On the use of supervised learning techniques to speed up the design of aeronautics components .

Gianluca Bontempi, Olivier Caelen

Université Libre de Bruxelles; Bruxelles, Belgium

Stephane Pierret, Cécile Goffaux

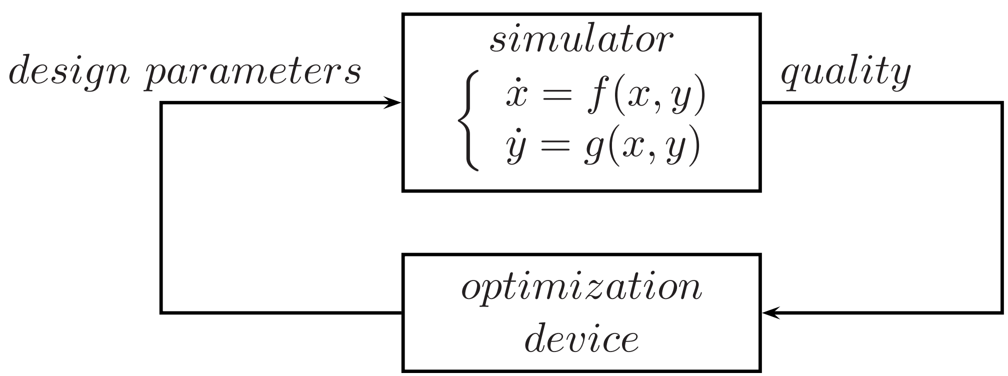
Centre of excellence in aeronautical research (CENAERO); Gosselies, Belgium

Machine Learning Group

Table of Contents

Design of a complex product(1)

- The design of a complex product can be seen as a search problem in the space of the design parameters which aims to maximize the quality of the product ($\Rightarrow$ optimization problem).
- The optimization device makes a great number of calls to a simulator to calculate the cost of the function to be optimized.

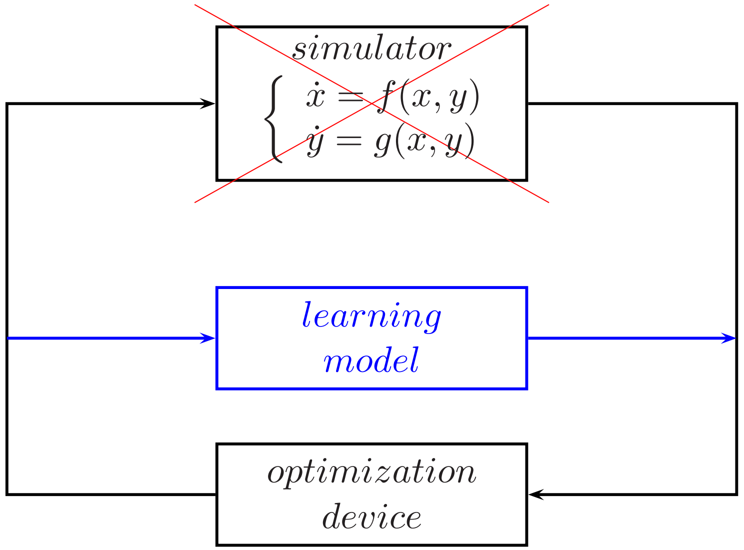


Design of a complex product(2)

- If the simulator is too complex, it can slow down the optimization process.
- This is often the case in aeronautics applications (turbine, heat pipe,...) where the relation between the design parameters and the quality criteria is modeled by a time consuming simulator.

