ATB
Leibniz-Institut für
Agrartechnik und Bioökonomie

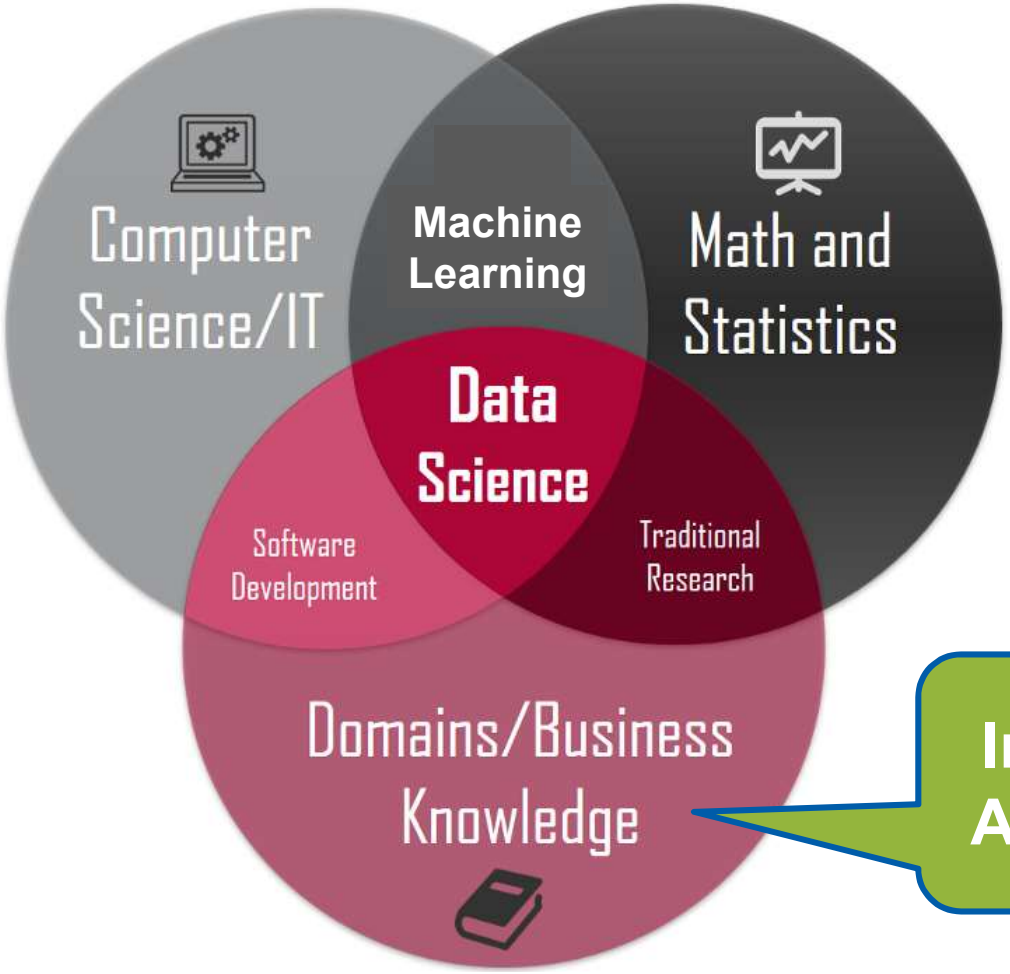
Data Science & Machine Learning in Agriculture

