Cyber-Physical Systems
in Mechanical Engineering TU Berlin

Applied Machine Learning in Engineering

Lecture 00 summer term 2025

Prof. Merten Stender

Cyber-Physical Systems in Mechanical Engineering, Technische Universität Berlin

www.tu.berlin/cpsme
merten.stender@tu-berlin.de

Introduction

01/2023: Professor (W1)
Chair of Cyber-Physical Systems in Mechanical Engineering
Technische Universität Berlin

10/2020: PhD from **Hamburg University of Technology**
Supervisor: Prof. Norbert Hoffmann

PhD Thesis: Data-Driven Techniques for the Nonlinear Dynamics of Mechanical Structures

Imperial College London
University of Technology Sydney
Sandia Laboratories Albuquerque/USA
AUDI AG Ingolstadt

Team CPSME

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

Chair of Cyber-Physical Systems in Mechanical Engineering (CPSME)

Homepage: www.tu.berlin/cpsme

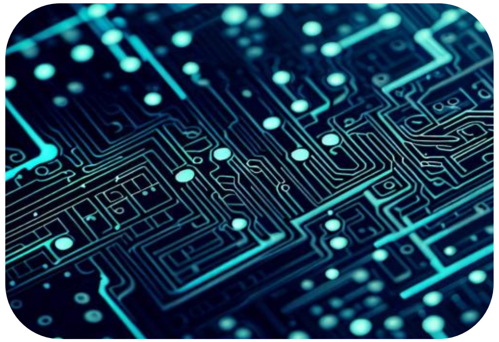
Teaching: https://www.tu.berlin/en/cpsme/study-and-teaching

- **Questions**:
  - 1st contact point: **ISIS forum** https://isis.tu-berlin.de/mod/forum/view.php?id=1987844
  - 2nd contact point: email to Prof. Stender

- **Student thesis projects**:
  1. Read https://www.tu.berlin/en/cpsme/study-and-teaching/final-theses
  2. **STROD** platform (https://www.theses.tu-berlin.de/de/theses) provides some projects
  3. You are invited to bring your own ideas and then get in touch

Teaching

Cyber-Physical Systems
in Mechanical Engineering TU Berlin



Applied Machine Learning in Engineering

Applied Deep Learning in Engineering

Cyber-Physical Systems

Hybrid Simulation Approaches

Research Focus

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

Complex Dynamics

Machine Learning

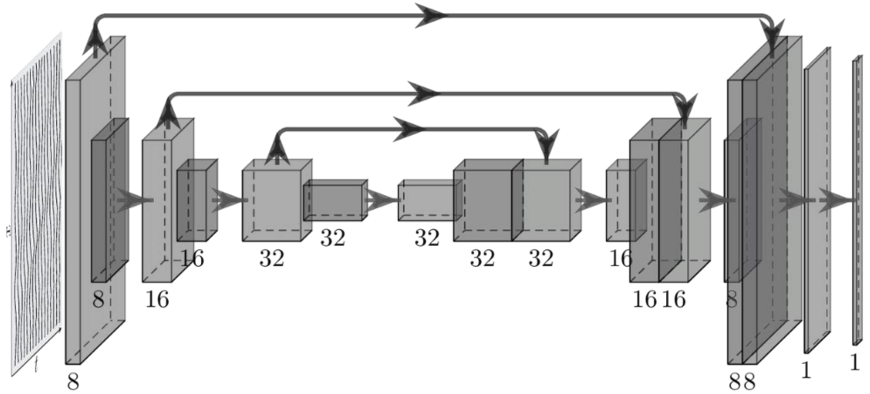
Engineering Systems

Research Profile

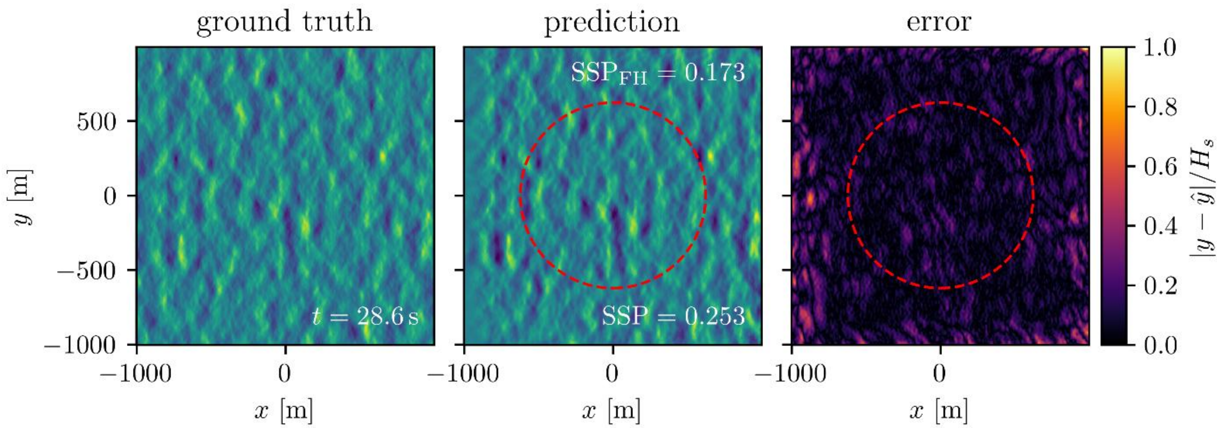
Cyber-Physical Systems
in Mechanical Engineering TU Berlin

- **Nonlinear complex dynamics**

- **Data-driven methods** (machine & deep learning)
  - Hybrid simulation approaches
  - Physics-based simulations (bottom-up)
  - Data-driven simulations (top-down)



- **Applications**
  - Ocean wave prediction
  - Automotive disk brake systems
  - (Offshore) wind energy systems
  - …



Organizational Details

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

- **Lecture**:
  - **Wednesday 10am-12pm CT, 90min, room EB 107**
  - In-person, no video recording available
- **Exercise**:
  - **Tuesday 12pm-02pm CT, 90min, room EB 202**
  - Hands-on coding exercises
  - Please bring your own computer
- **ISIS (LINK)**:
  - Slides, lecture notes and exercise sheets will be published ahead in time
  - **Poll on active course participation**: your choice until April 30th (otherwise: unsubscription)

- **Exam**:
  - Written digital exam **at TU Berlin**. Dates: 21.07.2025 (08:00am); 09.10.2025 (10:00am)
  - Exam registration: ONLY through Moses

