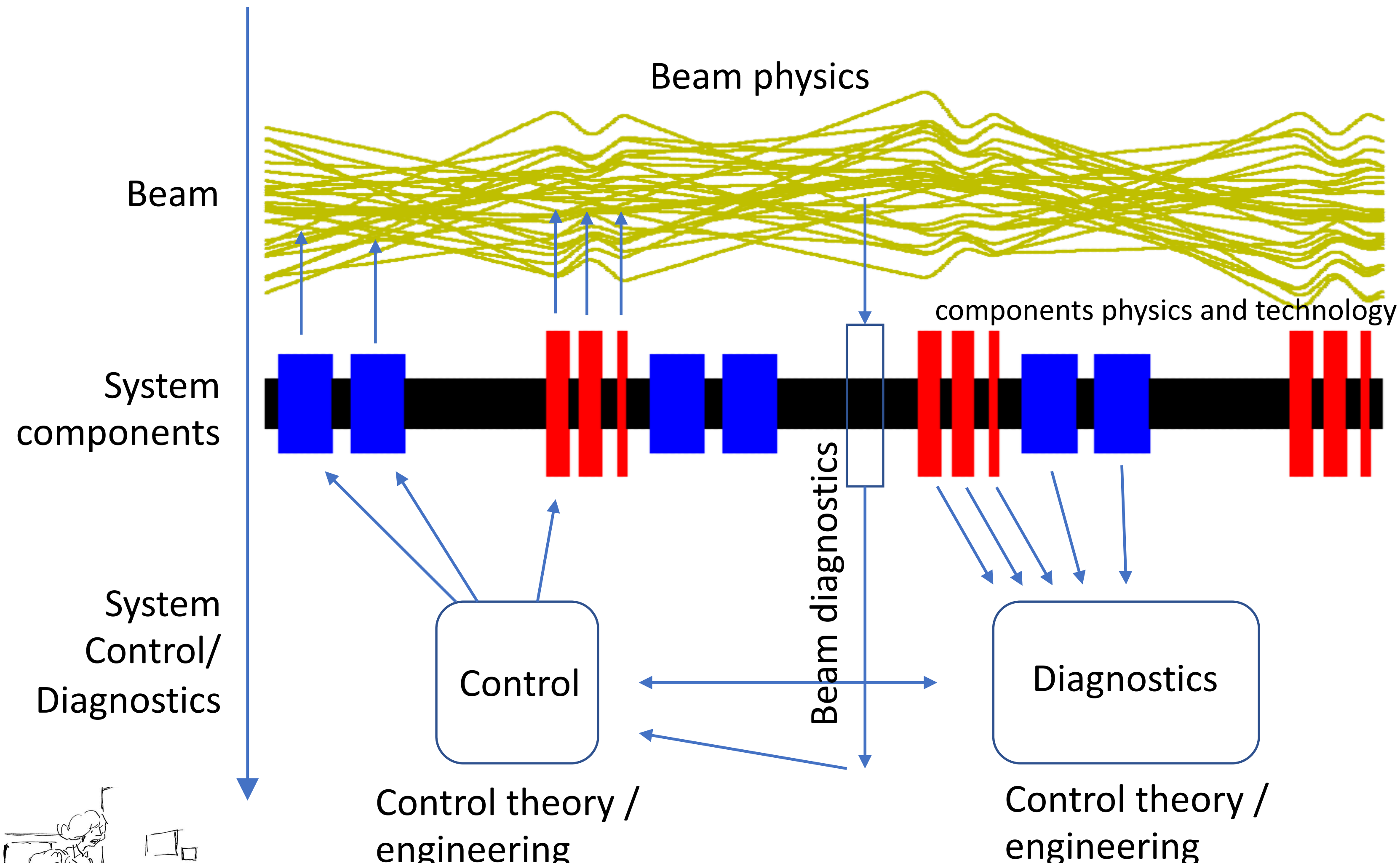
# Machine learning challenges

G. Franchetti, GSI

ARIES WP6 APEC & iFAST WP5.2 PAF
joint Brainstorming & Strategy Workshop (BSW22)

