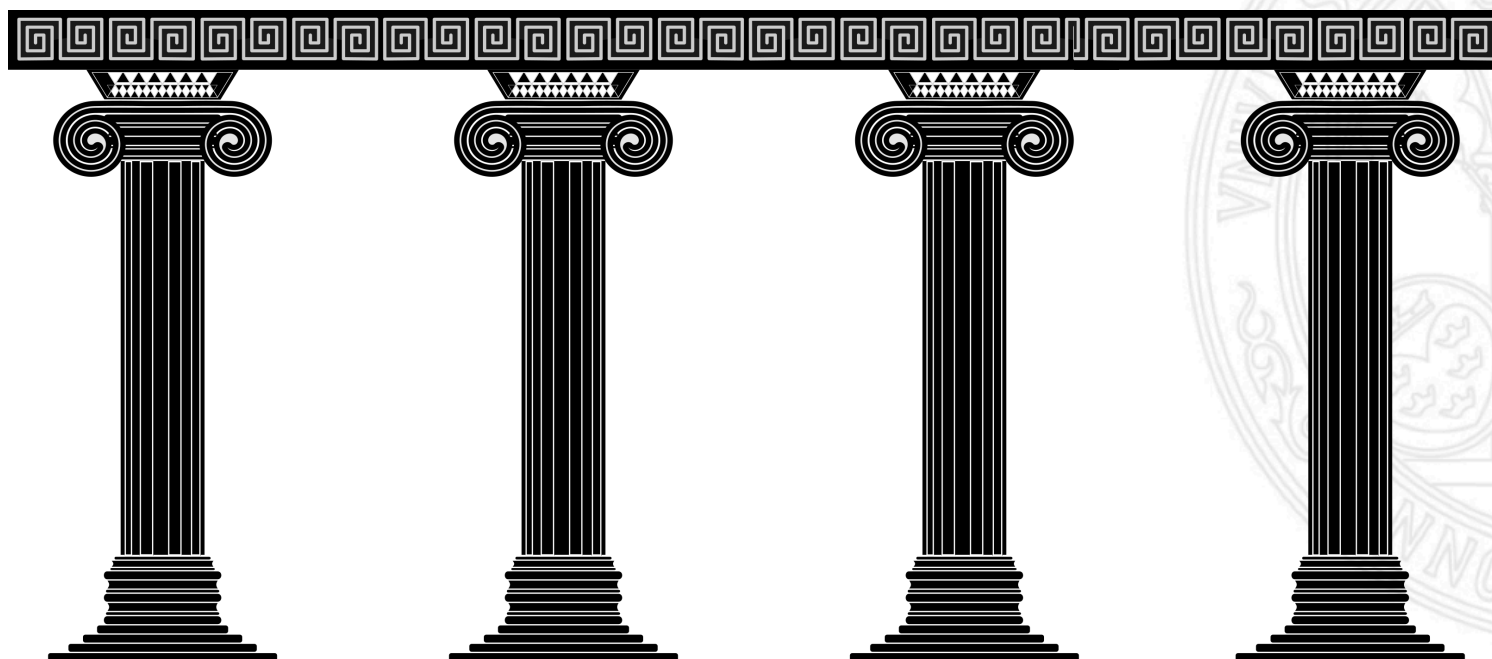
1. Medical problem✓
2. Proposal
3. Requirements & experiments
4. Conclusions & future work



- **PROPOSAL**

- Assess the impact of **ML models**
- For **VRE predictor** in a **CDSS**

User & Data REQUIREMENTS



Interpretability Concept drifting High dim. Imbalance classes

- **Requirements: INTREPRETABLE**
 - Vague term
 - Lipton's classification:
 - Transparent
 - **Post-Hoc explanation**
 - Trust and easy evaluation
 - **Approach:**
 - Logistic Regression
 - Decision Trees
 - Rule-based