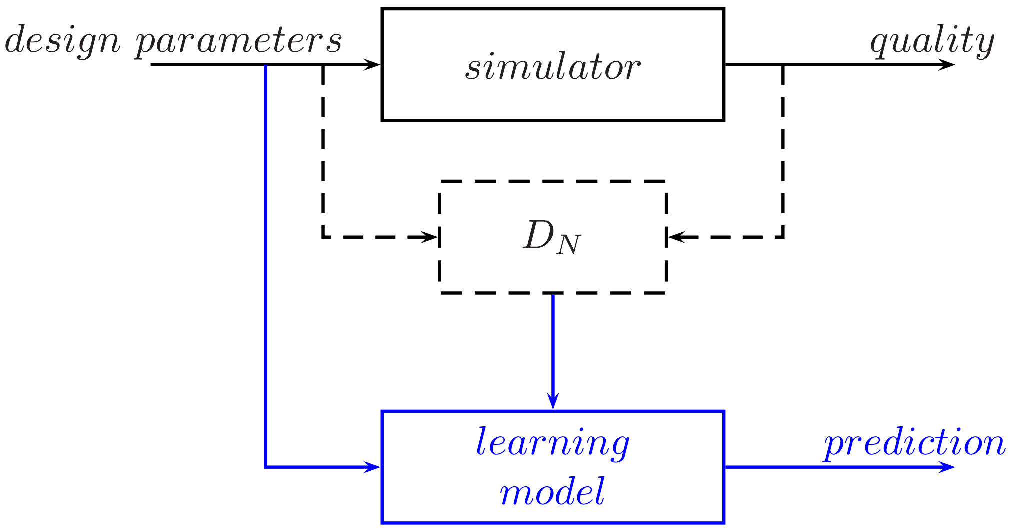
Speed up optimization process(1)

- A possibility to reduce the time required to evaluate each design configuration is the use of supervised machine learning techniques.
- The optimization device will try to use, as often as possible, the supervised machine learning.



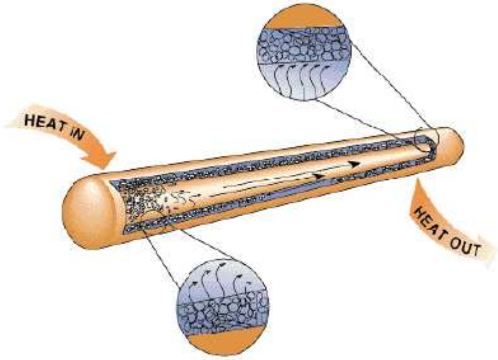
Speed up optimization process(2)

- Samples of data of a simulator are collected in D_N.
- D_N is then used to make a training on a supervised machine learning.
- After the training, the machine learning model can make predictions.



The Heat Pipe simulator

- We use a simulator of a heat-pipe to generate training samples.
- A heat pipe is a device for evacuating heat.
- One side of the heat pipe is in contact with the heat and the other one is in contact with outside.



Some notations

- Given two variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$.
- Let us consider the mapping $f : \mathbb{R}^n \rightarrow \mathbb{R}$, known only through a set of N examples $\{(x_i, y_i)\}_{i=1}^{N}$ obtained as follows:

$$y_i = f\left(x_i\right) + \epsilon_i$$

  where $\forall i, \epsilon_i$ is a random variable such that $E\left[\epsilon_i\right] = 0$ and $E\left[\epsilon_i \epsilon_j\right] = 0$, $\forall j \neq i$.
- we seek the learning model $\hat{f}\left(x\right)$ with the best prediction capacity.

Machine Learning Group

Global vs. local(1)

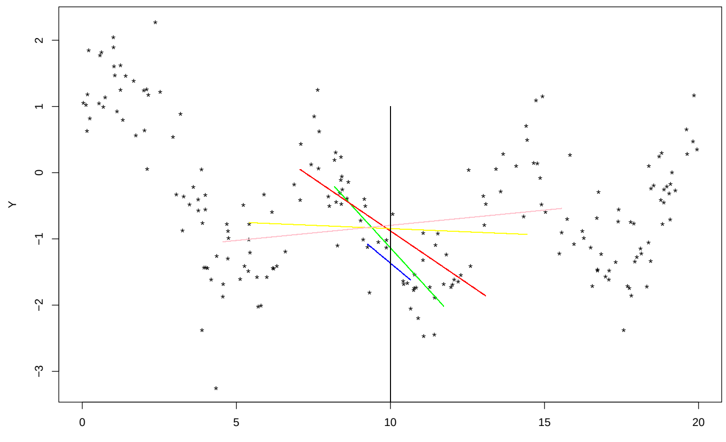
- The traditional approach to supervised learning is the *global* modeling which **describes** $\hat{f}\,(.)$ **with an analytical function over the whole input domain**. Examples:
  - linear regression models. $\hat{f}\,(x) = x^T\beta$
  - two level feedforward neural networks[2].