Valencia, Colegio Mayor Rector Peset, 30 March – 1 April 2022

# Accelerators




# Accelerators as a complex system

**Large expertize in:**

- Designing
- Maintaining
- Optimizing

Accelerator performance

Extreme control, or fast settings for optimization requires new techniques

**Complex system**

System: the full accelerator complex

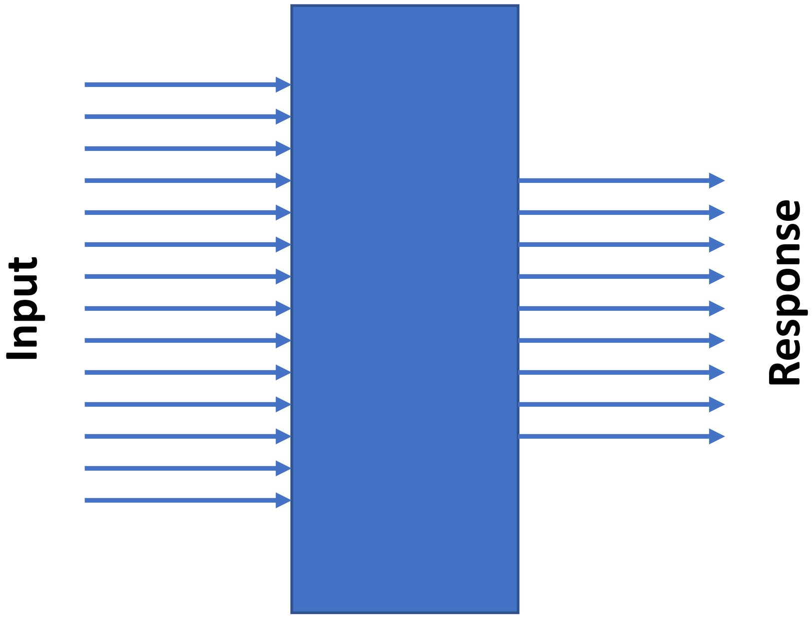
Complexity: thousends on interconntected components