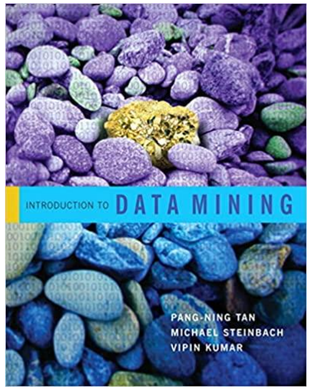
Literature

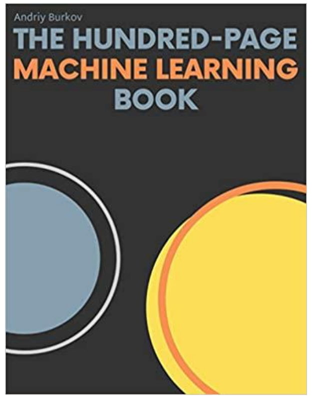
- **Introduction to Data Mining by Tan, Steinbach**
  - ISMB-10: 1292026154

- **The 100-page Machine Learning Book** by Andriy Burkov
  - https://themlbook.com/
  - Wiki http://themlbook.com/wiki/doku.php (read first, buy later)
  - GitHub Repo https://github.com/aburkov/theMLbook
  - ISBN-10 1999579518

- **Neural Networks and Deep Learning** by Michael Nielsen
  - Free access http://neuralnetworksanddeeplearning.com/
  - (only basic neural network part)

- **Several online ressources** (YouTube channels of Andrew Ng and others), online classes and lecturing materials by other universities





Online Programing Ressources

- **Khuyen Tran**: Efficient Python Tricks and Tools for Data Scientists (Link)
- **Arjan Codes**: Software design principles and Python programming (Link)
- **DataCamp** Classroom: coming soon (giving access to DataCamp classes)

Exercise sessions

- **Practical implementation of lecture concepts and methods**
  - Using Python
  - Please bring your own laptop (or collaborate with a peer) and ensure it is fully charged

- **No prior Python experience required**
  - However, regular and consistent practice is essential
  - 1.5 hours per week is not sufficient to develop the necessary skills

- Approach programming like learning a musical instrument:
  Progress requires frequent, hands-on practice

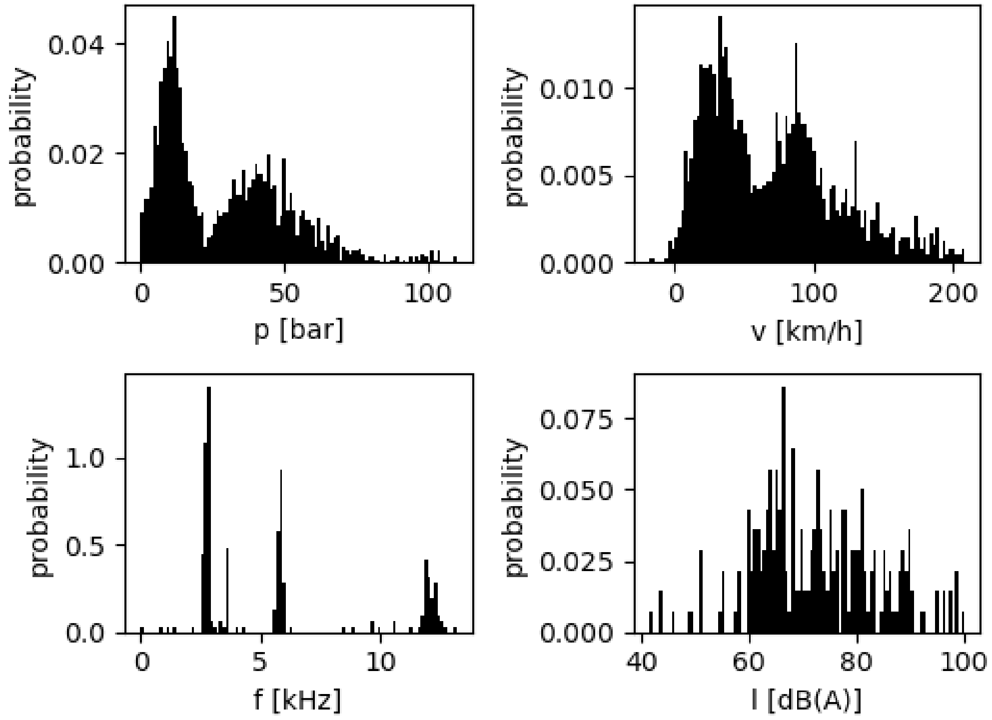
- **The exam will include programming tasks**

Course contents

- **Bits and Bytes:** basics of data, data types, data processing and programming

- **Exploratory data analysis**
  - Statistical characterization
  - Correlation analysis
  - Visualization
  - Outlier detection
  - Dimensionality reduction

- **Unsupervised learning**
  - Clustering algorithms
  - Cluster validity metrics
  - Engineering applications for clustering



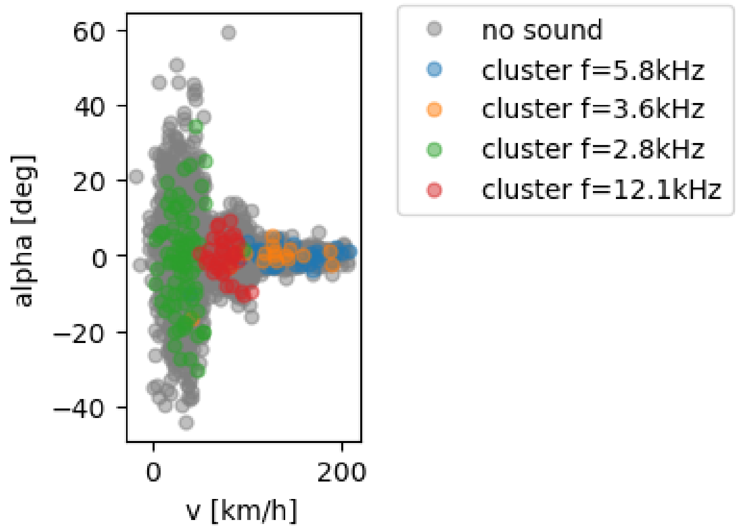
Course contents

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

- **Supervised learning**
  - Classification and regression scenarios
  - Simple (linear) regression
  - Decision trees
  - Feed-forward neural networks

- **Quality and evaluation metrics**
  - Critical model assessment
  - K-fold cross validation

- **Feature engineering**
  - Treating imbalanced data sets
  - Time series analysis



+ Hands-on **Python programming** (weeky exercises)

Terms We Will Cover (selection)

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

Bias-Variance Tradeoff

Gradient Descent

Generalization

Loss function

Data Cleaning

Feature Engineering

Latent Representation

Supervised Learning

Batch size

Overfitting

K-fold cross validation

Feature Extraction

Binary crossentroy

One-Hot Encoding

Random Forest

Clustering

Decision Tree

Questions?