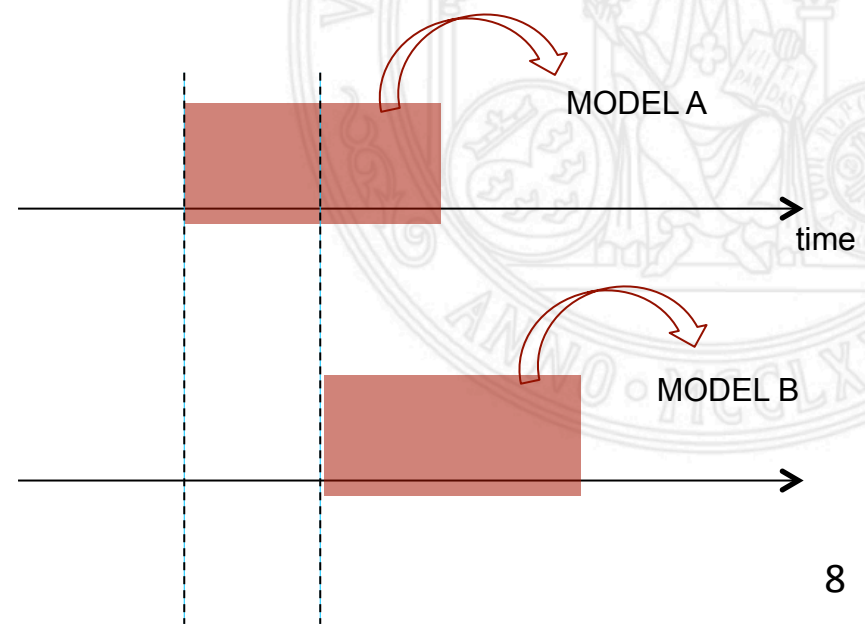


- **Requirements: Concept Drift**

- Influence of factors vary over time
- Bacteria develop resistance
- Seasonal variations dissemination bacteria
- Change clinical protocols

- Approach:

- Sliding windows
- Understand & Implement



- **Requirements: High Dimensionality**
 - Patient clinical record + Lab test results
 - Approach:
 - Filters (FCBF)
- **Requirements: Imbalanced Dataset**
 - Few observations Resistant class
 - Many Susceptible class
 - Approach:
 - Random oversampling (1:1)



Principal researcher

Other principal researchers



Bernardo Cánovas-Segura
PhD,
Univ. Murcia

The Condensed Nearest Neighbor Rule

The purpose of this rule is to introduce the condensed nearest neighbor decision rule (CNN rule) and to pose some unsolved theoretical questions which it raises. The CNN rule, one of a class of *ad hoc* decision rules which have appeared in the literature in the past few years, was motivated by statistical considerations

Manuscript received November 21, 1986; revised October 5, 1987.



Figure 2: A group of researchers, including Bernardo Cánovas-Segura, posing in a laboratory setting. A large blue arrow points from the 'Other principal researchers' header to the group photo.

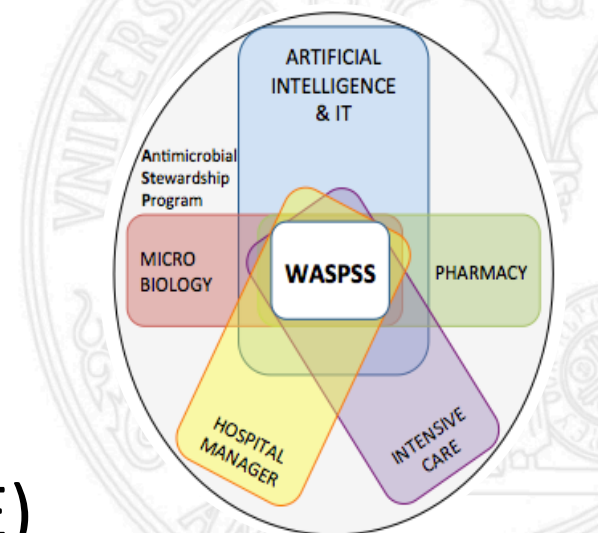
- **EXPERIMENTS**

- VRE:

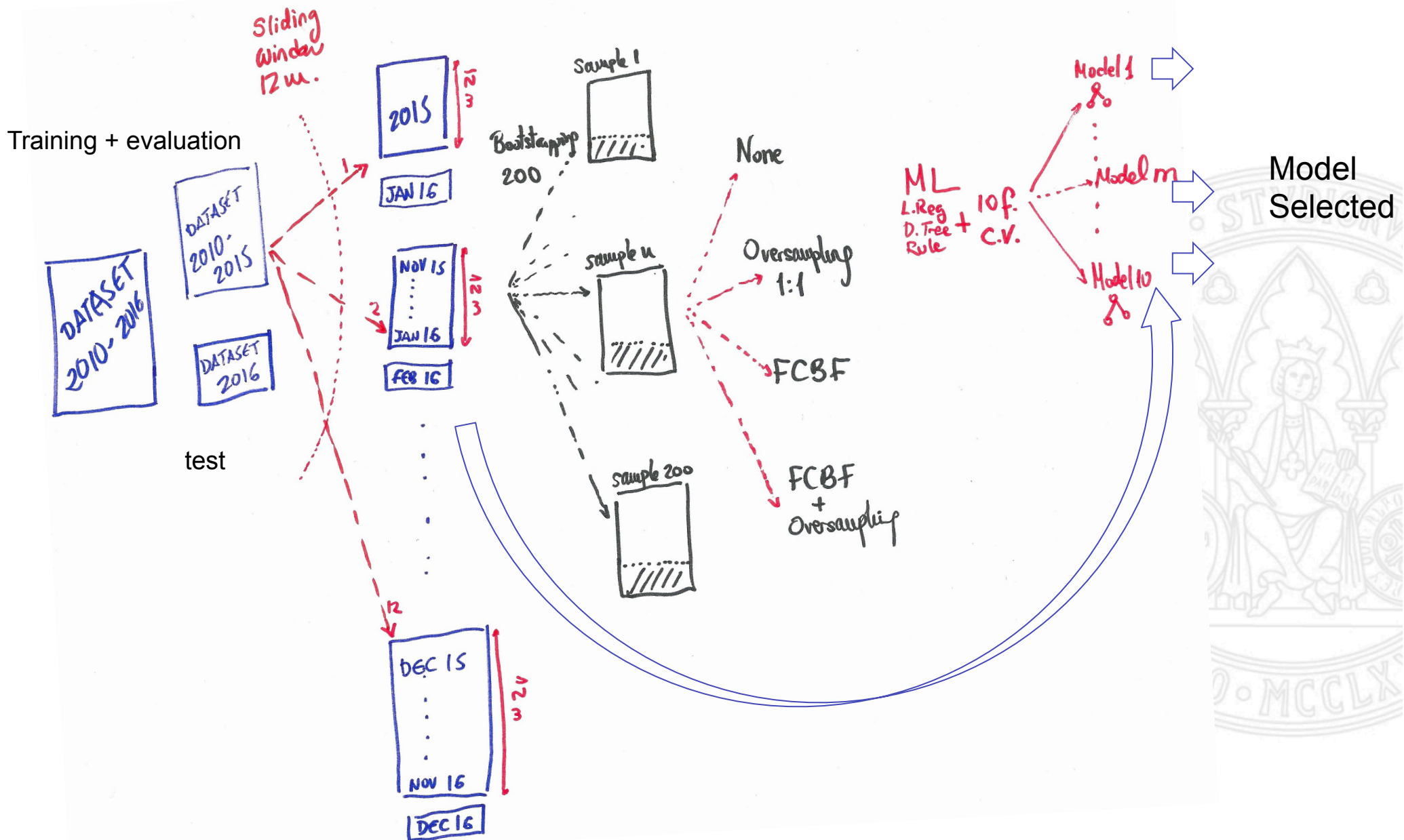
- Bacteria: Enterococci faecium + Enterococci faecalis
 - Antibiotic: Vancomycin

- Dataset:

- WASPSS platform
 - University Hospital Getafe
 - 2010-2016
 - 1393 + Enterococci cultures (42 VRE)
 - 571 attributes + binary class (resistant|suscept.)

