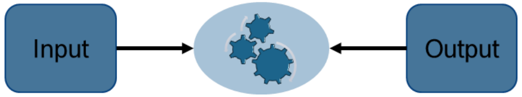
INTRODUCTORY TRAINING TO BIG DATA ANALYTICS AND ARTIFICIAL INTELLIGENCE FOR AE LEADERS

2019 – 05 – 06

BOSCH

Introduction
Learning Objectives

Depositphotos, Bosch License



BD & AI:
What and why?

- Basic definitions
- Why AI and why now?
- Real world applications
- Bosch Use Cases

Data as an asset:
Value creation

- Relevance of data
- Data mining
- Big Data handling

Technical foundations:
Machine Learning

- What is machine learning?
- Strategies and tasks
- Method focus on Deep Learning overview

Implement:
Data mining in your BU

- Value of data
- Data-driven mindset
- Roles and competences
- Trainings and platforms
- Discuss & exchange

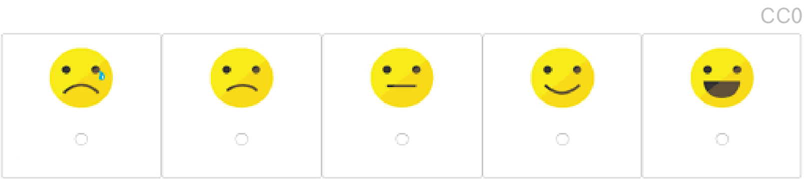
BOSCH

Introduction
Round of introduction – 30 second challenge

1. Who are **you** and where do you work?

2. What is your **expectation** of this day?

3. What is your **connection to AI** and how do you feel about AI on a **scale from 1 to 5**?



WHAT IS ARTIFICIAL INTELLIGENCE?

What is Artificial Intelligence?
Basic Definition

Artificial intelligence is a subfield of computer science that models intelligent behaviour.

What is Artificial Intelligence?
Definitions of AI

AI addresses problems which can be solved by humans, but for which we don’t have good algorithms for solving.

AI is a moving target; once a task solved by algorithms, it’s no longer perceived as AI!

Defining artificial intelligence isn’t just difficult; it’s impossible, not the least because we don’t really understand human intelligence.

Artificial intelligence is the science of making machines do things that would require intelligence if done by men.

What is Artificial Intelligence?
Is Artificial Intelligence possible?

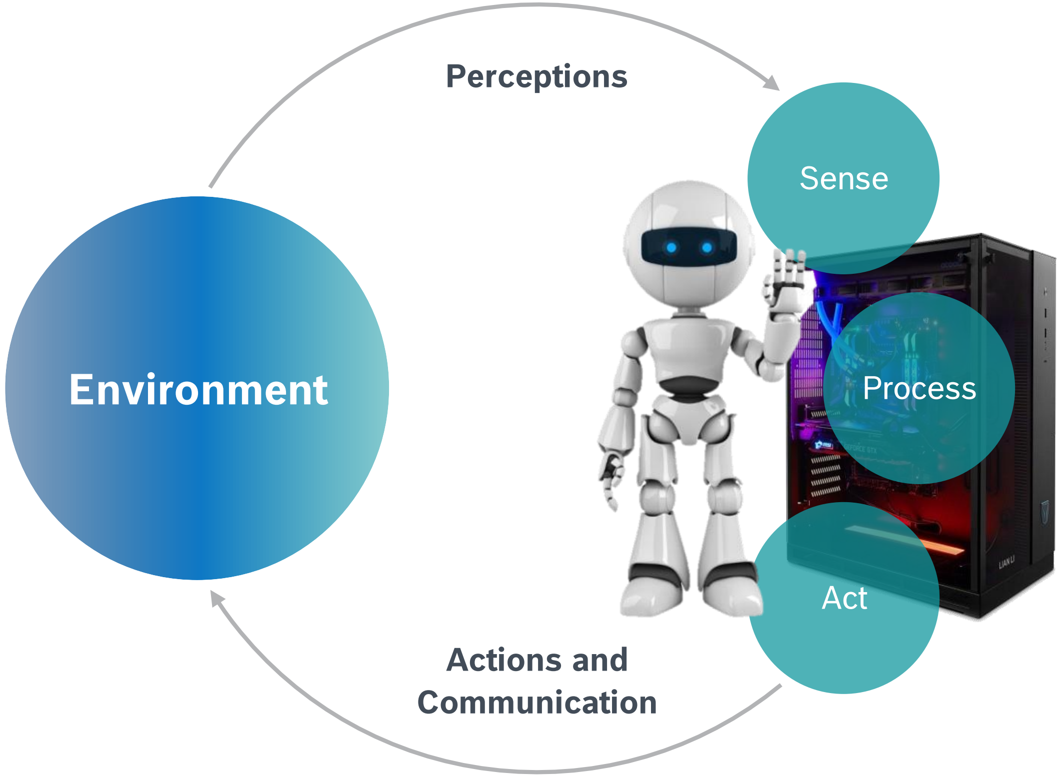
Dartmouth Conference (1956):

“Every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.”