Agenda Lecture 00

- Terms: Artificial intelligence, machine learning and deep learning
- Paradigms for using data-driven methods in engineering
- Various application scenarios in engineering disciplines

Learning outcomes Lecture 00

Learn to …

- differentiate structured and unstructured data.
- recognize supervised and unsupervised learning tasks.
- be data-aware and recognize the hidden value in data.

Know about …

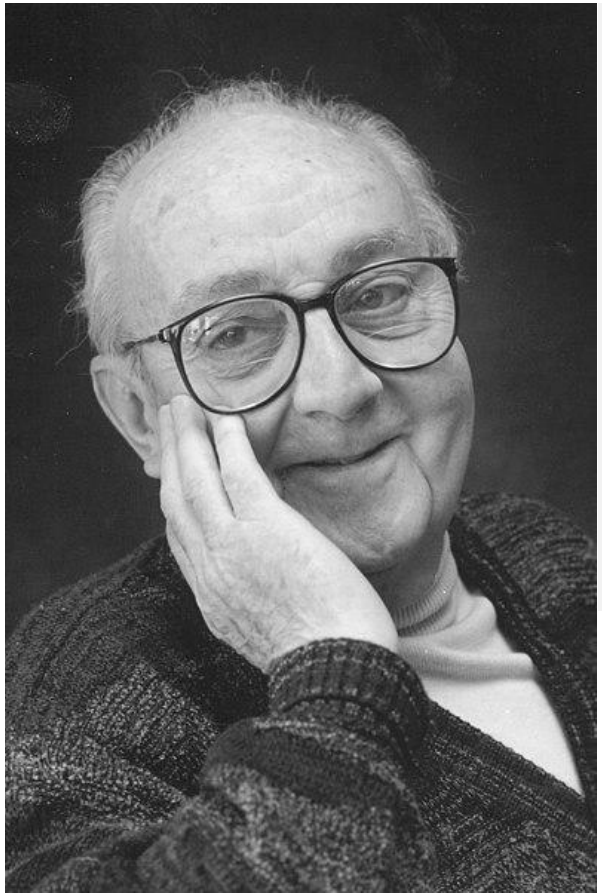
- the terminology of machine learning and deep learning.
- basic and underlying ideas of models with trainable parameters.
- common use cases of machine learning in engineering disciplines.

Machine Learning

“All models are wrong,
but some are useful*.”*

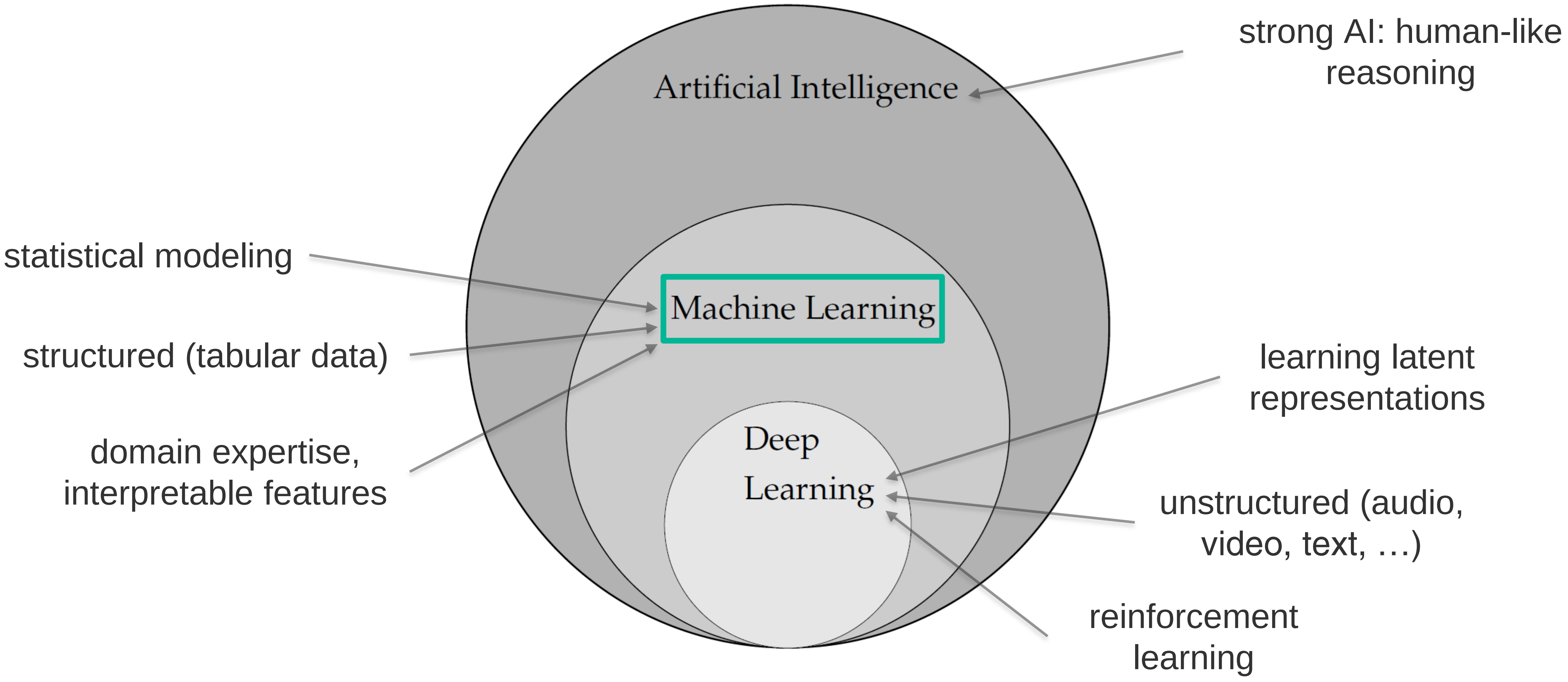
George Box (1919-2013)

1976: "Science and statistics" Journal of the American Statistical Association, 71



wikipedia:
“statistical models always fall short of the complexities of reality but can still be useful nonetheless”

Overview


strong AI: human-like reasoning
Artificial Intelligence
statistical modeling
Machine Learning
structured (tabular data)
domain expertise, interpretable features
learning latent representations
Deep Learning
unstructured (audio, video, text, …)
reinforcement learning