$$\hat{f}\,(x) = h\left(\sum_{m=0}^{M} w_{1m}^{(2)} g\left(\sum_{k=0}^{n} w_{mk}^{(1)} x_k\right)\right)$$

Machine Learning Group

Global vs. local(2)

- An other approach is the local modeling.
- An example is the *k-Nearest Neighbour* model.
- The problem in the k-Nearest Neighbour is to find the best value of ’k’.
- The *lazy-learning* model [1] finds automatically, by *PRESS*[4], the best value of *k*.



Global vs. local(3)

The lazy-learning has some promising features:

- The reduced number of assumptions.
  Example:
  - No assumption on the existing of a global function.
  - No assumption on the properties of the noise.
- On-line learning capability
- Effective feature selection
  - Reducing the cost of feature selection by the *Hoeffding race*[3] .

Machine Learning Group

Tests of two models

- In this study, we compare the generalization capacities of a feedforward neural networks with these of a lazy learning on data generated by the Heat Pipe simulator.

Experimental Process

- The experiments use two datasets; I call them
  - the set D1
  - the set D2
- Two sessions of experiments are carried out:
  - the first with no feature selection.
  - the second with a preliminary selection of the relevant design parameters by Hoeffding race.

The two datasets

- The set D1:
  - it is composed of $N = 1260$ samples.
  - it has $n = 3$ inputs and $m = 2$ outputs.
- The set D2:
  - it is composed of $N = 820$ samples.
  - it has $n = 6$ inputs and $m = 2$ outputs.

The validation procedure

- We adopt the following procedure in order to assess the prediction capacity:
  - The total of samples is randomly divided into two halves.
  - The first half is used for the training.
  - The second half is used for the validation.
  - The estimation of the MISE is returned by :

$$\frac{1}{N/2} \sum_{\langle x,y \rangle \in V(D)} \left(y - \hat{f}\left(x\right) \right)^2$$

Where $V\left(D\right)$ is the validation subset of the sample set D_N.

Machine Learning Group

Experimental Results (1)

- Before feature selection : Mean square prediction error for the two outputs of $D1$:

Learner	Output 1	Output 2
Lazy-learning	2.9e-04	2.2e-05
Neural network	5.0e-03	1.2e-04

- Before feature selection : Mean square prediction error for the two outputs of $D2$:

Learner	Output 1	Output 2
Lazy-learning	1.4e-02	1.0e-05
Neural network	2.2e-02	4.8e-05

Experimental Results (2)

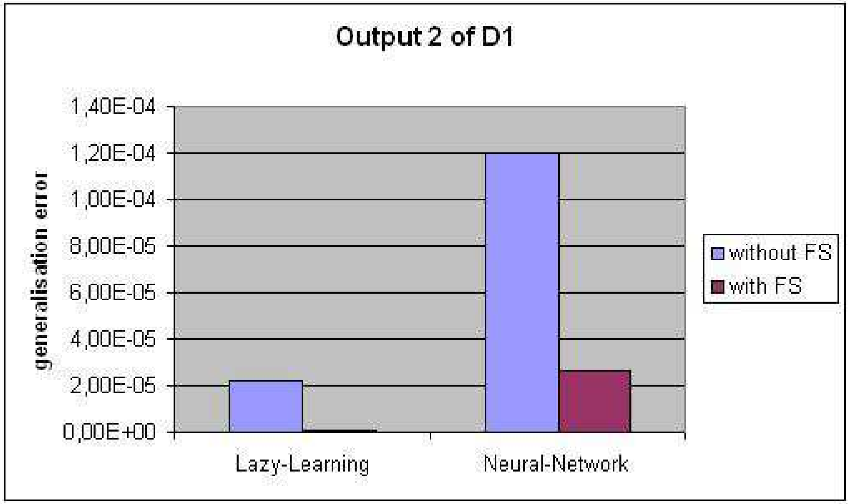
- Next, a feature selection process has been made on the two training sets.
- On the first output variable of $D1$, the best input subset is the complete set of input variables.