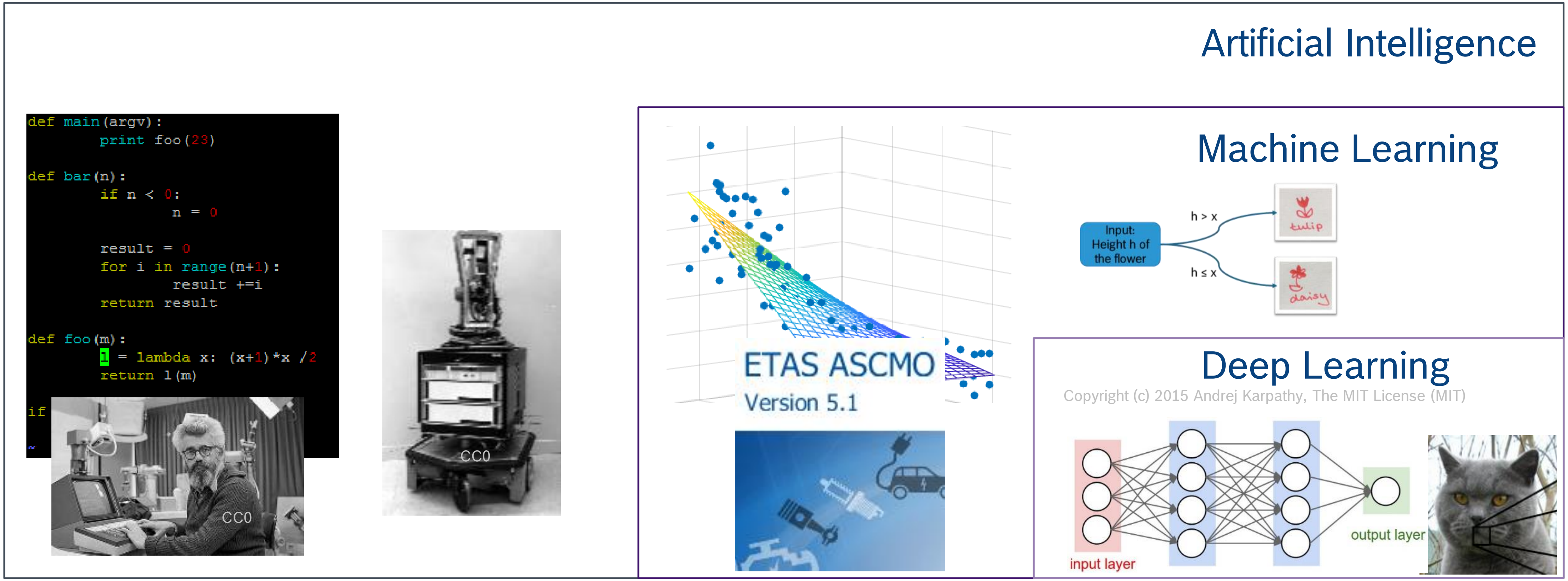
What is Artificial Intelligence?
What is necessary for Artificial Intelligence?

Which skills make humans intelligent?

Which tasks do we have to model to make machines behave intelligently?