Network: components interconnected also via “physics” coupling

Nonlinearity: yes we have

Emergence: optimizations are sometimes resulting of empirical rules

# Modeling a complex system

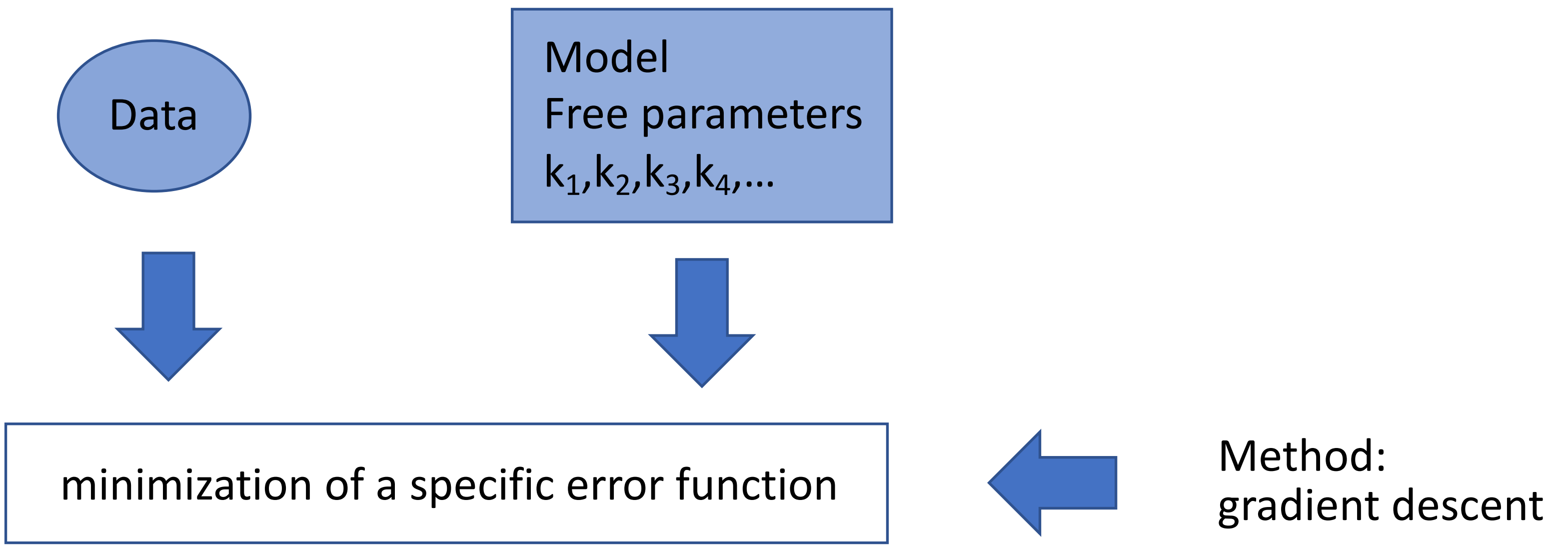



**Interpolation:**
finding the mathematical function that goes through the sampling

**Fitting:**
being compatible, similar or consistent

**Generalization:** constructing the "true" system response by sampling the response.
Generalization →
the capability of explaining new cases

# Change of paradigm: "learning"

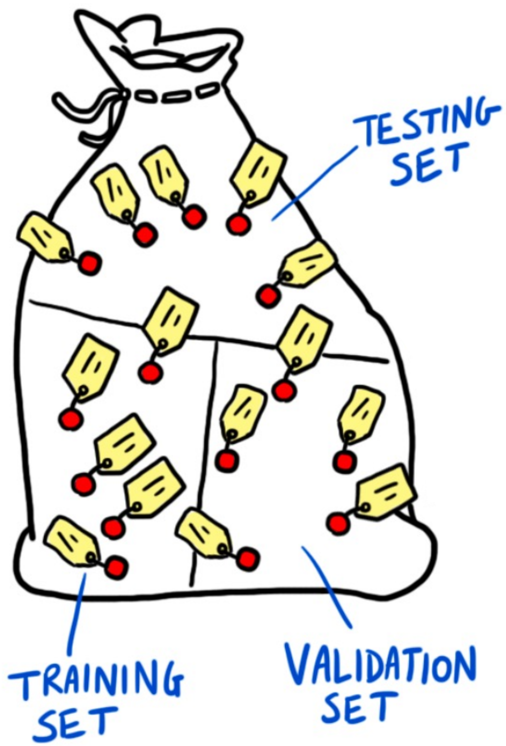



**The paradigm**

Minimization of the fitting errors is here interpreted as "learning" by example but the generalization is not a easy thing to prove....

# Does it learn?

**careful experimental procedures** are necessary to measure the effectiveness of the learning process




How do you know if **generalization** is reached?

- Accuracy
- Precision
- Recall
- Confusion matrix

The larger the data used the more it learns

Roberto Battiti · Mauro Brunato
**The LION Way Machine Learning *plus* Intelligent Optimization**