Machine Learning Group

Experimental Results (3)

- On the second output variable of $D1$, the feature selection process finds another input subset.

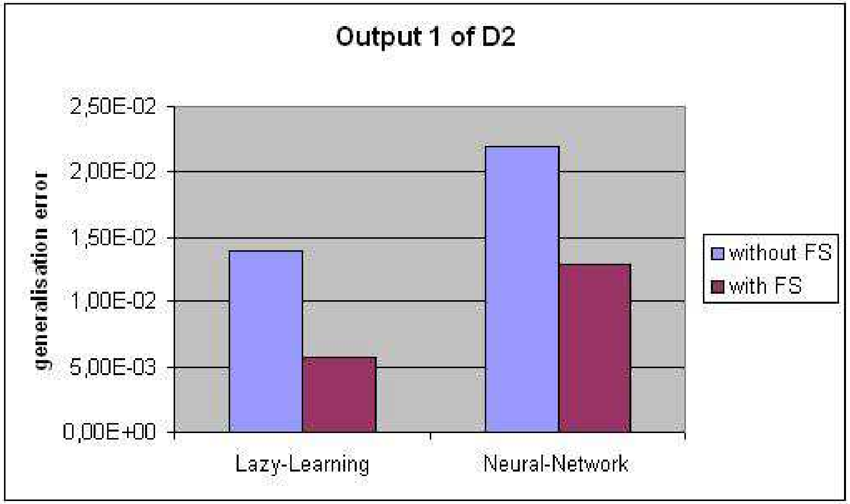
Learner	Output 2
Lazy-learning	7.0e-07
Neural network	2.6e-05



Experimental Results (4)

- On the first output variable of $D2$, the feature selection process finds another input subset.

Learner	Output 1
Lazy-learning	5.7 e-03
Neural network	1.3 e-02

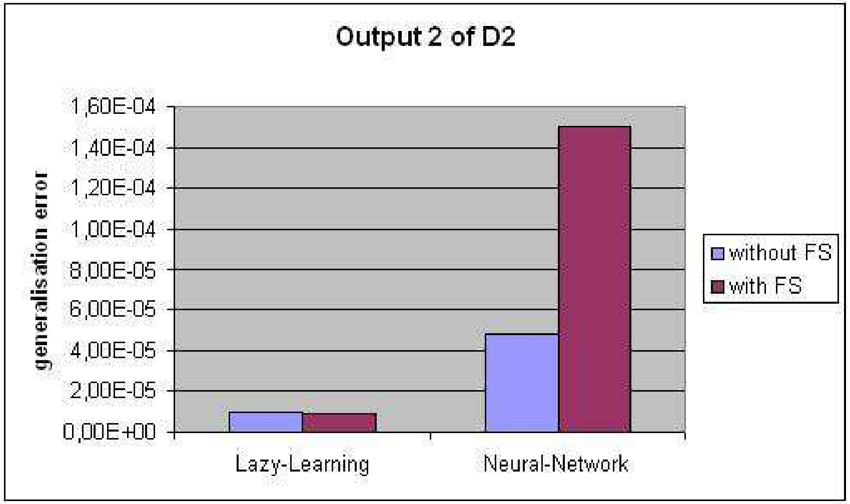


Machine Learning Group

Experimental Results (5)

- On the second output variable of $D2$, the feature selection process finds another input subset.