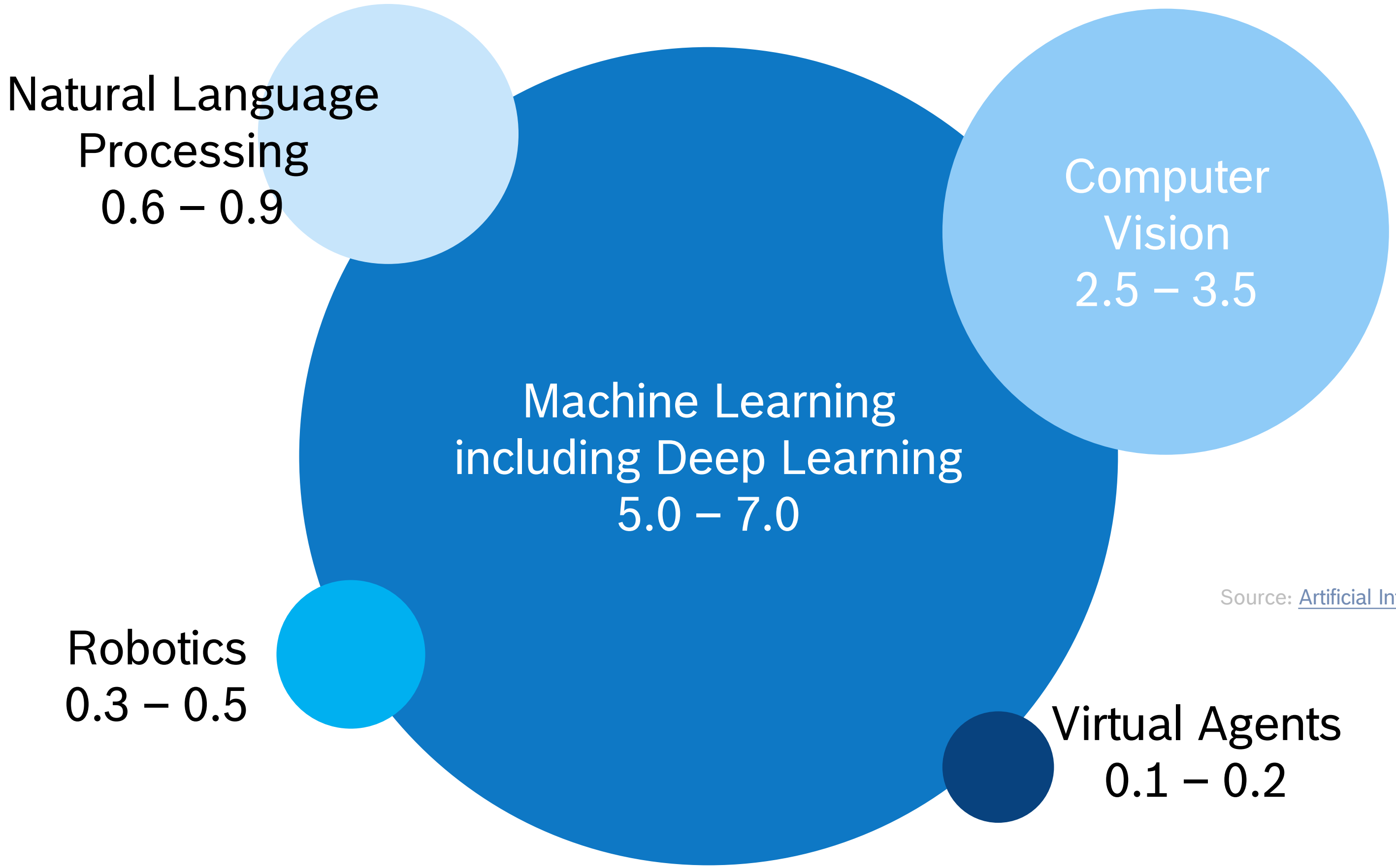


What is Artificial Intelligence?
A Brief Look at History



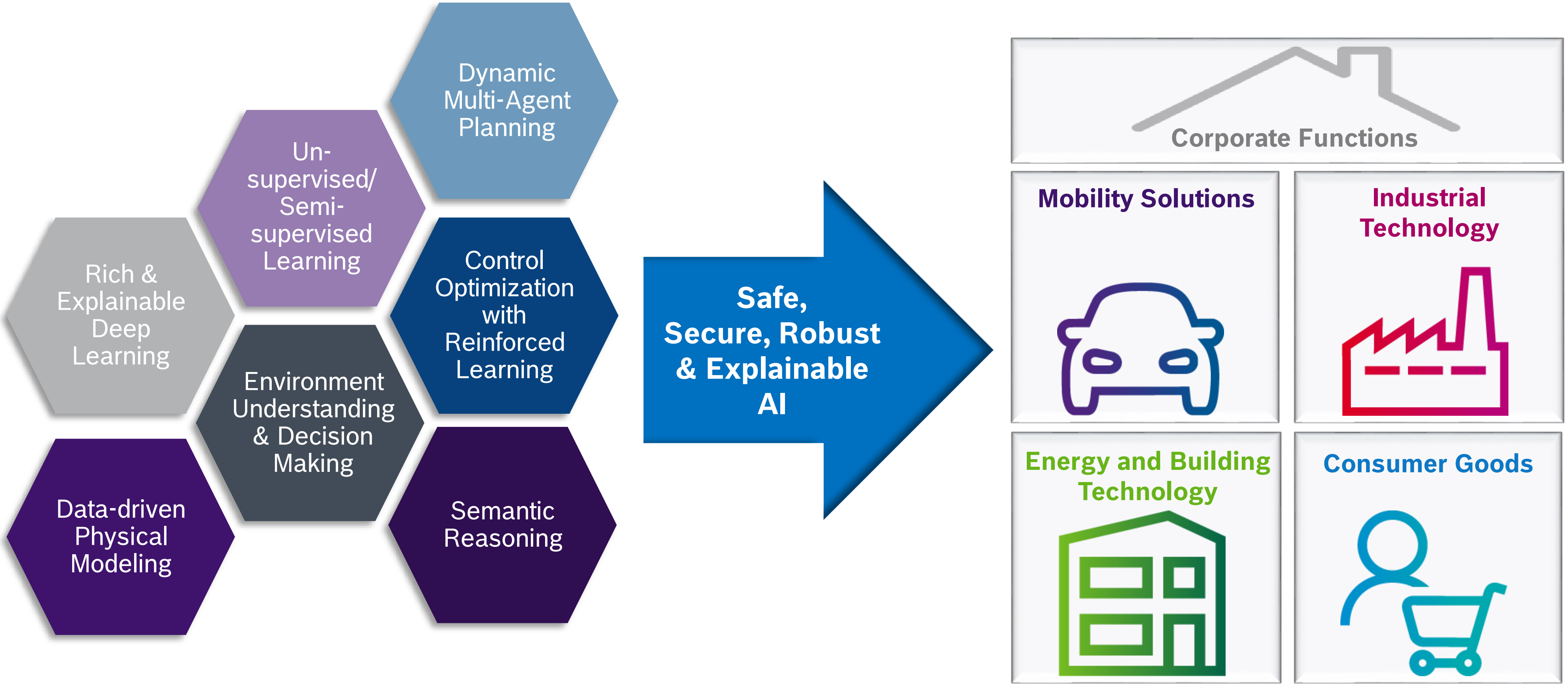
1950s | 1960s | 1970s | 1980s | 1990s | 2000s | 2010s | **2016**

What is Artificial Intelligence?
External Investment in AI focused Companies 2016 in Billion $