Machine Learning

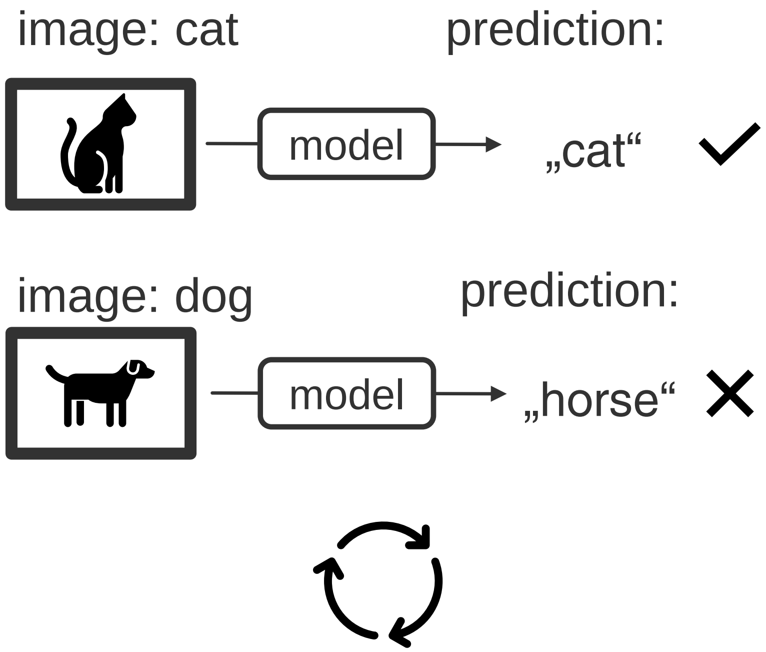
=

maximizing

similarities or dissimilarities through optimization

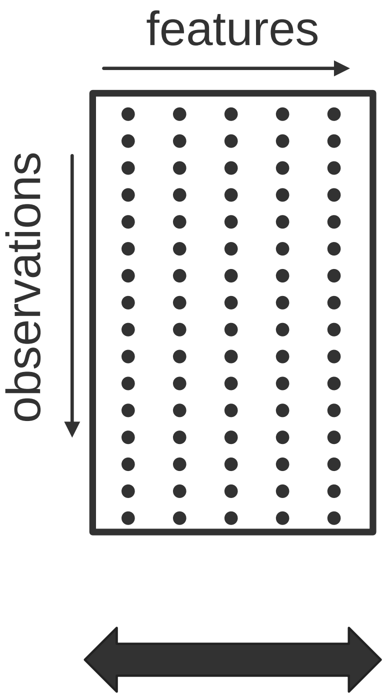
Supervised and Unsupervised Learning

supervised learning
(predictive task)

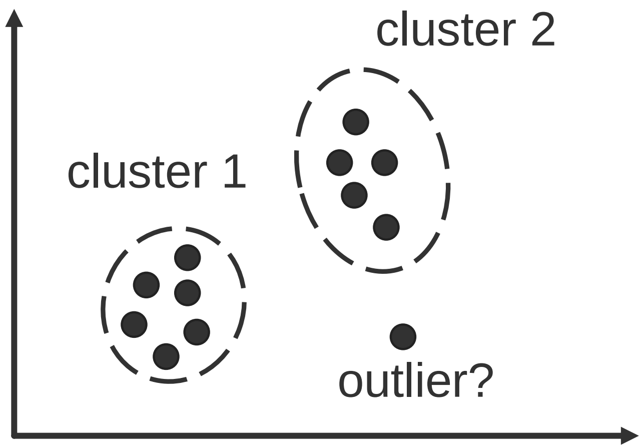


model training = reduce prediction error

data
(tabular)



unsupervised learning
(descriptive task)



