Visual Analytics *for* Machine Learning

Alexandru Telea
University of Groningen, the Netherlands

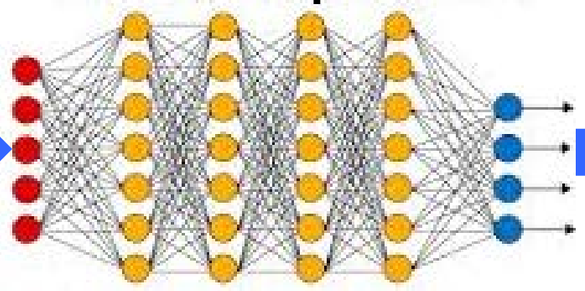
a.c.telea@rug.nl

Machine learning *vs* manual exploration

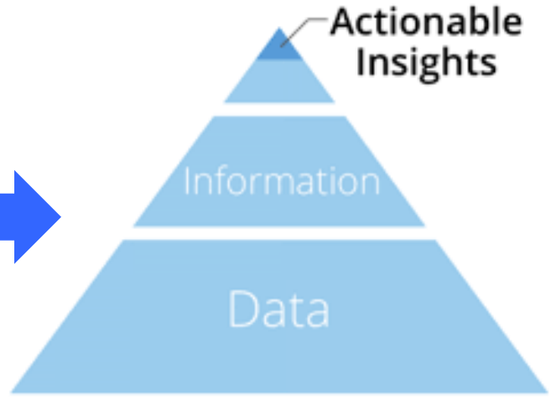
expensive!

input data

manual exploration



machine learning



output (insights)

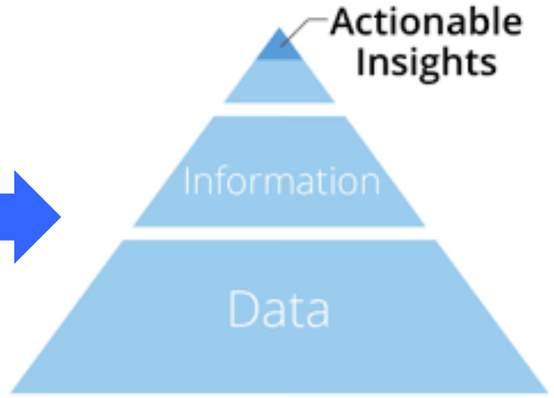
Machine learning *vs* manual exploration

input data

manual exploration

machine learning

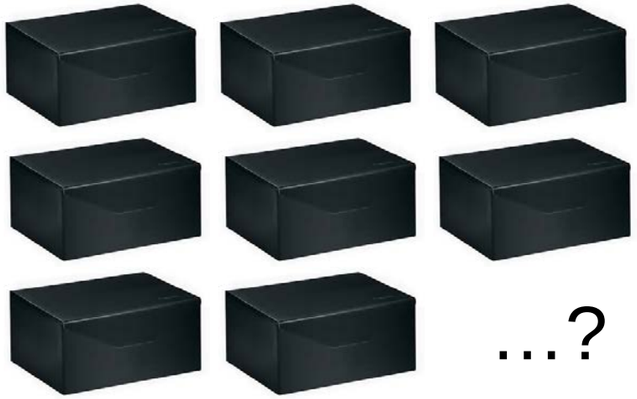
expensive too!



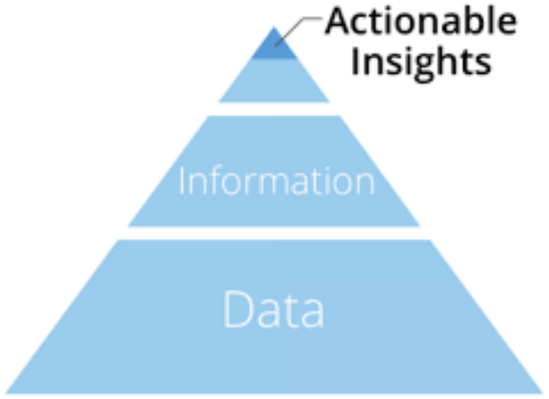
output (insights)