Impuls-Talk

Dr. Julian Adolphs
Department Data Science

Leibniz
Association

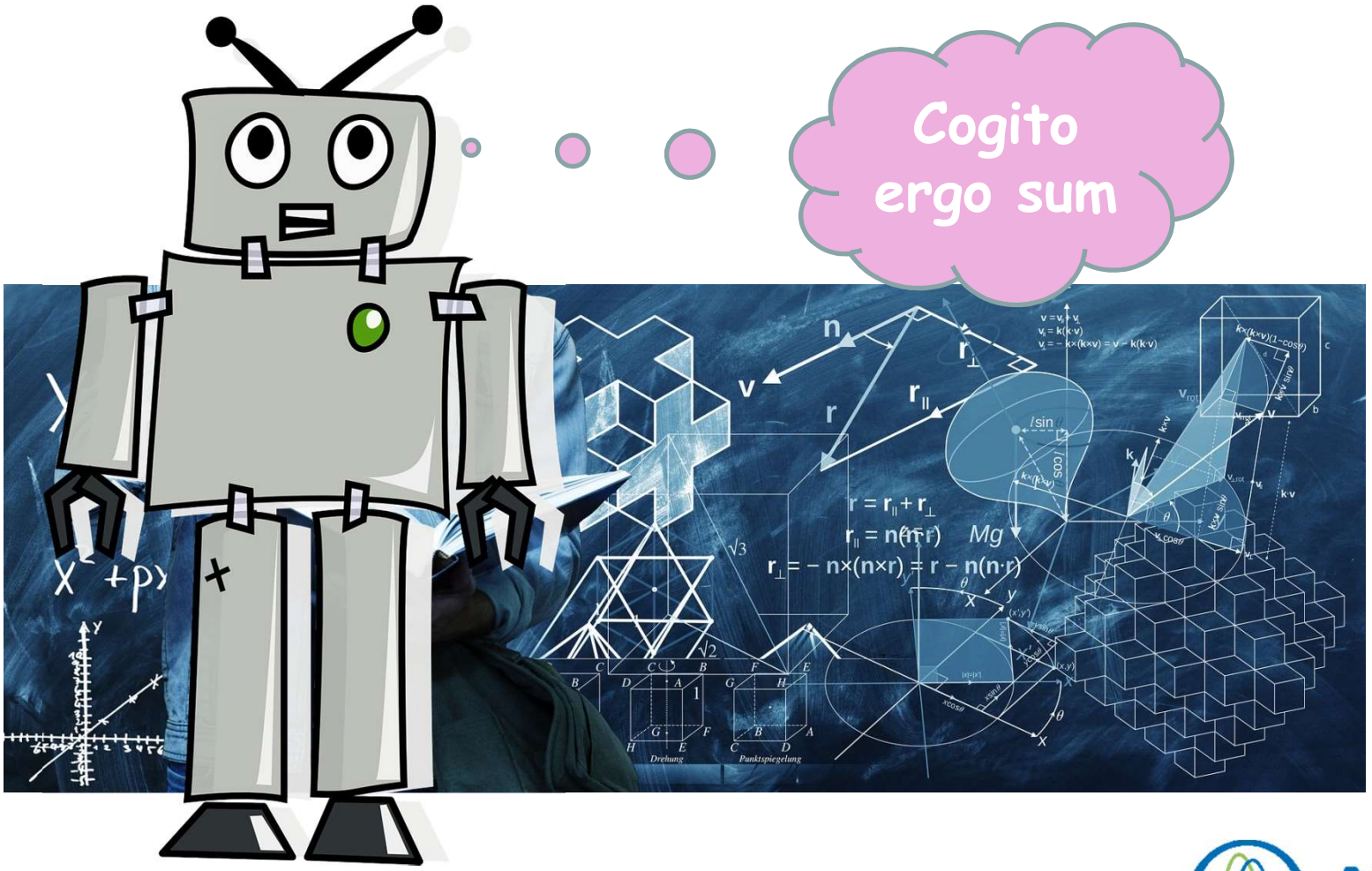
What is Data Science and Machine Learning ?



In our case: Agriculture !

Picture from https://towardsdatascience.com

Machine Learning is not about Teaching Machines to Think or Develop Consciousness.



What is Machine Learning (ML)?

Extraction of **Knowledge** from **Data**.

ML is a subset of artificial intelligence.

ML algorithms build models based on **training data**, in order to make **predictions** without being explicitly programmed for the task.