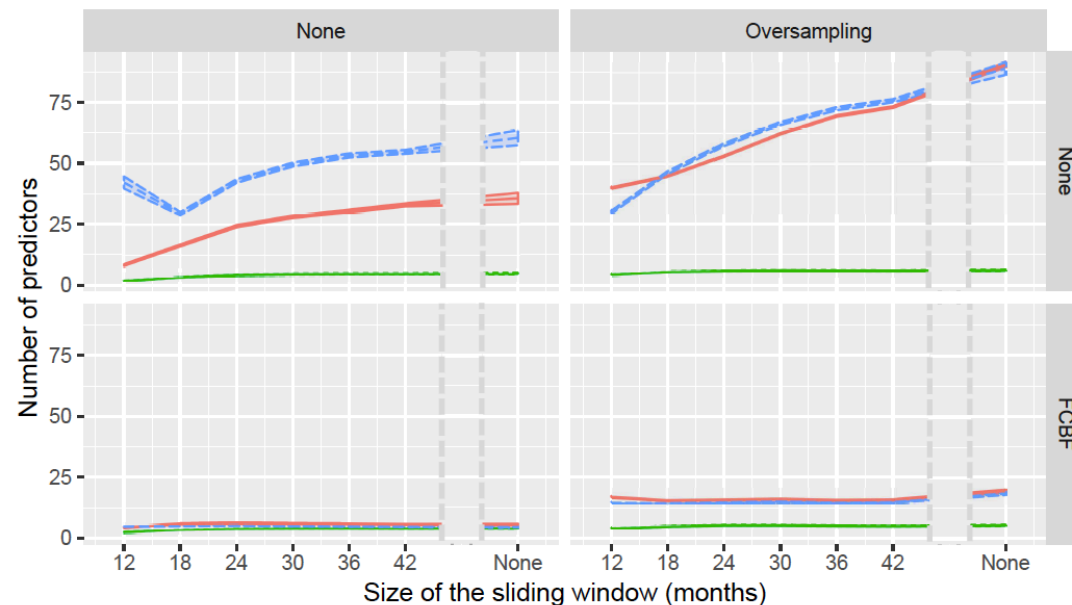


• EXPERIMENTS:



• Results

Mean (95% C.I.) number of predictors per model



Model: — Logistic regression - - - Conditional tree - - - Rule-based model

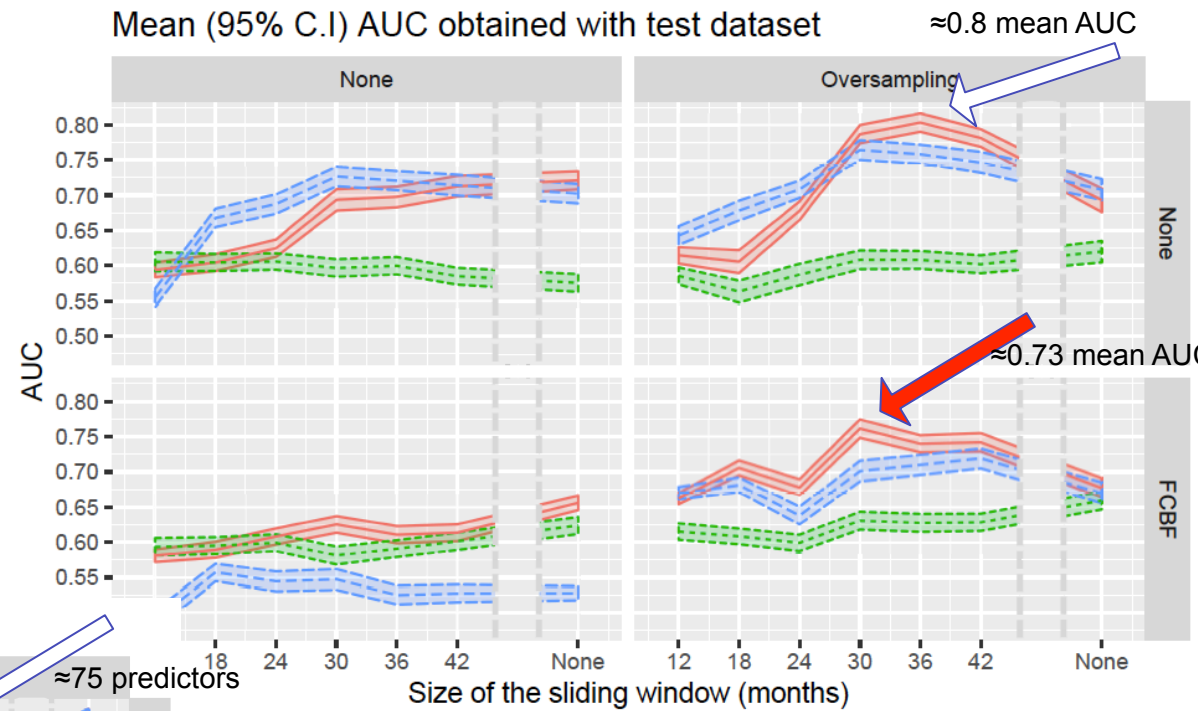
Mean (95% C.I.) AUC obtained with test dataset