VIS 2017

Machine learning challenges: Algorithm choice



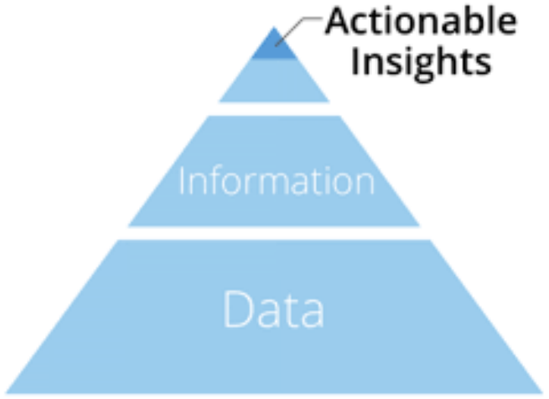
…?



How to choose the best *algorithm/architecture?*

(SVM, KDD, LVQ, RFC, ANN, XYZ,…)

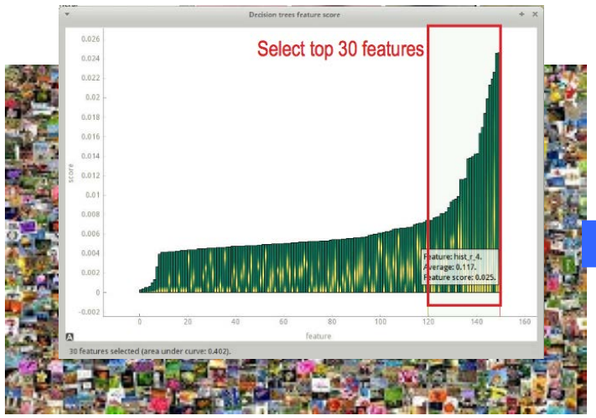
Machine learning challenges: Parameter choice

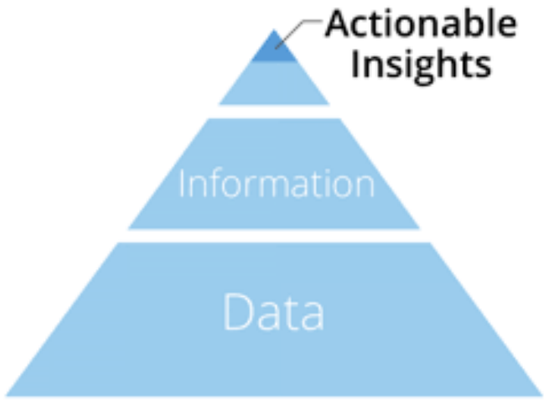


How to choose the best algorithm *parameters*?