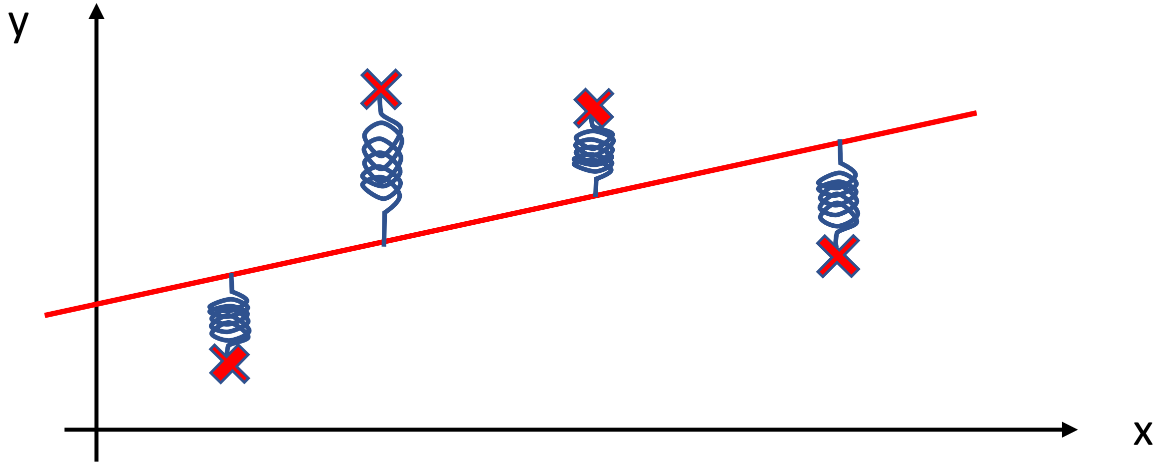
# Linear models

* Linear Regression. (Linear models for classification)

* Ansatz → smoothness

* Learning → minimizing the sum of squared errors

* Tricks for numerical sensitivity → ridge regression



# Learning through statistics

The goodness-of-fit is given by the chi-squared →

$$\chi^2 = \sum_{i=1}^{N} \left( \frac{y_i - f(x_i)}{\sigma_i} \right)^2$$

* Difficulties in the relation between **model, interpolation and generalization**