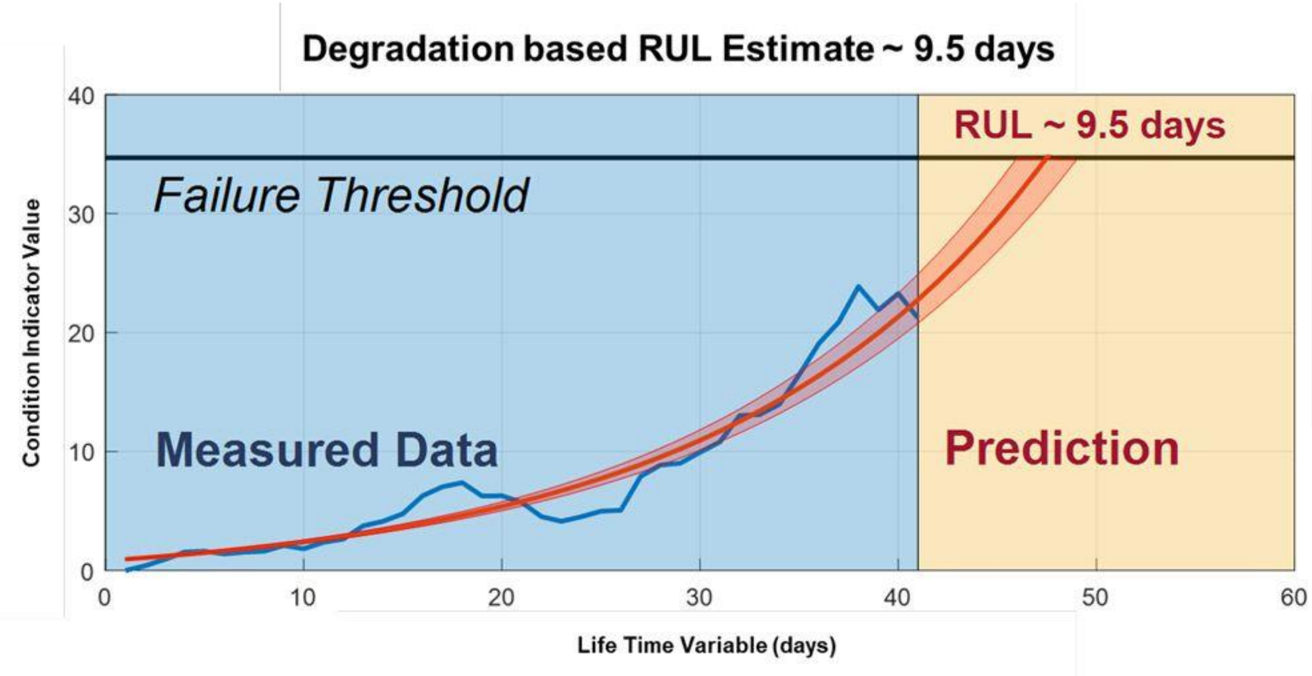
finding clusters, groups and anomalies

Engineering Examples

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

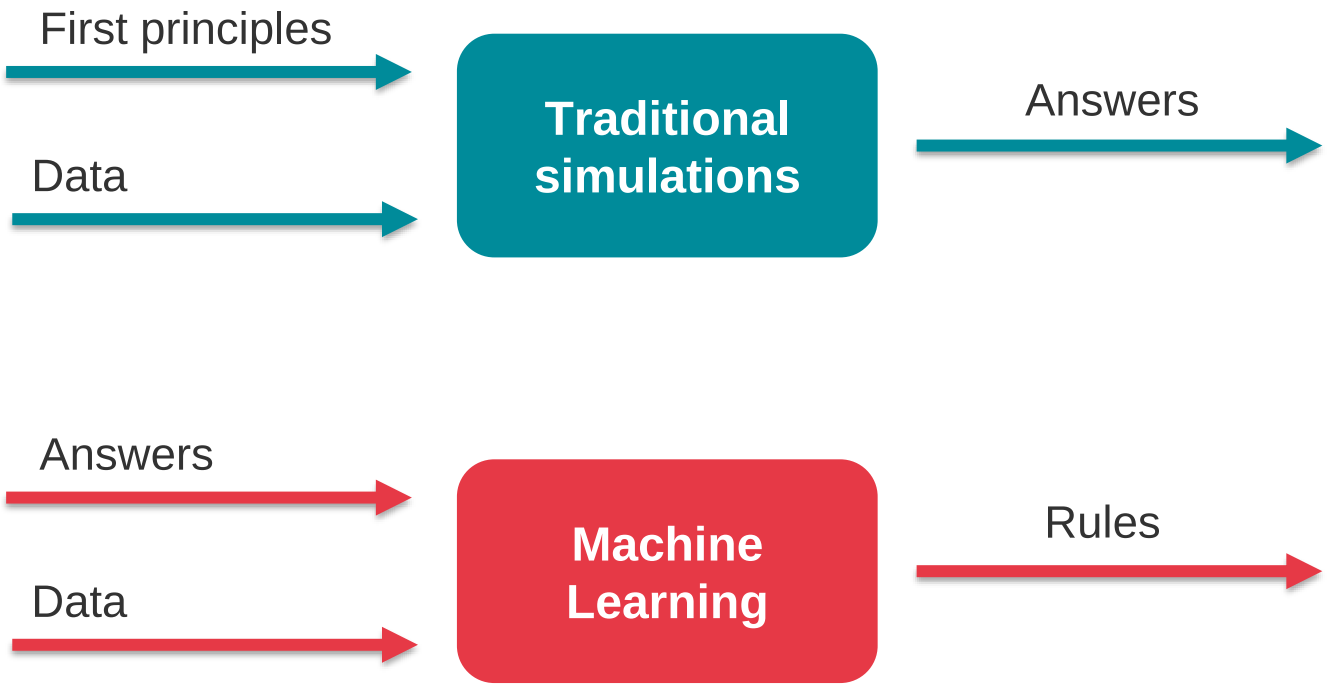
- **Supervised Learning**:
  - Remaining useful lifetime prediction (RUL)
  - End-of-line quality checks (computer vision)

- **Unsupervised Learning**:
  - Grouping customer behavior
  - Outlier detection



Matlab

Learning Methods in a Nutshell



Supervised Learning

- Statistical methods $\mathcal{M}$ for approximating an unknown input ($\mathbf{X}$) – output ($\mathbf{y}$) function based on observations (data samples)

$$\mathcal{M}: \mathbf{X} \mapsto \mathbf{y}$$

- Theoretical basis: **Universal Approximation Theorem**

- Classic example: Boston house price prediction (regression task)

CRIM (per capita crime rate)

NOX (nitric oxides concentration)

DIS (distance to Boston center)

PRRATIO (pupil-teacher ratio)

…

$ Median value of home

Structured Data

- Also denoted as **tabular data**
- **Features** (attributes): quantities that describe measurements or characteristic properties of an individual data sample (record). Example: ambient temperature at specific point in time.

rows: data samples

columns: features
CRIM, NOX, …

Unknown multivariate and nonlinear relationship (to be approximated)

targets

$$\underbrace{\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \cdots & \vdots \\ x_{N,1} & \cdots & x_{N,n} \end{bmatrix}}_{n \rightarrow} \downarrow N \longrightarrow \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

- High-dimensional data sets: n very large
- Big data: N very large (and n potentially, too)

implies

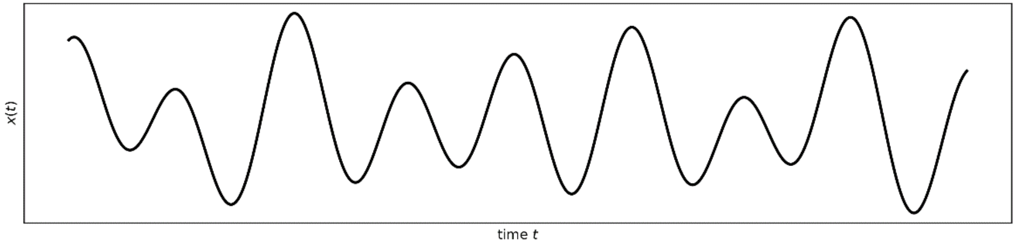
Unstructured Data

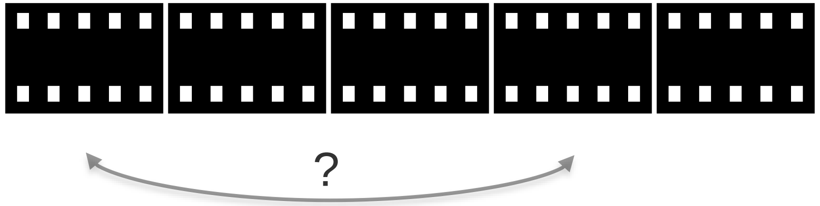
- Also denoted as non-tabular data
- <u>Examples:</u>
  - Text
  - Audio
  - Video
  - Images …

- Special about unstructured data:
  - **Additional latent dimension(s)**
  - **Order matters** (latent dim.)

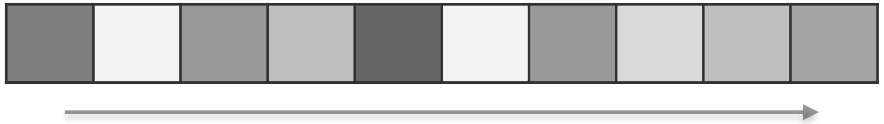
[I | live | in | Munich | but | work | in | Berlin]

?





audio can be stored in an array but is not structured!



latent dimension = time

Order of features not interchangeable!