Machine Learning Algorithms

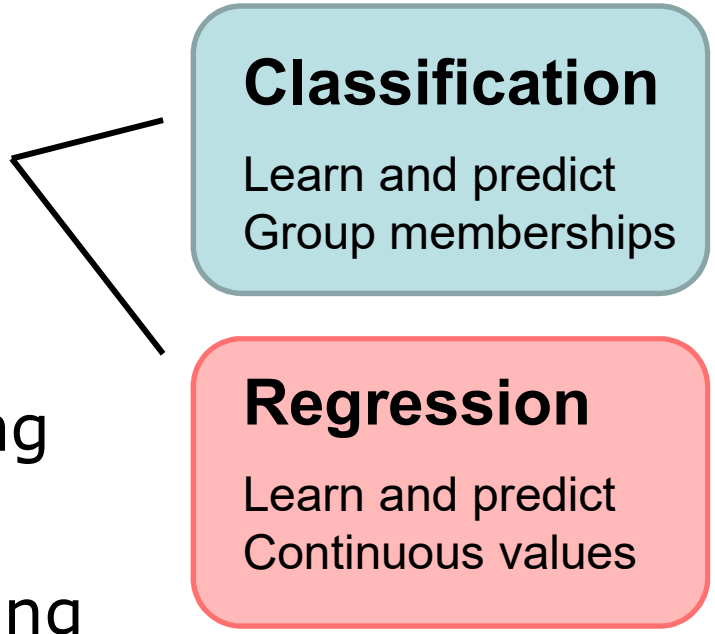
Three categories of Machine Learning (ML):

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Machine Learning Algorithms

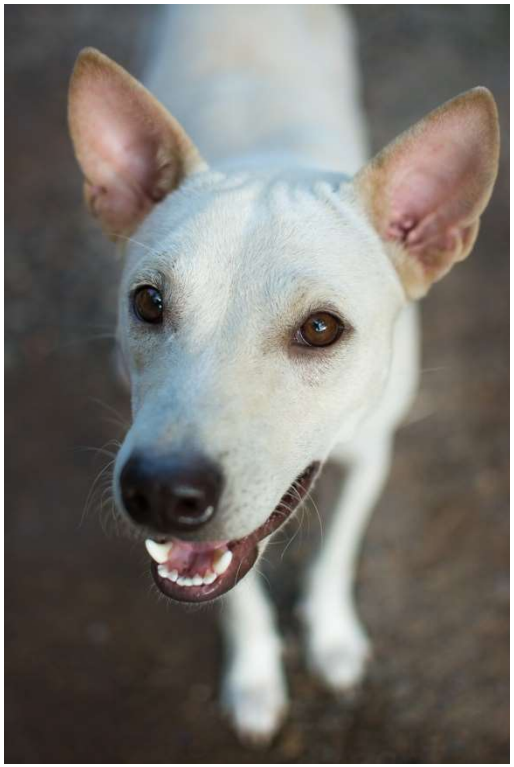
Three categories of Machine Learning (ML):

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning



Supervised ML – Classification (Discrete Values)

Dog or Cat ?!