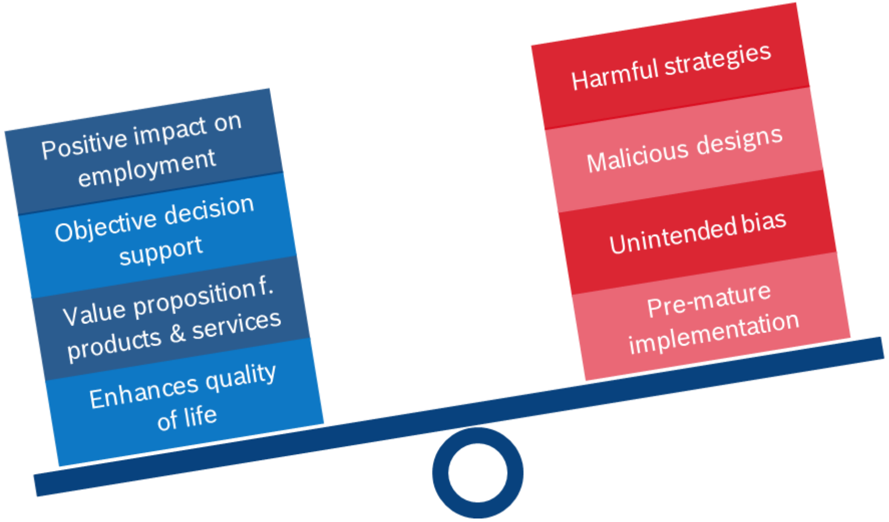
Source: Artificial Intelligence, the next digital frontier, McKinsey Global

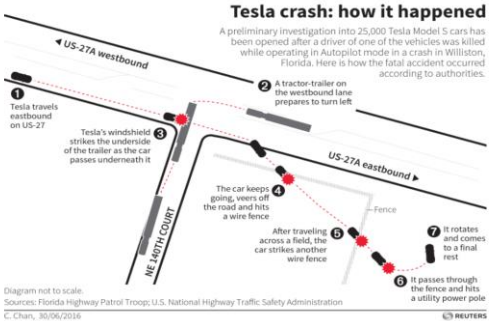
What is Artificial Intelligence?
Current Differentiating BCAI Research Fields



What is Artificial Intelligence?