VIS 2017

Machine learning challenges: Training



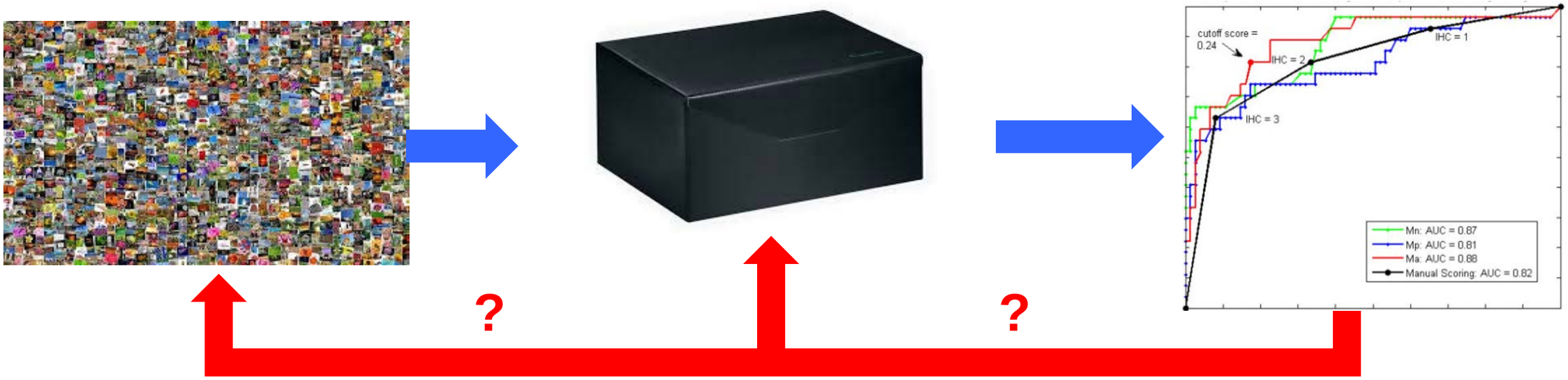


How to understand which *parts* of the input were good/bad for training?

- instance classes / feature classes
- individual instances
- parts of instances

VIS 2017

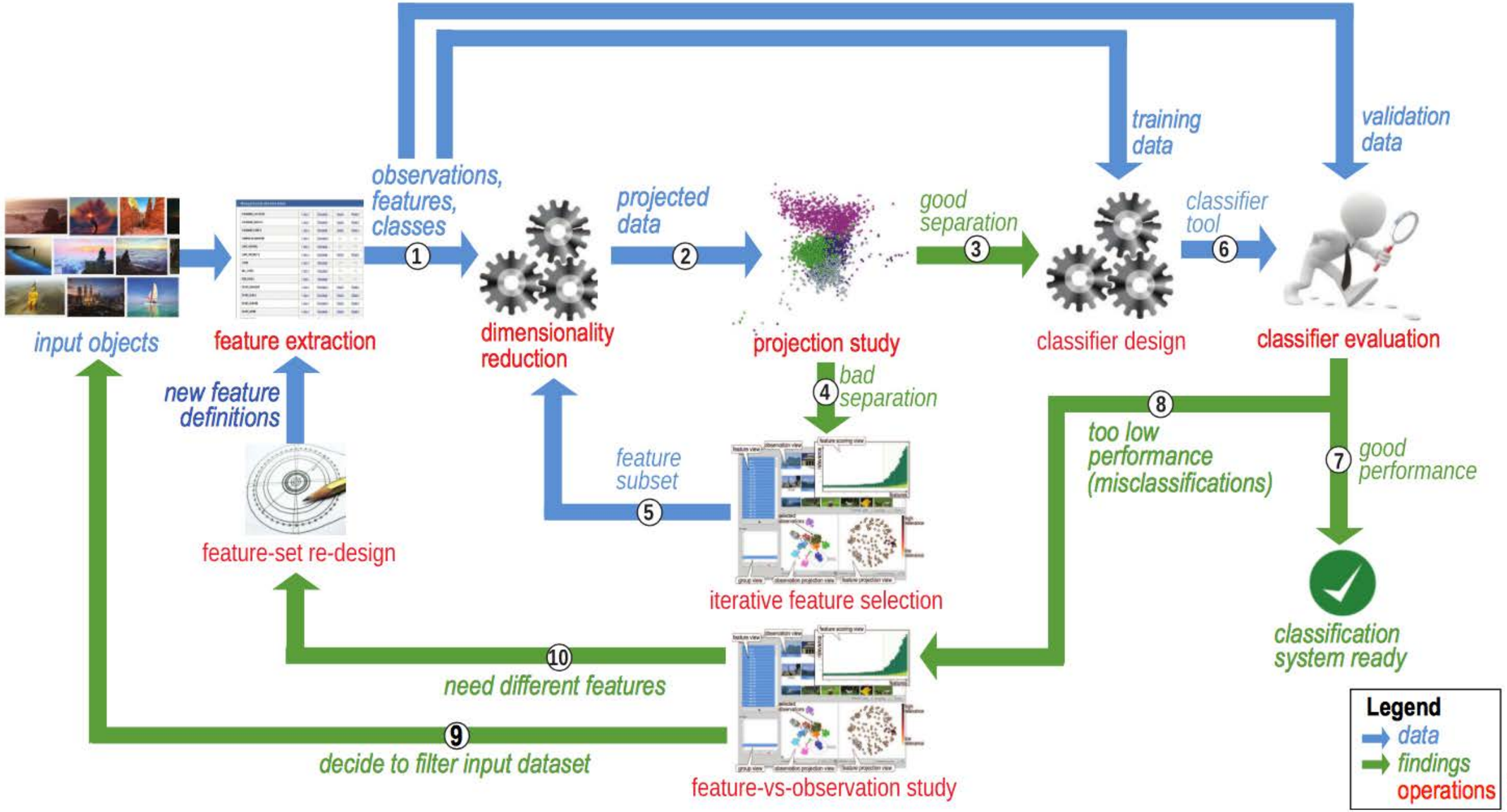
Machine learning challenges: Optimization