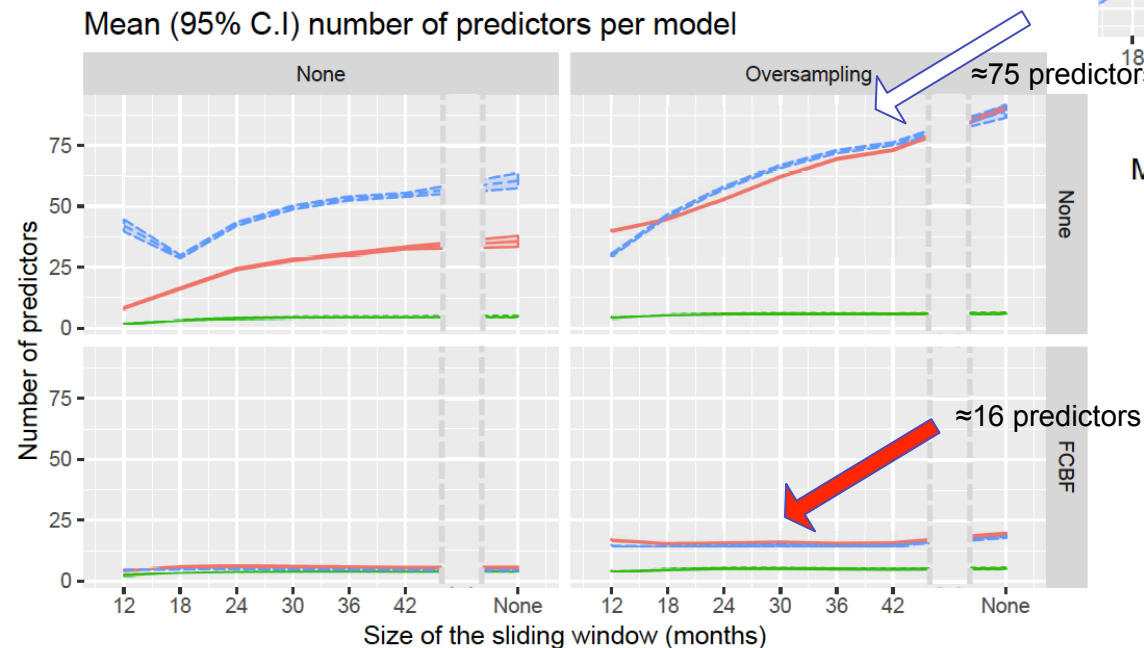
Model: — Logistic regression - - - Conditional tree - - - Rule-based model

