

Stacking Prediction for a Binary Outcome

C. Gomes, H. Noçairi, M. Thomas, F. Ibanez, J.F. Collin & G. Saporta

- **Introduction**
- **Specific methodology**
 - **Visualization of the methodology**
 - **Process of validation rules**
 - **Parametrization process for each model**
- **Data and Application**
- **Conclusions and Perspectives**

Objective : Prediction of binary outcome



A large number of supervised classification models have been proposed in the Literature

Which One To Choose?



Bias induced by the use of one single statistical approach



Solution "stacking" meta-model.

- **"Stacking" meta-model**

Combining models → Logistic PLS-DA instead of OLS



Strong correlation between predictions



- **Choice of different models**

Boosting, Naïve Bayes, SVM, Sparse PLS-DA, and Expert Scoring

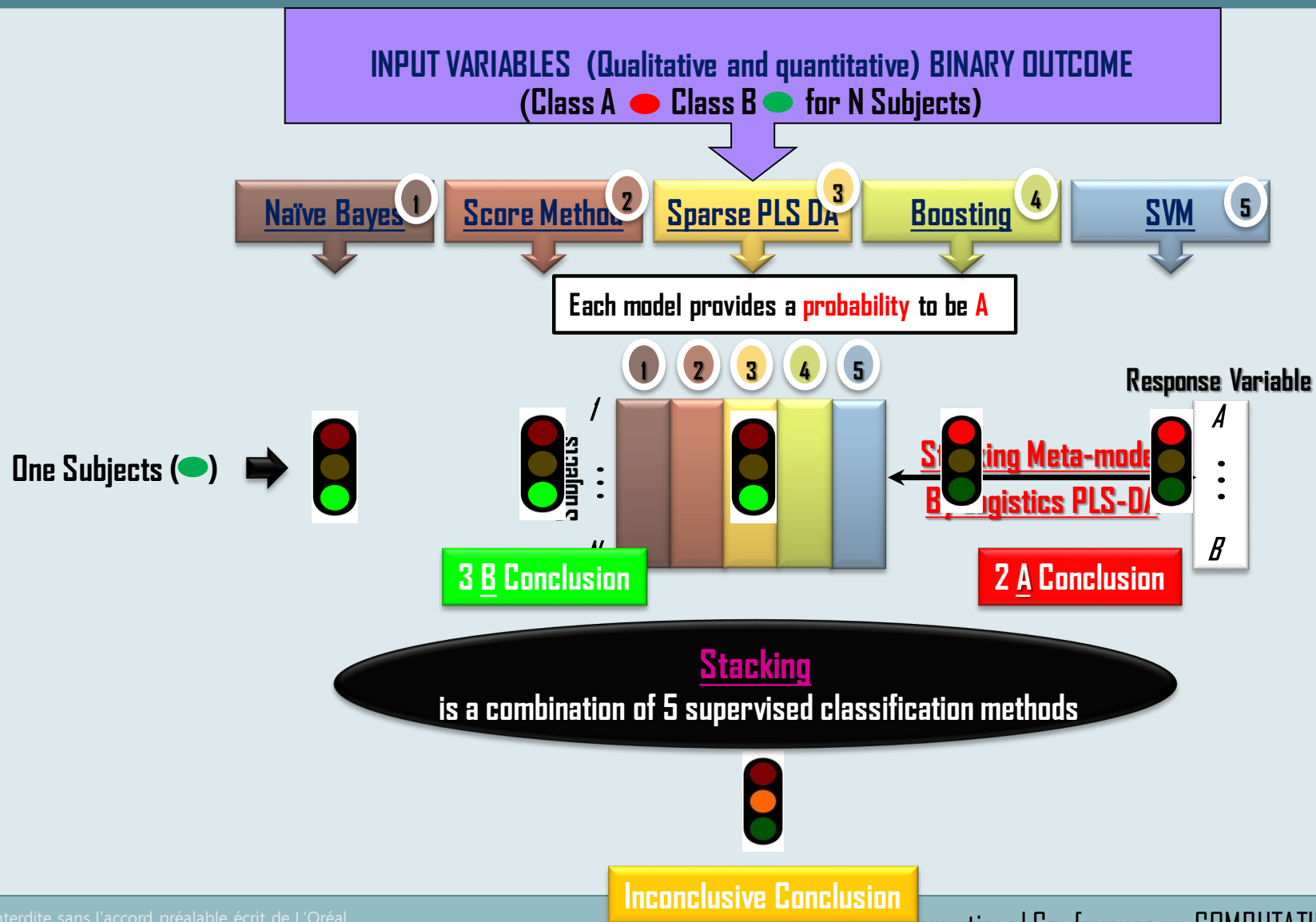


- **Small number of observations**

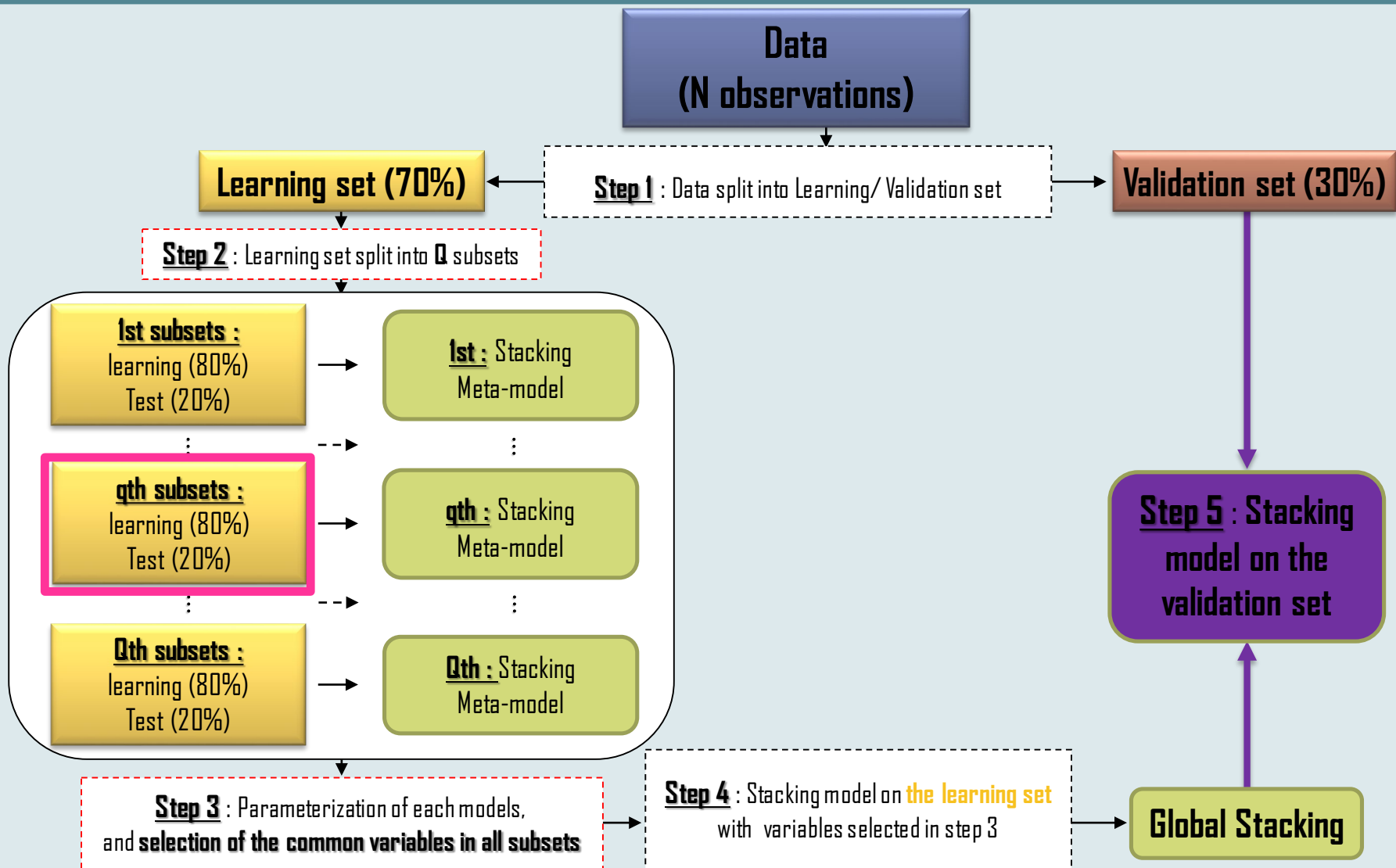
Repeated sub-sampling for variables selection