Artificial Intelligence and Impact on Society





Bosch responsibility:

- ❖ Development and implementation of safe, secure and robust AI
- ❖ Communicate AI as technology to support “invented for life”

Internal | CR/PJ-AI-E & CI/OST | 2019

BOSCH

What is Artificial Intelligence?
Levels of artificial intelligence

Narrow AI

Computed intelligence that matches or exceeds human level in specific tasks

General AI

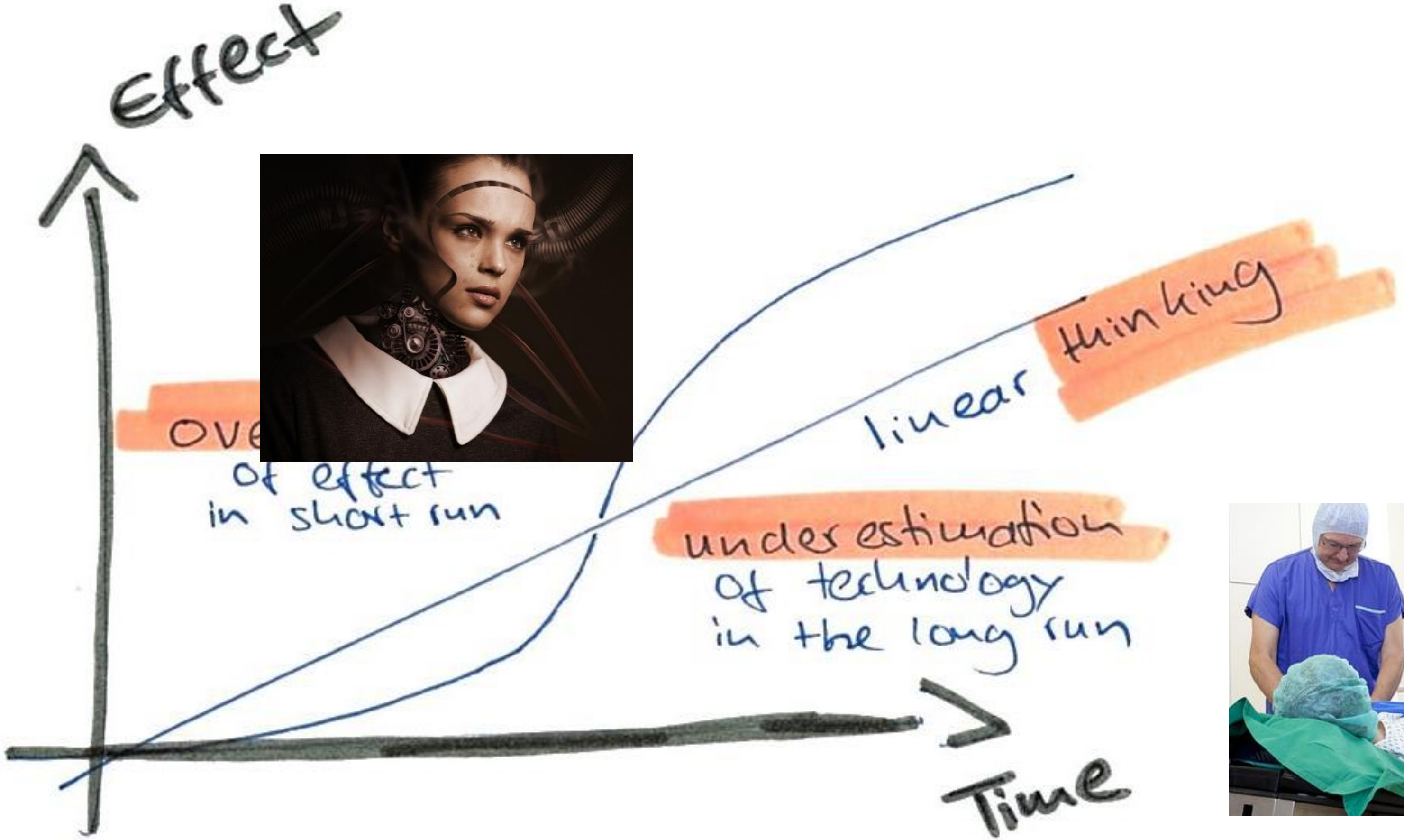
A machine with general human-level intelligence

AI surpassing humans

(Collectively) smarter than humankind on every challenge – singularity, superintelligence, ...