* Least-squares is a maximum likelihood estimator

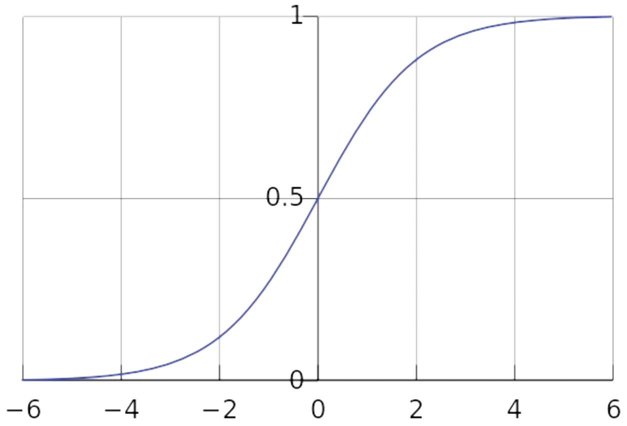
* Null hypothesis testing

* Bootstrapping → strange approach.. (*random sampling with replacement)*

* Statistics is a minefield if assumptions are wrong

# Nonlinear Models

Based on logistic regression $P(t) = \frac{1}{1+e^{-t}},$



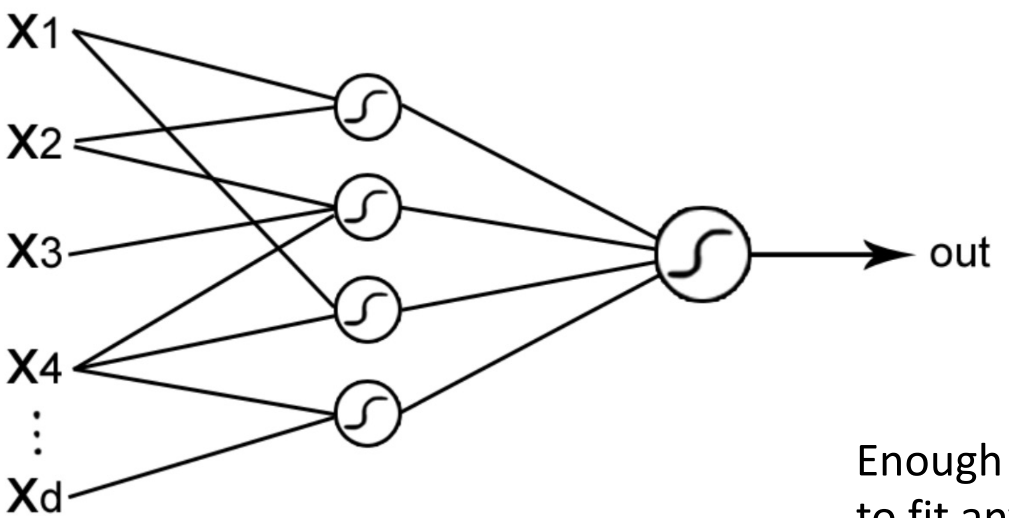

Nonlinear model

$$P(\boldsymbol{x}) = \frac{1}{1+e^{-(\boldsymbol{w}^T\boldsymbol{x})}}.$$

weights features

Learning by maximum-likelihood

# Neural Networks

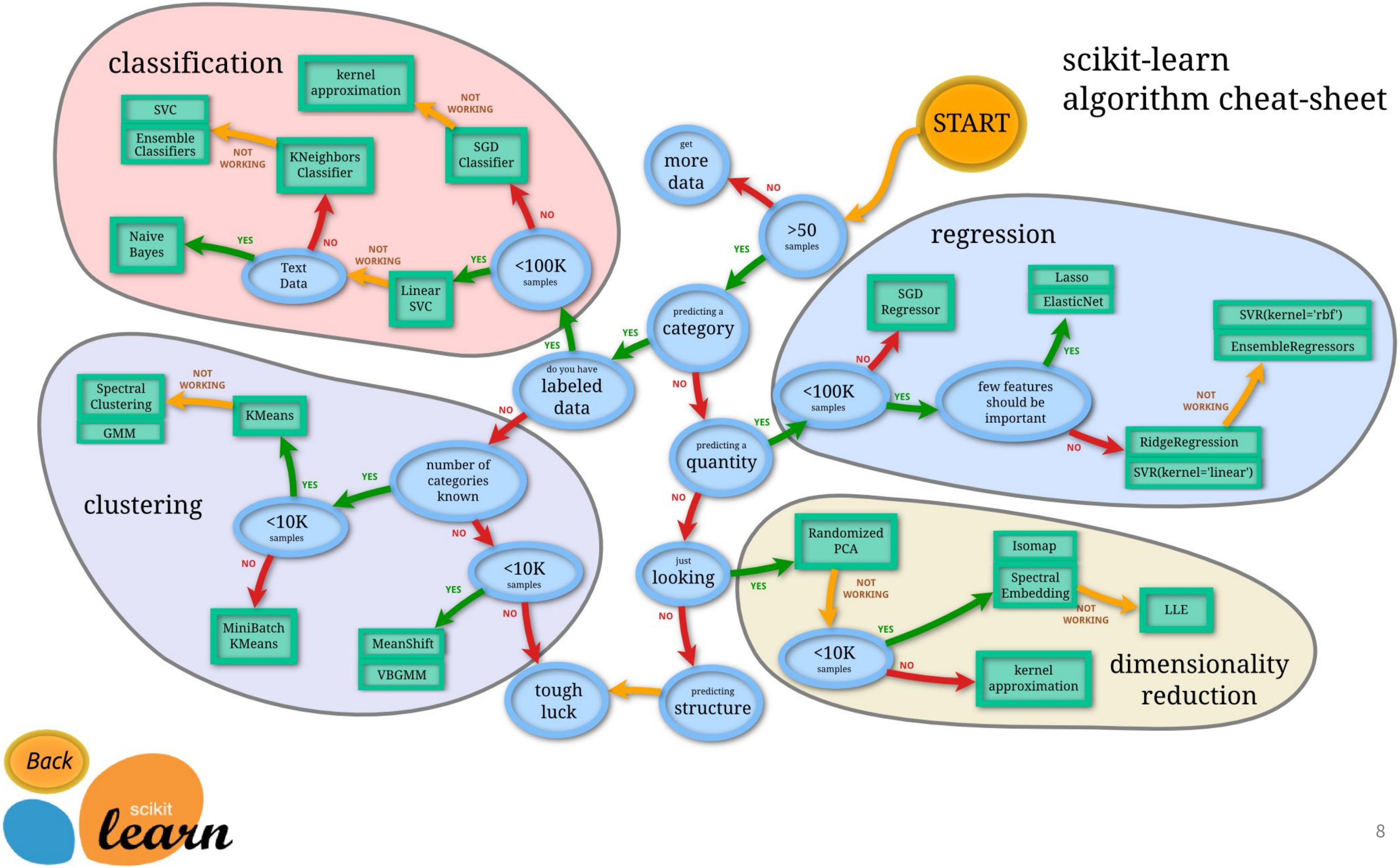



Enough layer provide flexibility
to fit anything

Fit by training (backpropagation)
→ large data set

Issue → does generalization works?