Visualization of the methodology

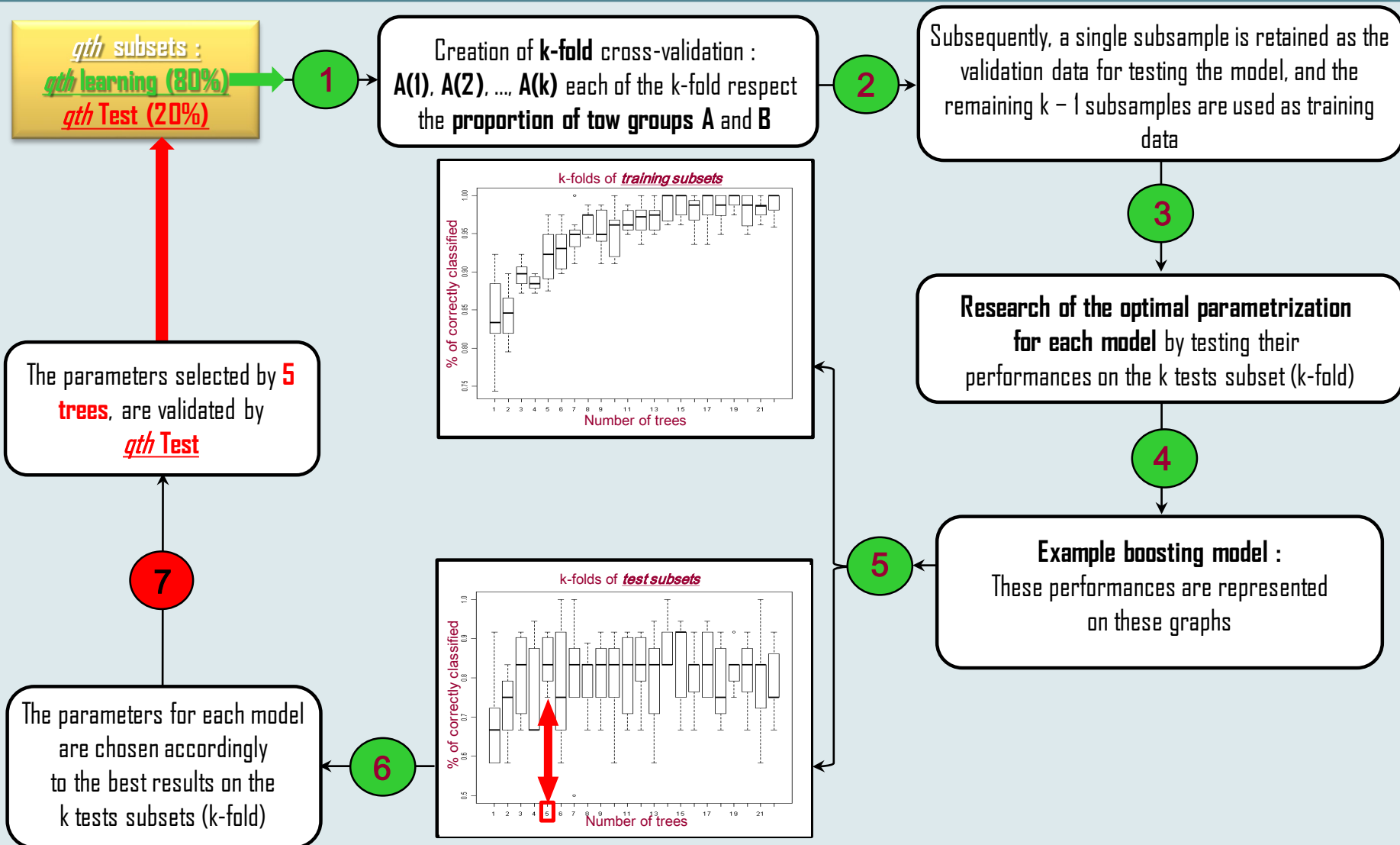
L'ORÉAL



Process of validation rules



Parametrization process for each model



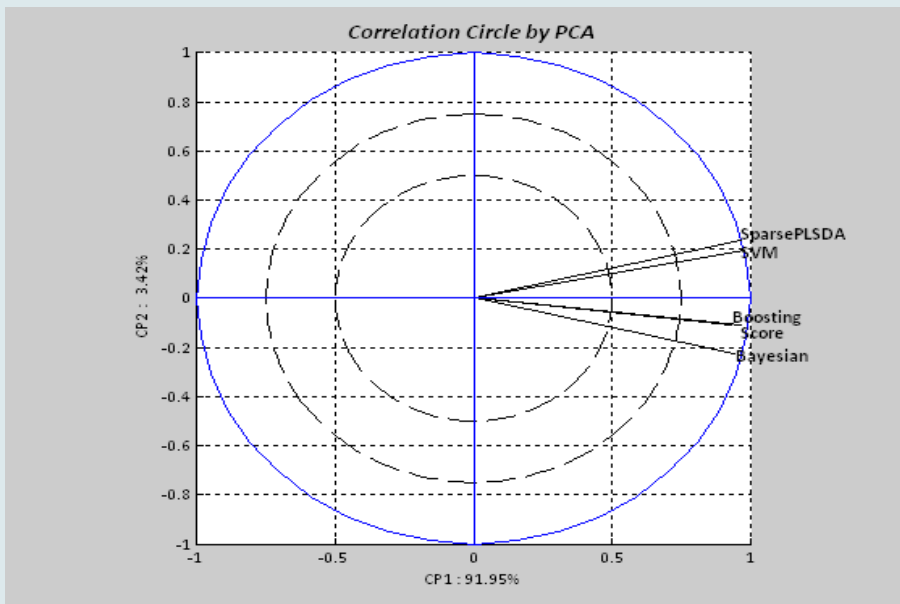
- Comparisons between the five methods and the combined models are done according to two features:
 - global performance with ROC analysis
 - Concordance assessed by Principal component analysis (PCA)
- A decision system with three intervals is finally proposed to the expert, with a no-decision region.

- L'Oreal is developing approaches for safety evaluation (the evaluation of skin sensitization) of ingredients by combining multiple *in vitro* and *in silico* data.