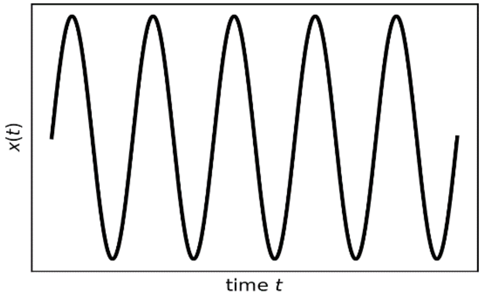
Feature Engineering

- **Machine learning**: structured data
- **Deep learning**: structured + unstructured data
- How to process unstructured engineering data (here: time series data) with ML?

$$[x(t_1) \quad \cdots \quad x(t_n)] \xrightarrow{\text{feature extraction}} [f_1, f_2, \ldots]$$

unstructured data
(time series)

structured data
(vector of features)

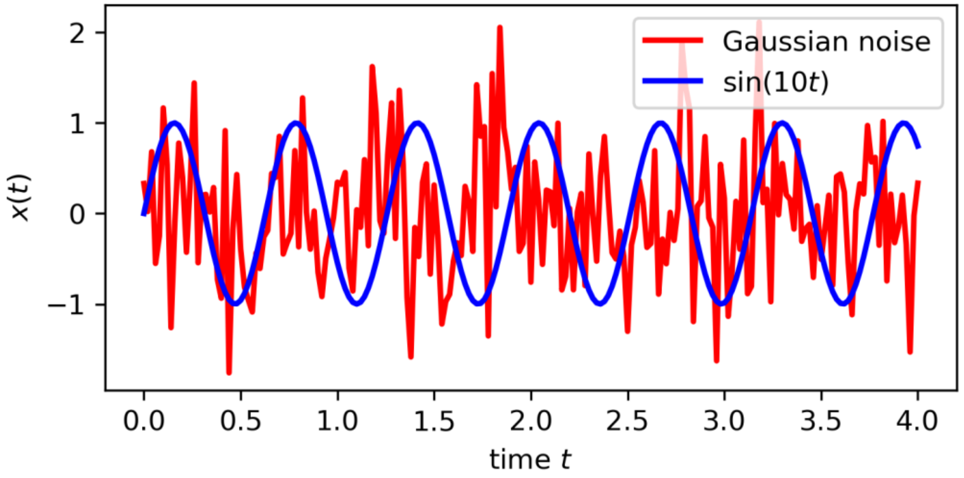


Requires domain expertise!

$$\longrightarrow [\max(x(t)) \quad \text{mean}(x(t))]$$

Example: Signal Classification

- Classify vibration data: a) **harmonic signal** and b) **random (Gaussian) process**
- Obvious to the human eye – how to teach the computer?
- **Without** feature engineering → deep learning required
- **With** feature engineering → simple decision tree



Statistical characterization

Not generic (amplitude / frequency variation?)

```
maximum of noise signal:        1.726001130529702
maximum of sine signal:         0.9999118601072672
mean of noise signal:           -1.7675192431843288e-17
mean of sine signal:            0.04318132362329466
median of noise signal:         0.017201635395270597
median of sine signal:          0.10071709699250761
standard dev. of noise signal:  0.7093145682637918
standard dev. of sine signal:   0.7093145682637918
```

Better: signal processing methods

Fourier analysis, Nonlinear Approaches

Machine Learning in Engineering

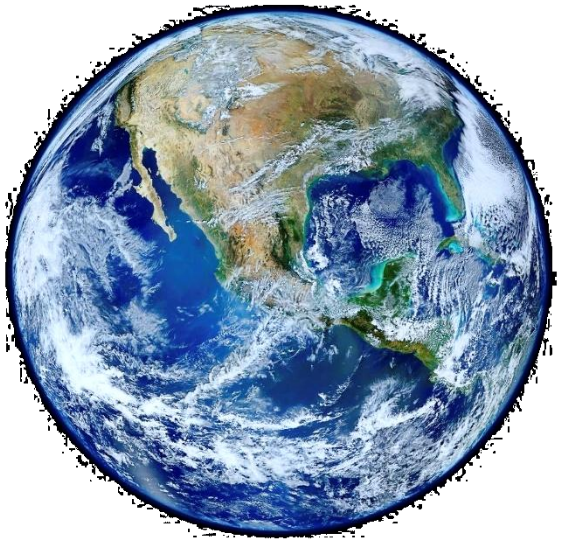
Data-Driven Models

When to use data-driven models (a selection of motives)

- There is **no analytical description available** for the problem
  - Language translation, e.g. English → German
  - …
- **Causal relationships are not** (fully) **understood** today
  - Climate variability, e.g. El Nino
  - …
- **Searching for patterns in huge data sets**
  - Pedestrian detection for autonomous driving
  - …

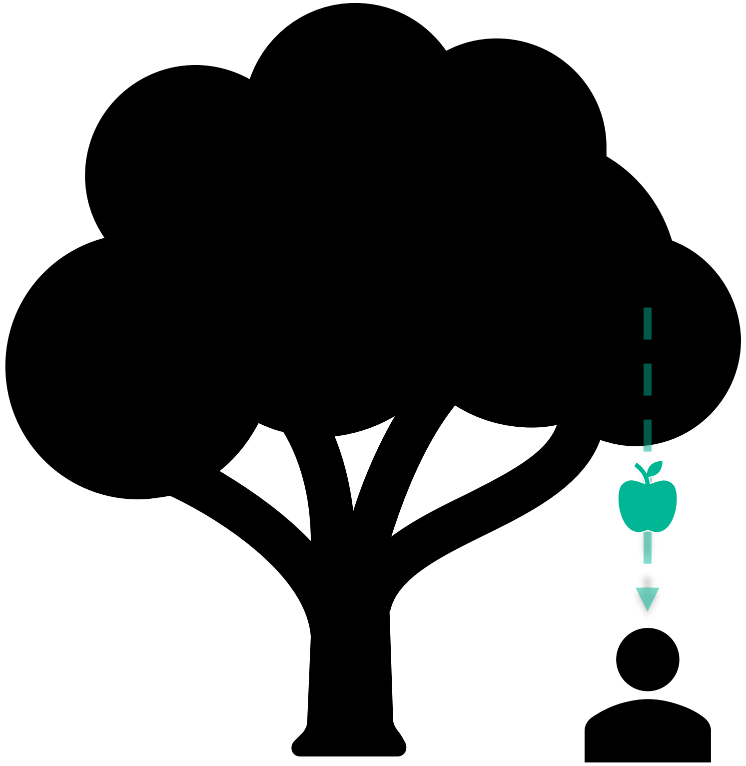
amazon

You may also like …



Data-Driven Models in Engineering

Cyber-Physical Systems
in Mechanical Engineering TU Berlin



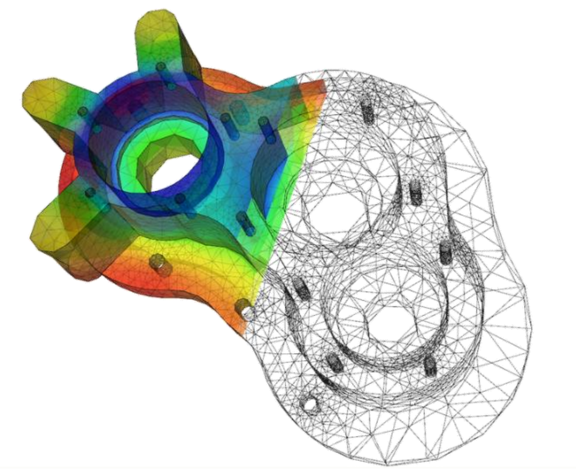
First principles and governing equations

Mathematics	$\mathbf{A}\mathbf{x} = \mathbf{b}$
Mechanics	$F = m \cdot \ddot{x}$
Fluid dynamics	$p + \frac{1}{2}\rho v^2 + \rho g h = \text{const}$
Thermodynamics	$\Delta U = \Delta Q + \Delta W$
Electrical engineering	$P = U \cdot I$

We can compute / simulate everything?