# A jungle of methods




8

# Feature selection and information

Reducing the number of input attributes → smaller model size and better human understandability, faster training and running times, possible higher generalization.

What are the relevant features? → difficult without a specific modeling method and their mutual relationships. Physics model we can make

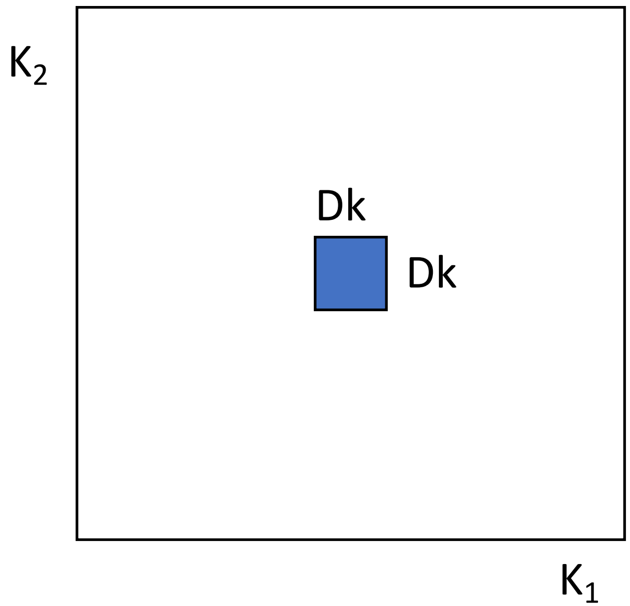
* Trust the correlation coefficient only if there are reasons for *linear* relationships,

* Use chi-square to identify possible dependencies between inputs and outputs by estimating probabilities of individual and joint events.

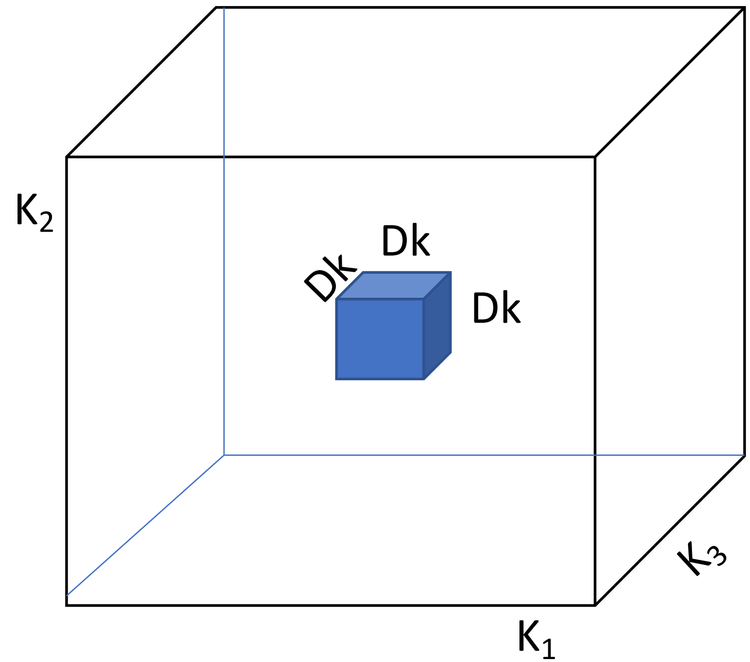
**Mutual information** to estimate arbitrary dependencies