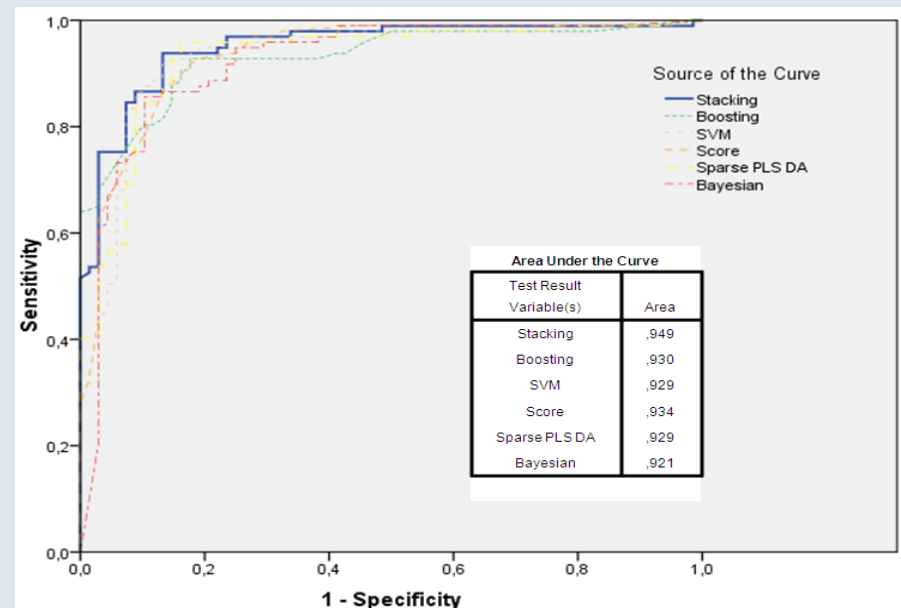
- Data :
 - For this purpose we used a full data set on 165 chemicals composed of 35 different variables, representing
 - the results from *in silico* predictions (Derek, TIMES, Toxtree), from DPRA, MUSST,
 - Nrf-2 and PGE-2 *in vitro* assays as well as numerous physico-chemical experimental or calculated parameters

- In order to predict substances into two groups (**Sensitizer**/**No-Sensitizer**)

Experimental result Context



Predictions provided by the five models are obviously highly positively correlated as shows the following PCA analysis

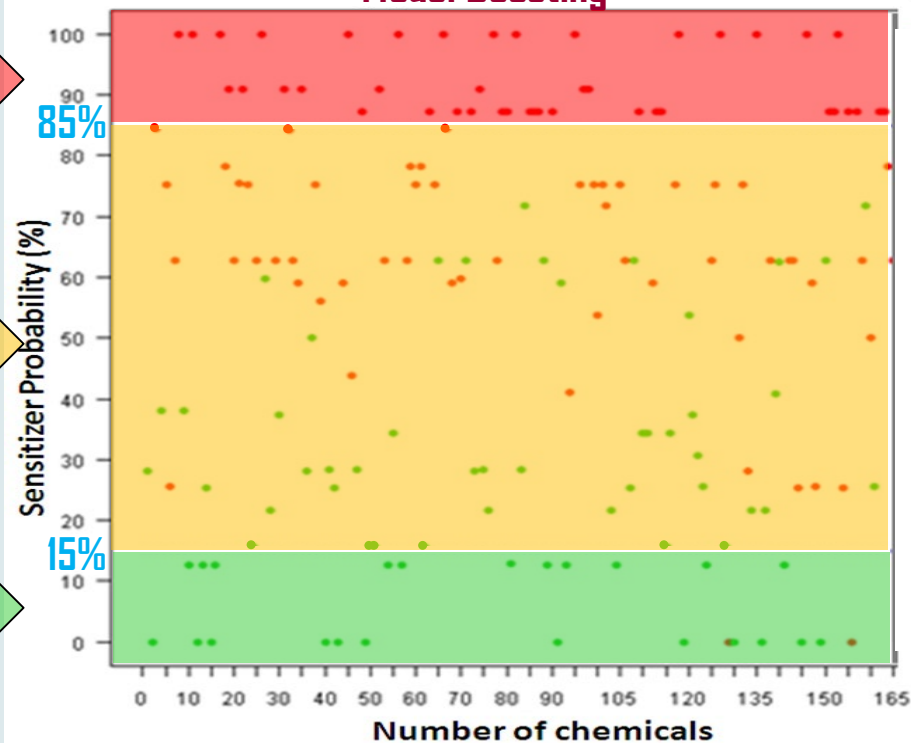


The stacking appears to be the most efficient (blue curve) with the highest area under the curve (0.949).