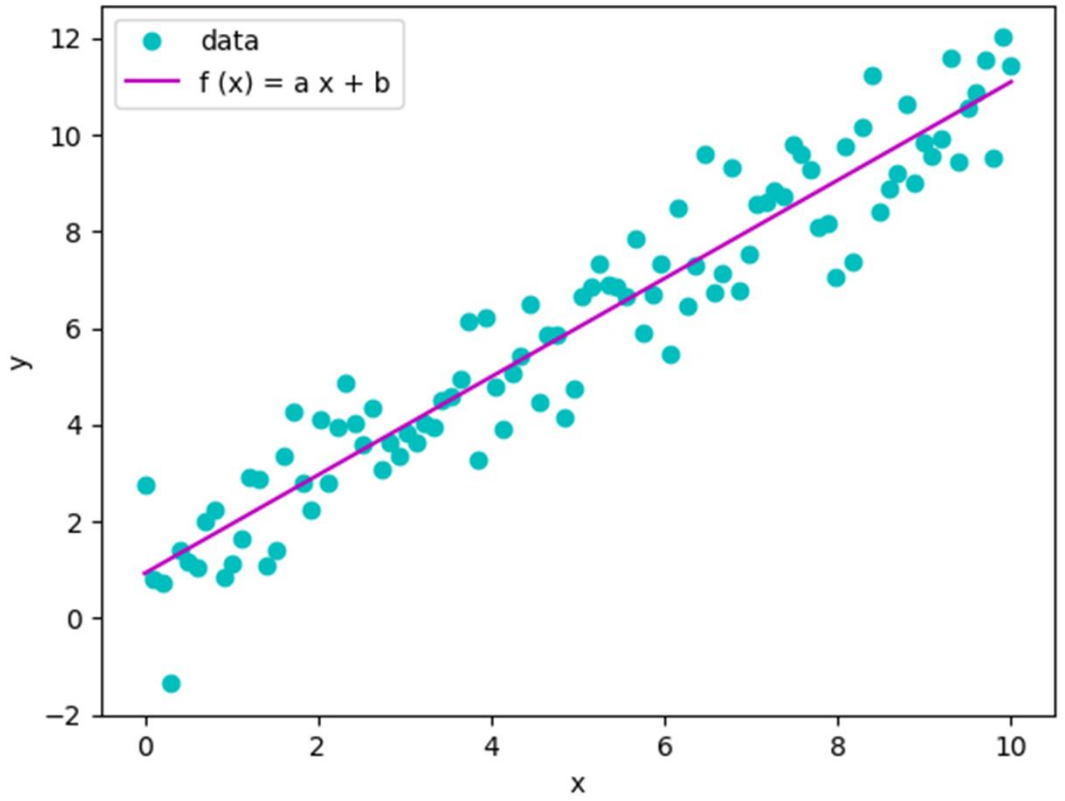
Classification in Agriculture – Disease Detection

Desease or no desease ?!

Disease: stipe rust

Supervised ML – **Regression** (Continuous Values)

Linear Regression of the Data with linear Function: $f(x) = a\,x + b$



generalised Model

$a = 1.0167$

$b = 0.9280$

Supervised ML – **Regression** (Continuous Values)
non-linear, multidimensional

Input-data	Output-data
Features	***Labels / Targets***
(size, # rooms, age, …)	*(House price)*

Features (of the houses):

- size / sqm
- # rooms
- # bathrooms
- condition
- age / years
- garden / sqm
- U-station / km
- garage / sqm
- …

number of features: d

Targets:

Prices of the houses / $

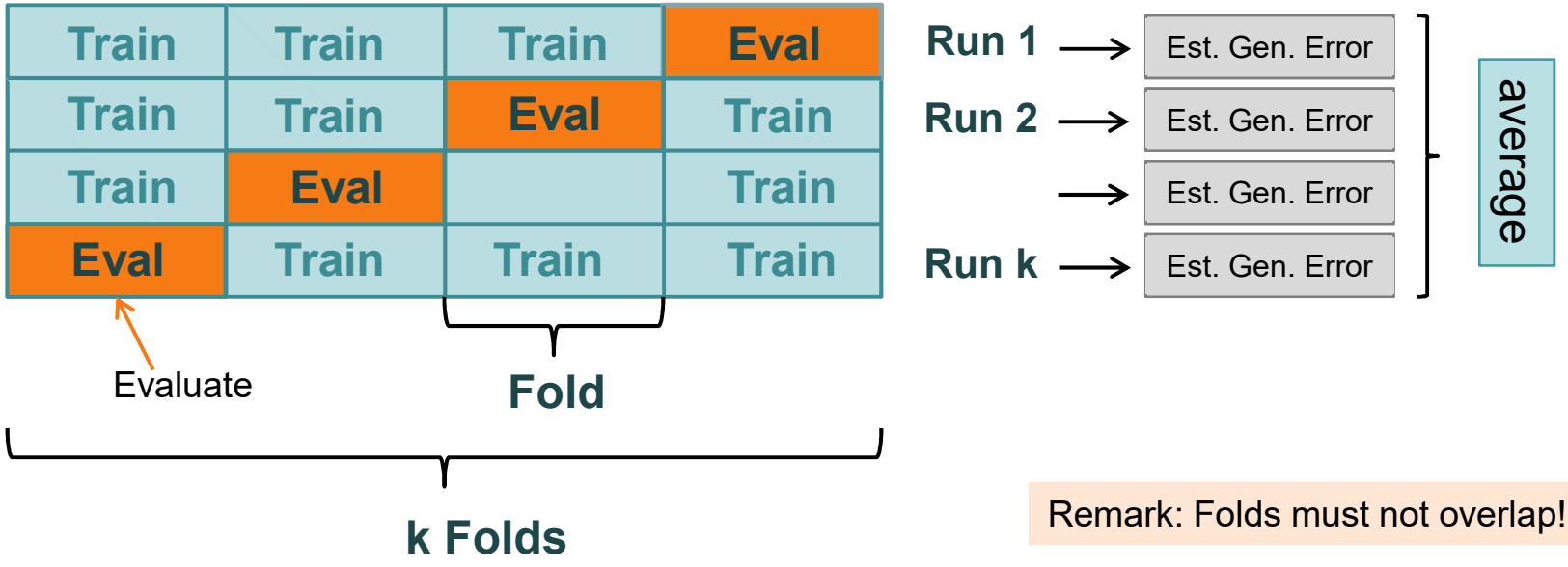
ℓ **samples**

Cross Validation

How to estimate how good our model is?

Split data into **train set** and **test set**:

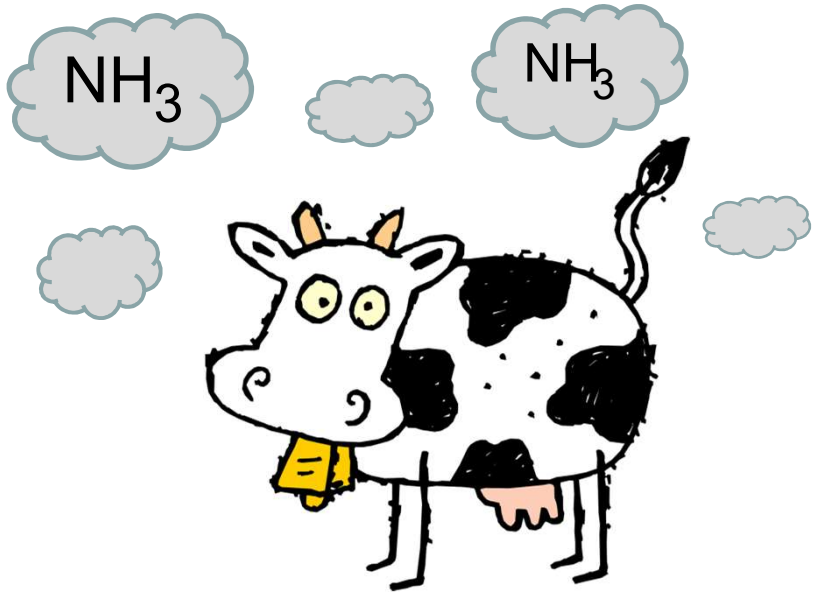
Test error is an estimate for the **generalization error**.



Remark: Folds must not overlap!

Regression with Machine Learning Methods

In the **Training Session** we study real **emission data**, measured in a cow barn in germany.



We use Machine Learning Algorithms:

- Linear Regression
- Polynomial Regression
- Random Forest
- Support Vector Machines (SVM)
- Artificial Neural Networks (ANN)

We use the very useful programming language python

(previous knowledge in any programming language helpful)

python™

Trainers in the Machine Learning session: Julian Adolphs & Sabrina Hempel

ATB