It looks far fetching, but ultimately is based on statistics and information theory

# The curse of dimensionality




$(Dk/k)^2$




$(Dk/k)^3$

Scaling as $(Dk/k)^d$ → for large dimension proper finding of good parameters is very hard

ML good for large parameter space? → how big should the training be?

# Machine learning for accelerators

→ Developing/adapting of a sub class of ML research for accelerator

The problems in accelerators

Problems of complex systems with large number of knows

Specific issues of accelerator, which can be almost be modeled starting with physics model

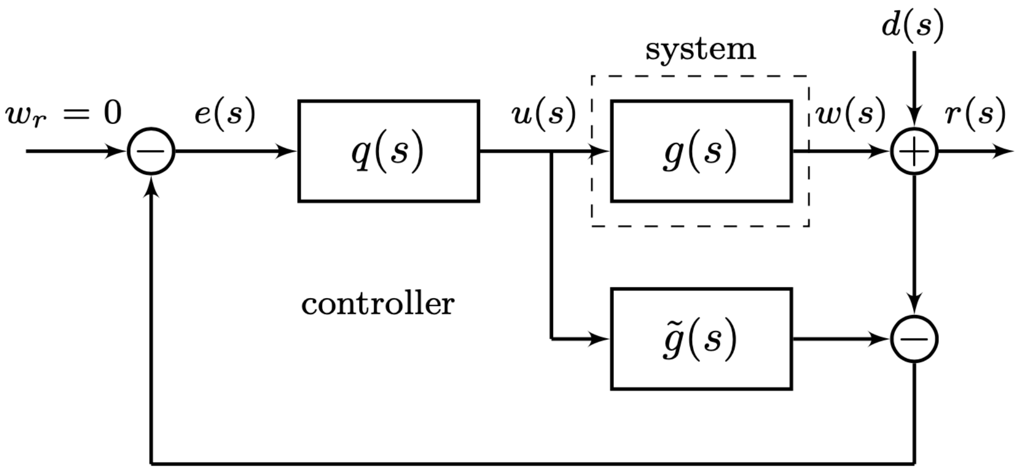
# Time dependency

Time dependence change the system → learning by training more difficult

Extra requirement → not only it is required generalization, but also prediction of the dynamics…. Very hard

Dynamic training?

Replacement of a feedback system controller with a fast learning ML ?
Example:
orbit control during fast acceleration is better with ML techniques?




Mirza Sajjad PhD thesis

# Magnet sorting for DA optimization

Given each magnet equipped with its unique multipolar set, one can compute DA

For N magnets there are N! permutations → the task is huge for standard codes

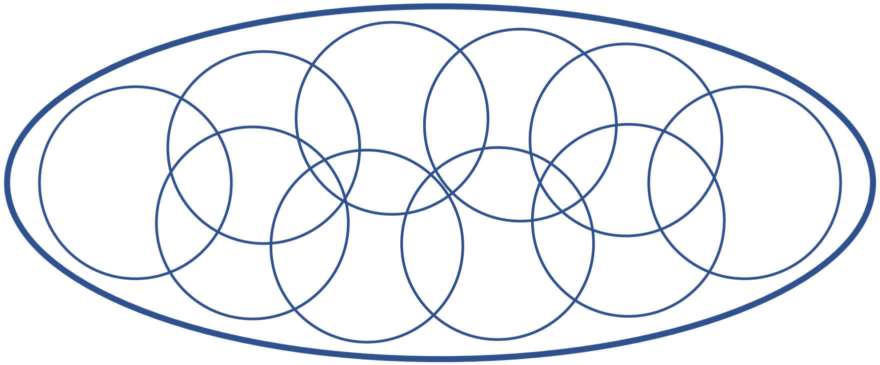
Can ML help to make sorting?