Confidence area of the boosting model and of the stacking model

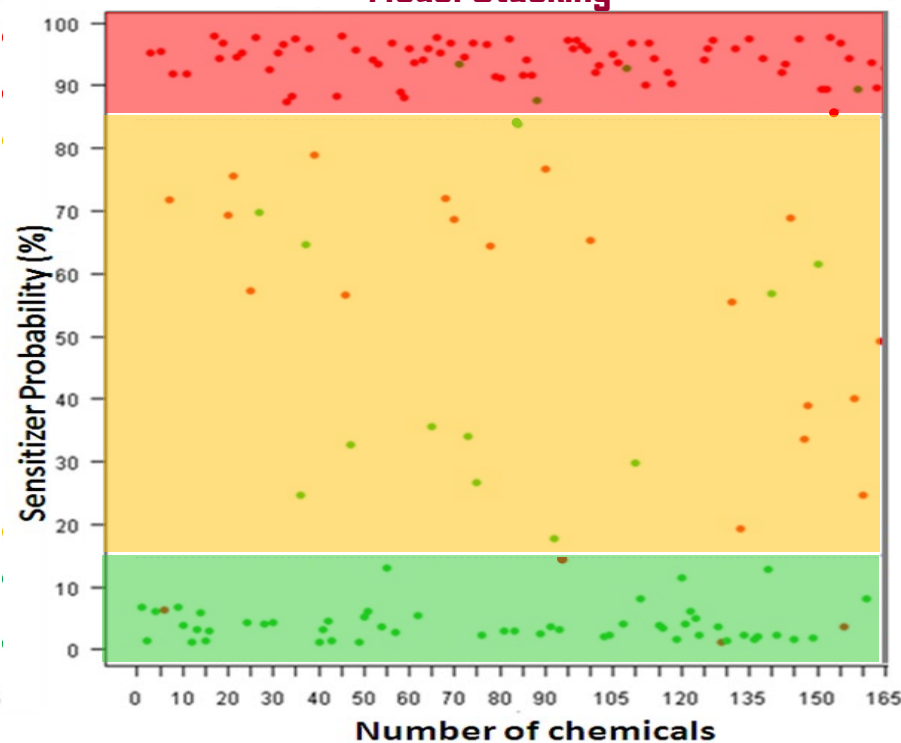
(Sensitizer ● No Sensitizer ●)

Model Boosting



(N=67: $\geq 85\%$ and $\leq 15\%$)

Model Stacking



(N=135: $\geq 85\%$ and $\leq 15\%$)

Sensitizer
Conclusion

Inconclusive
Conclusion

No Sensitizer
Conclusion

85%

15%

Performances on the validation set (N = 50)

- Performance comparisons on a validation set (25 Sensitizer and 25 No Sensitizer) :
 - Take into account only high probabilities ($\geq 85\%$ and $\leq 15\%$) :

Performances of the prediction	Boosting	Score	Sparse PLSDA	SVM	Naïve Bayes	Stacking
Sensitivity	84.61	87.50	84.21	86.36	<u>95.00</u>	<u>91.00</u>
Specificity	<u>100</u>	<u>100</u>	93.75	93.75	89.00	<u>94.44</u>
Concordance	90.00	<u>95.24</u>	88.57	89.47	92.00	<u>92.50</u>
Kappa	79%	<u>89%</u>	77%	78%	84%	<u>85%</u>

Results show that stacking model has better performance than all the other models taken separately on a larger set

■ Conclusions :

- The **Stacking Meta-Model** gives a prediction model with **better performances** for the development of alternative approaches in safety evaluation of chemicals the each of initial five models separately
- The **binary outcome** revealed the **difficulty** to split the data into **learning/test subsets**. To overcome this problem we have proposed a **specific methodology** based on **samples stratification**

■ Perspectives :

- Implementation of another model prediction for Stacking meta-model
- Integration of other Sub-models to improve stacking. for example:
 - Decision Tree based model on C4.5, Neural Network, Multiblock Redundancy Analysis, ...
- Extension of stacking prediction for a multi-class case ...