• Results

Mean (95% C.I) AUC obtained with test dataset

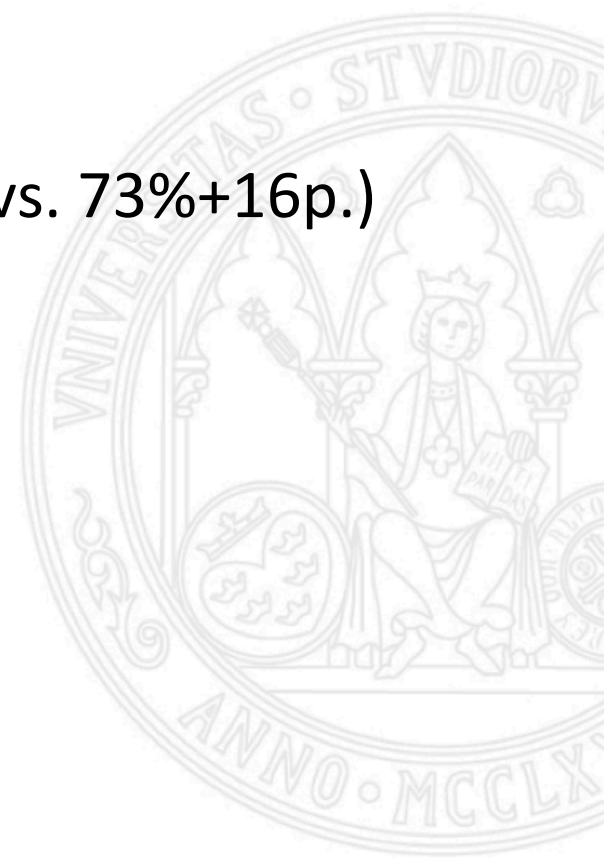


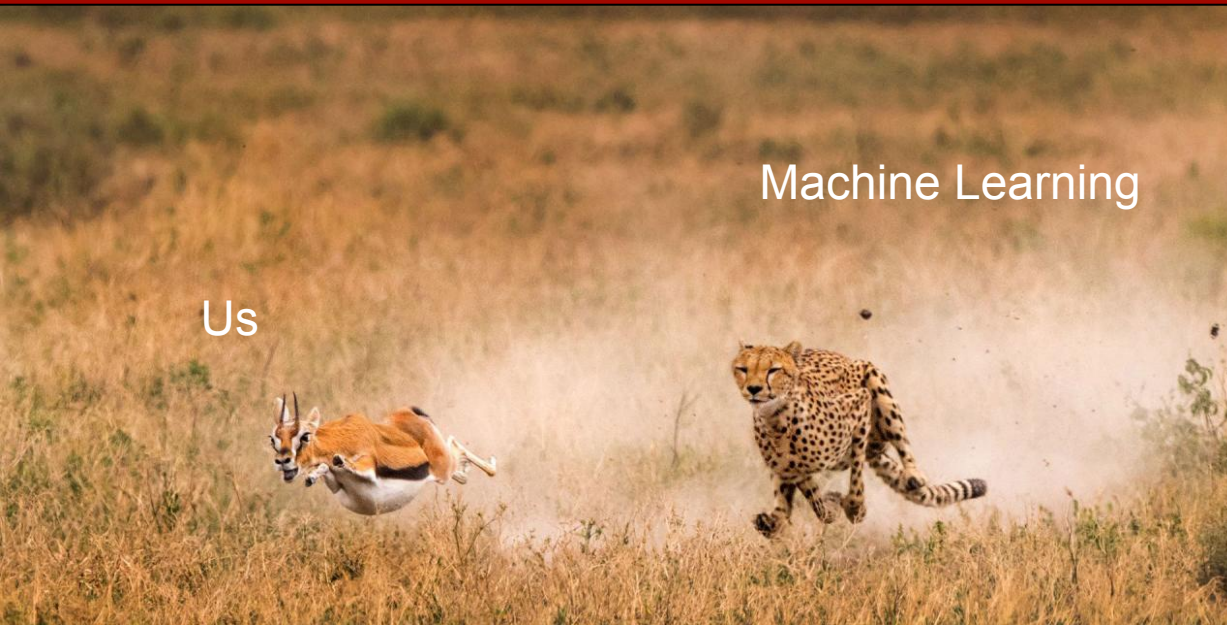
Mean (95% C.I) number of predictors per model



- **Conclusions & future work**

- 3 interpretable models for VRE problem
- Temporal dependence: concept drifting
- ✓Oversampling +FCBF in VRE
- Performance vs. interpretability (80%+75p. vs. 73%+16p.)
- How evaluate interpretability of models
- Models in CDSS and clinical activity flow





HUNTING CHEETAHS...
why not?





Dziękuję za uwagę

(THANK YOU FOR YOUR ATTENTION)

Jose M. Juarez

jmjuarez@um.es