Example: →

application of a hierarchical feature selection to cluster magnets according to significant correlation

Through ML identification of the relevant subspaces in the magnet permutation space in order to reduce the brute force configuration space.

# Advanced magnet modeling



Standard multipole expansion works in the convergence radius

$$B_y + iB_x = \sum_n (b_n + ia_n)(x + iy)^n$$

There is no mathematical prove that the field is well described outside the circle

Multiple measurements provides multiple dataset of multipoles, which merging is quite difficult without creating artificial Field huge fluctuation

ML could be trained to merge several multipoles providing the best magnet modeling

# Diagnostics

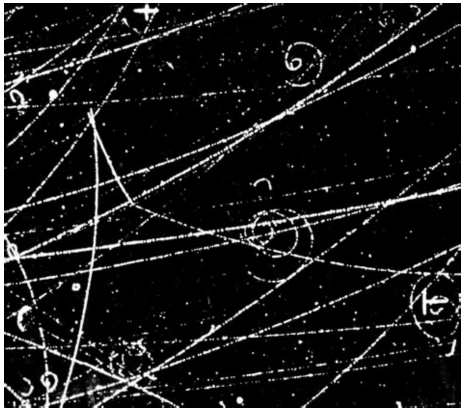
diagnostics

Correction of measurements by unwanted effects: example cleaning of beam profile in a IPM when measuring an high intensity bunch.

Denoising (auto-encoder) signals for high precision measurements

????

Pattern recognition



# ML classification of operation

Physics experiment proposal → Approval by the Physics Advisory Committee (PAC) → Scheduling by the Beam time coordinator

↕

Discussion with all Technical group

Could one use ML to rank the difficulty of a proposed experiment?

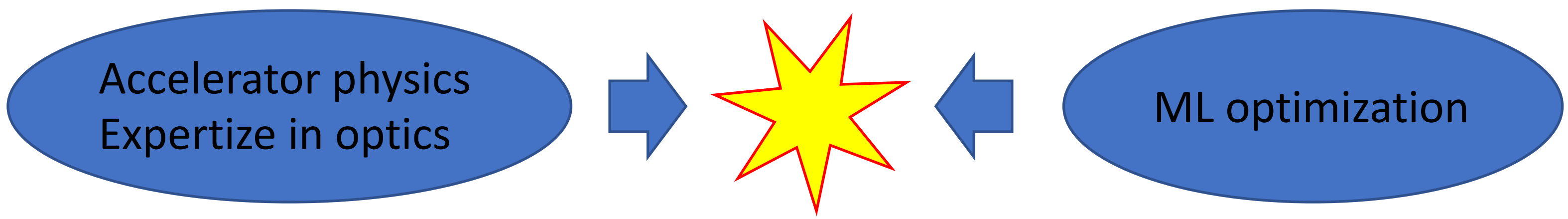
ML will take into account of real time specific issues and give ranking: for example including the probability of failures

# Design with ML ???

May be or may be not possible …

design

Design optimization might be possible, but it requires a highly specialized approach: only experts in ML  may be able to dig out the features and make the training of a deep neural network




How can ML be trained by a human?  Is it possible asking Oide-San to produce 100K model accelerator design for ML, NN training?….)

# Non-smooth systems ??

Can machine learning help in characterizing the systems with sensitivity threshold?

Example: Can the dependencies of the threshold of beam Instabilities be learnt?

# Summary

The learning is obtained by training

Features are the key and the proper training too

Generalization **cannot by proven**, but not contraddicted by **validations** and **tests**

**The advantages are in using a highly flexible model that**
**the training fix into a data-defined characterization (learning).**

The learning it is a classical **minimization problem**, but the structure
of the general method allows for **"fitting"** potentially any problem.

A plethora of anomalies is affecting the use of these methods for a blind user,
which requires an "learning" of what machine learning technique should be used