Learner	Output 2
Lazy-learning	8.8e-06
Neural network	1.5e-04



Conclusion

- The machine learning model can be used to speed up the design process of a complex product.
- The Leazy-Learning technique appears competitive with more conventional machine learning technique, like feedforward neural networks.

Machine Learning Group

Future work

- Extending the experiments to a larger number of design parameters and quality objectives.
- Integrating the machine learning in the optimization process.

References

[1] G. Bontempi. *Local Learning Technique for Modeling, Prediction and Control*. PhD thesis.

[2] Bishop C. *Neural Networks for Pattern Recognition*. Oxford UP, 1995.

[3] Oden Maron and Andrew W. Moore. The racing algorithm: Model selection for lazy learners. *Artificial Intelligence Review*, 11(1-5):193–225, 1997.

[4] R. H. Myers. *Classical and Modern Regression with Applications*. PWS-KENT, Boston, MA, 1990.

THANK YOU

for your attention!

You can find the slides of this presentation at this adress:

http://www.ulb.ac.be/di/map/ocaelen/

For more information on the Lazy Learning R Package see here :

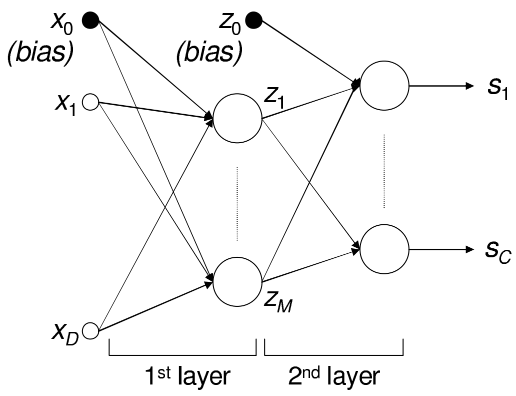
http://iridia.ulb.ac.be/∼lazy/

Machine Learning Group

ANNEXE

The feedforward neural networks (1)

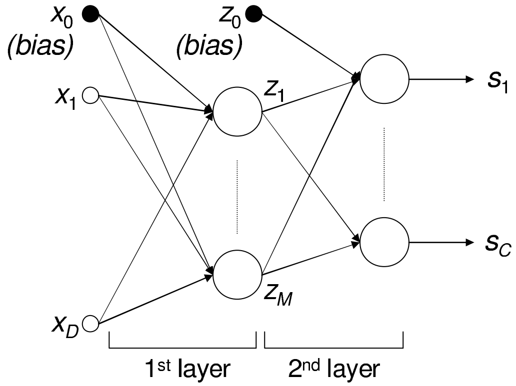
- The feedforward neural network, that we use, has two layers.



- The first layer is based on a sigmoidal transfer function and the second layer on a linear transfer function.

Machine Learning Group

The feedforward neural networks (2)



- The number of hidden nodes (M) is determined using an empirical relation which is a function of :
  - the number of training data.
  - the input dimension.
  - the output dimension.

The two datasets : D1

$N = 1260$ samples, $n = 3$ inputs and $m = 2$ outputs.

- Input parameters
  - the internal diameter of the heat pipe.
  - the diameter of the groove (d_{hyd}).
  - the inclination angle of the heat pipe.
- Output parameters
  - The power (in Watt) released by the heat pipe.
  - The external diameter of the heat pipe.

The two datasets : D2

$N = 820$ samples, $n = 6$ inputs and $m = 2$ outputs.

- Input parameters
  - The internal diameter of the heat pipe.
  - The number of groove in the heat-pipe.
  - The diameter of the groove (d_{hyd}).
  - The width of the bottom of the grooves (w_b).
  - The width of the top of the grooves (w_t).
  - The depth of the grooves (h) .
- Output parameters see D1

Machine Learning Group

The two datasets : figures