First Principles and Governing Equations

- Centuries of research have resulted in
  - **First principles and governing equations**, e.g. conservation of energy
  - **System description** in terms of ordinary/partial differential equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$
  - **Numerical discretization schemes**, e.g. Finite Element Methods
  - **High-performance computing** and parallel algorithms

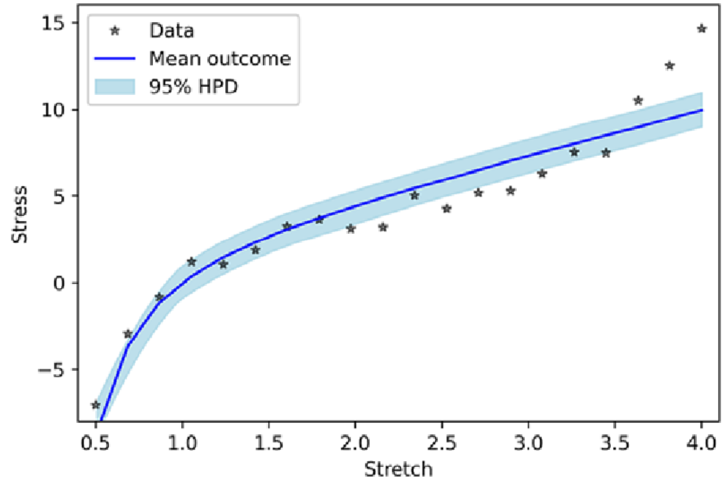


Wikipedia, Creative Commons Attribution

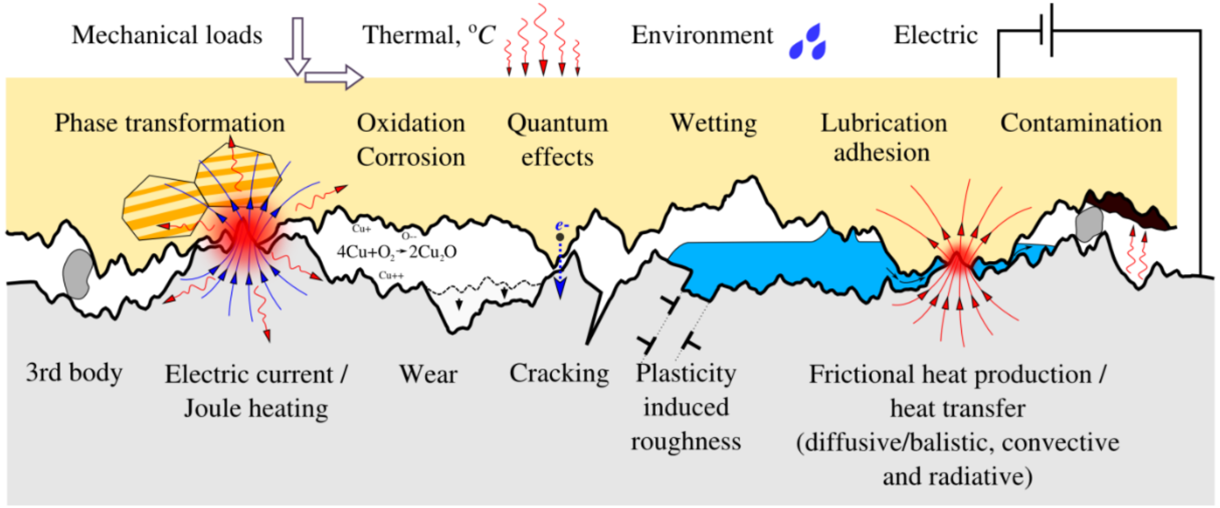
Stuttgart, HRLS

Data-Driven Methods: Use Cases

- **Uncertain model parameters**
  - Structural damping of metals, …
  - Nonlinear behavior (elastomers stiffness, …)
- **Inherent modeling assumptions and limitations**
  - Simplified constitutive models, …
  - Idealized assumptions on homogeneity, …
- **Speed and energy-efficiency**
  - Homogenization of materials
  - Low-order yet fast surrogate models



doi:10.1017/dce.2021.9



doi.org/10.1016/j.triboint.2018.02.005

Machine Learning in Engineering: Examples

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

- **Wind turbines**
  - Gear box defects
  - Rolling bearing faults
  - Crack propagation in blade roots
  - Blade icing
  - Optimal control and energy yield

- **Railways and infrastructure**
  - Fiber optic sensing (newspaper article)
  - Supervision of by-passing trains with video and audio recordings
  - Digital services for operation „Digitale Schiene“

Machine Learning in Engineering: Examples

End-of-line-tests (EOL) for product quality

1. **Object recognition**: what kind of product?
   - Computer vision
   - Image classification
   - Object detection
   - Pose estimation (angles, position)

2. **Failure detection**
   - Cracks, specific failure modes

3. **Factory-level data analysis**
   - Pattern recognition
   - Predictive maintenance

Wikimedia CC

Machine Learning in Engineering: Examples

Autonomous Driving at different levels

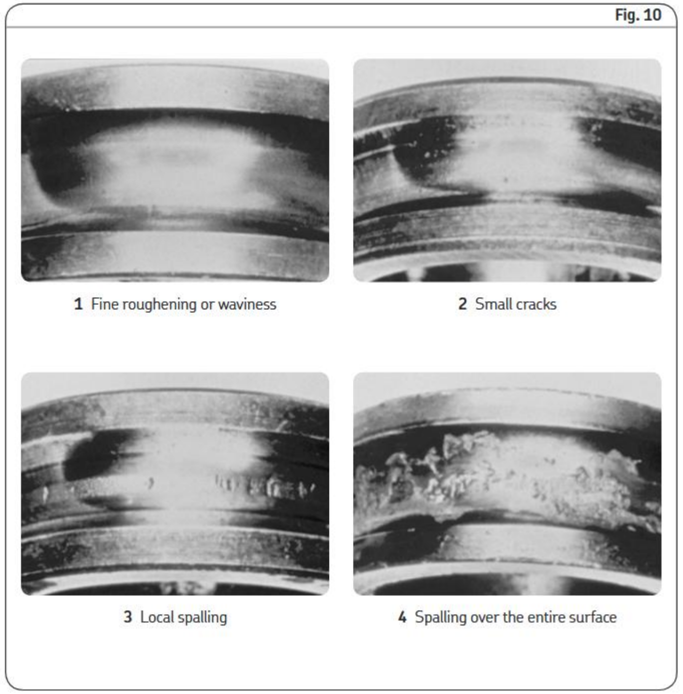
- **Perception**
  - Detection of traffic signs, lanes, obstacles, other traffic participants
  - „looking through walls“ by observing other cars and humans
  - Prediction of future actions by pedestrians and other cars
  - …