How to understand what to change in input/parameters to guarantee a given change in output performance?

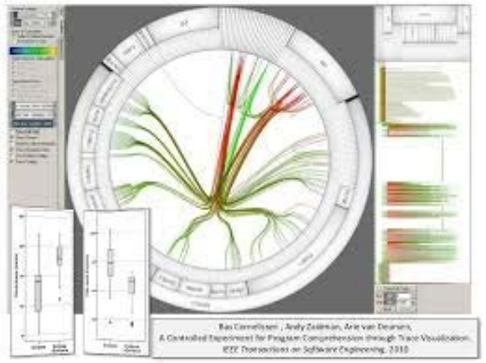
- different features?
- different training set?
- different classifier parameters?

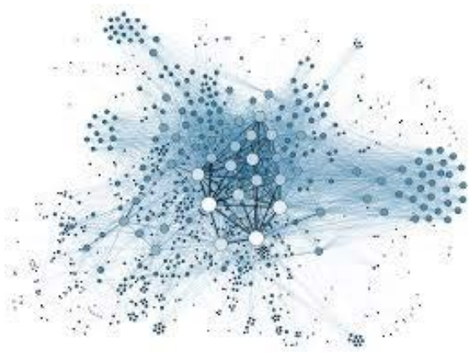
Workflow optimization is essential!

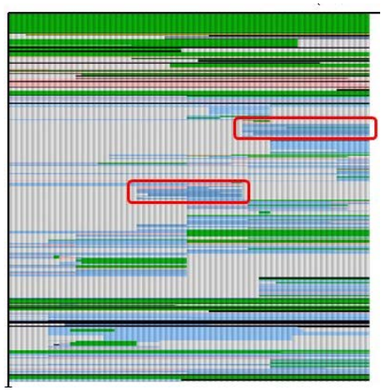


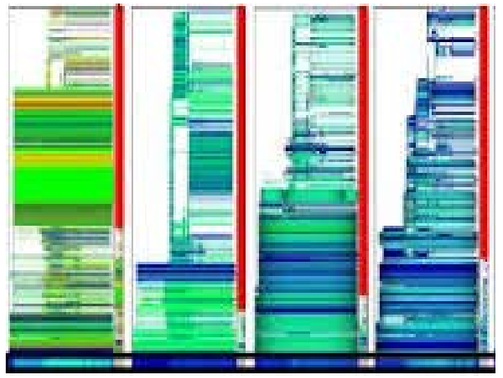
Rauber *et al.* (2016) Projections and visual aids for classification system design; doi.org/10.1177/1473871617713337

Idea: Learn from Program Comprehension!









Input data

- complex
- hybrid
- multivariate
- time-dependent

Program

Output (insights)

- high-level
- data/execution
- aim: debug
- aim: optimize

VIS 2017