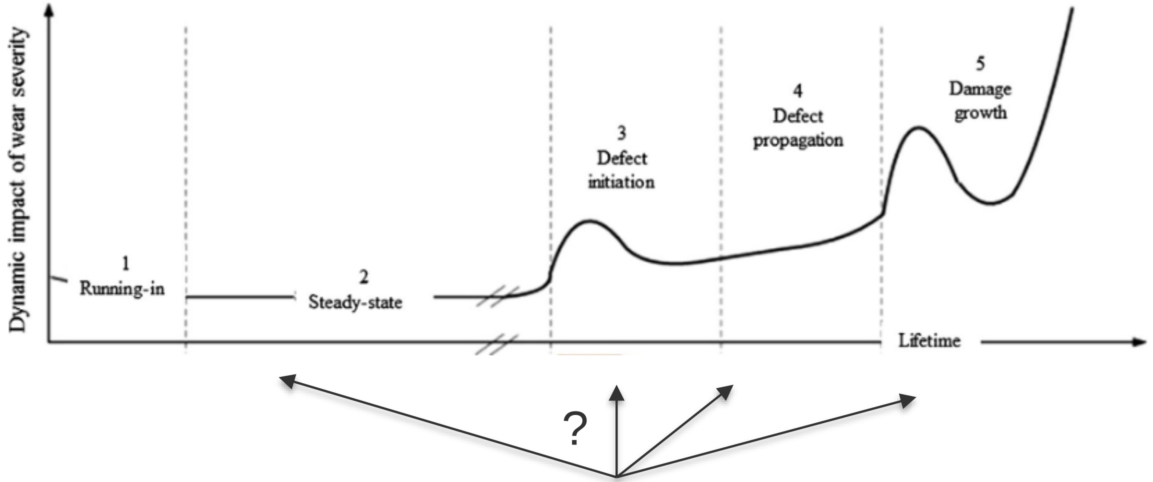
- **Behavior & decision making**
  - Route planning
  - Steering, braking, accelerating
  - …

Machine Learning in Engineering: Examples

Predictive maintenance (PdM), Health Condition Monitoring, Structural Health Monitoring, Remaining Life Time Estimation

Example: rolling bearings





How to estimate the state of the bearing from measurements during operation (variable speed, ambient vibrations, measurement noise)?

SKF® Bearing damage and failure analysis
Cerrada et al. (2018): A review on data-driven fault severity assessment in rolling bearings, Mechanical Systems and Signal Processing 99, 169-196

Machine Learning in Engineering: Examples

- Reinforcement learning (vacuum cleaner robots)
- Precision farming
- Smart building heating/cooling
- Heartbeat signal analysis
- Infrastructure monitoring
- …

©PublicDomainPictures, pixabay

©flo222, pixabay

© PublicDomainPictures, pixabay

7 perspectives on Machine Learning

(1) Supervised Machine learning **is about building prediction models**.

(2) Machine learning **is about learning patterns from data**.

(3) Machine learning **is automated decision-making at scale**.

(4) Machine Learning **is optimization**.

(5) Machine learning **is soft computing**.

(6) Machine learning **is compression**.

(7) Machine learning **is about algorithms producing algorithms**.

by Christoph Molnar

Machine Learning in Engineering: Examples

… and many many more!

The overarching principle is: **data-driven methods are particularly promising and powerful when a handcrafted algorithm is not existent or extremely difficult to formulate.**

How To Study Machine Learning Methods

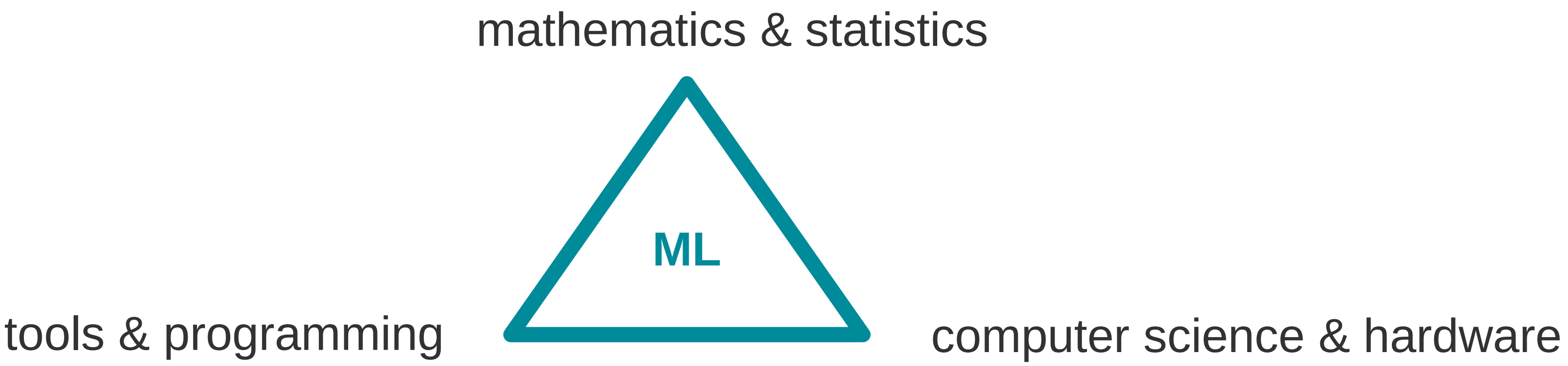
1. Books
2. Online resources (blogs, videos, …)
3. Online classes and mini-degrees: coursera, udacity, udemy, …

UDACITY

coursera

udemy

- **This class:**
    - Basic understanding of algorithms and their limitations
    - Application cases in (mechanical) engineering



Routes to successful ML

Cyber-Physical Systems
in Mechanical Engineering TU Berlin

How to achieve great ML results?

1. **Brute force** (unleash the power of statistics)
   - Large data
   - Large deep learning models
   - Large computers

missing data

CO_2 emissions

2. **Domain expertise**
   - Feature engineering based on domain expertise
   - Careful model selection and optimization
   - Critical assessment of ML results

engineering students

ML knowledge

programming skills

Exercise 00

Exercise 00

- **Basic Python:**
  - Functions
  - Type hints
  - Loops
  - Conditional statements

- **Matplotlib**
  - Plots
  - Axis properties

Questions?