AI TODAY

AI today
Trend over- and underestimation



 Internal | CR/PJ-AI-E & CI/OST | 2019

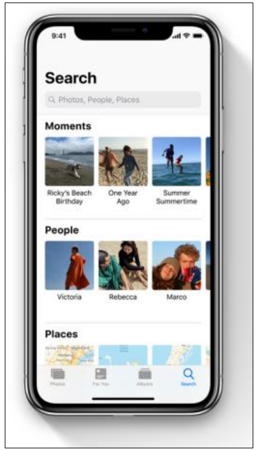
BOSCH

AI today
AI breakthroughs

2005
Autonomous cars manage to drive 213 kilometers at the DARPA Challenge

2016
Deepmind software AlphaGo defeats one of the world's best players in the game GO

Speech recognition



Face recognition and photo search

breakthroughs in research & development

ubiquitous AI applications

AI today
Narrow AI vs experts

"I believe human intuition is too advanced for AI."

"I think (AlphaGo's) level doesn't match mine."

Final Score:
Alpha Go 4 – 1 Lee Sedol

AI defeats elite doctors in diagnosis competition

China Daily/Asia News Network / 03:14 PM July 02, 2018

Radiologist Zhang Junhai from Shanghai Huashan Hospital reads a medical image display during a competition with BioMind, an artificial intelligence system, in Beijing on Saturday. CHINA DAILY

BEIJING — An artificial intelligence system recorded a 2-0 victory against

00:00 Click to hear the article ...

LATEST STORIES | MOST READ

NEWSINFO
Go denies getting special treatment at Comelec
OCTOBER 17, 2018 02:15 PM

SPORTS
Warriors get rings, then beat Thunder on NBA opening night
OCTOBER 17, 2018 02:12 PM

POP
Revamp, rediscover, and recharge with ChELTS 2018
OCTOBER 17, 2018 02:10 PM

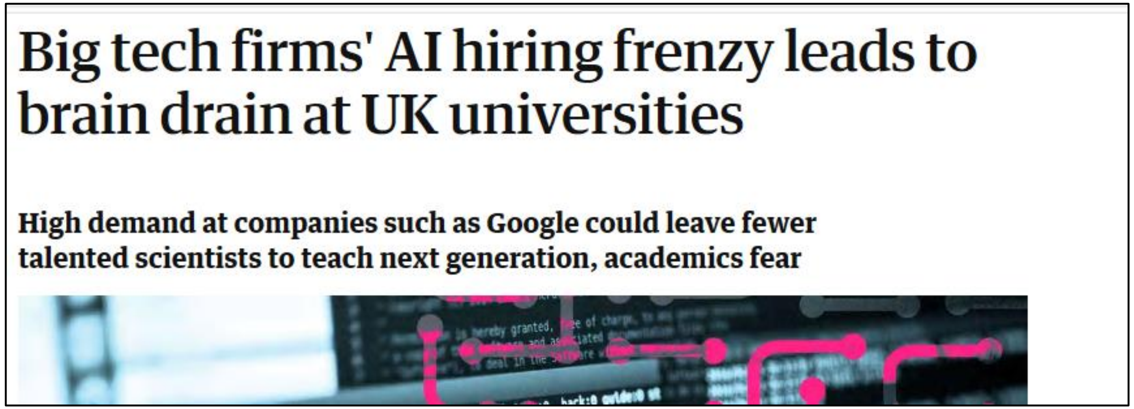
NEWSINFO
CHEd: No suppression of anti-gov't sentiments in universities
OCTOBER 17, 2018 02:10 PM

From: inquirer.net

Nature Feb 2017

Roles, Training and Consulting
AI researcher is the new soccer player

Skysports.com


Big tech firms' AI hiring frenzy leads to brain drain at UK universities

High demand at companies such as Google could leave fewer talented scientists to teach next generation, academics fear

The Guardian

- **AI** talent is scarce
- Experts expect another intensification in the „war on talents“ in the AI field
- Median income at Facebook 240.000 USD (Economist)

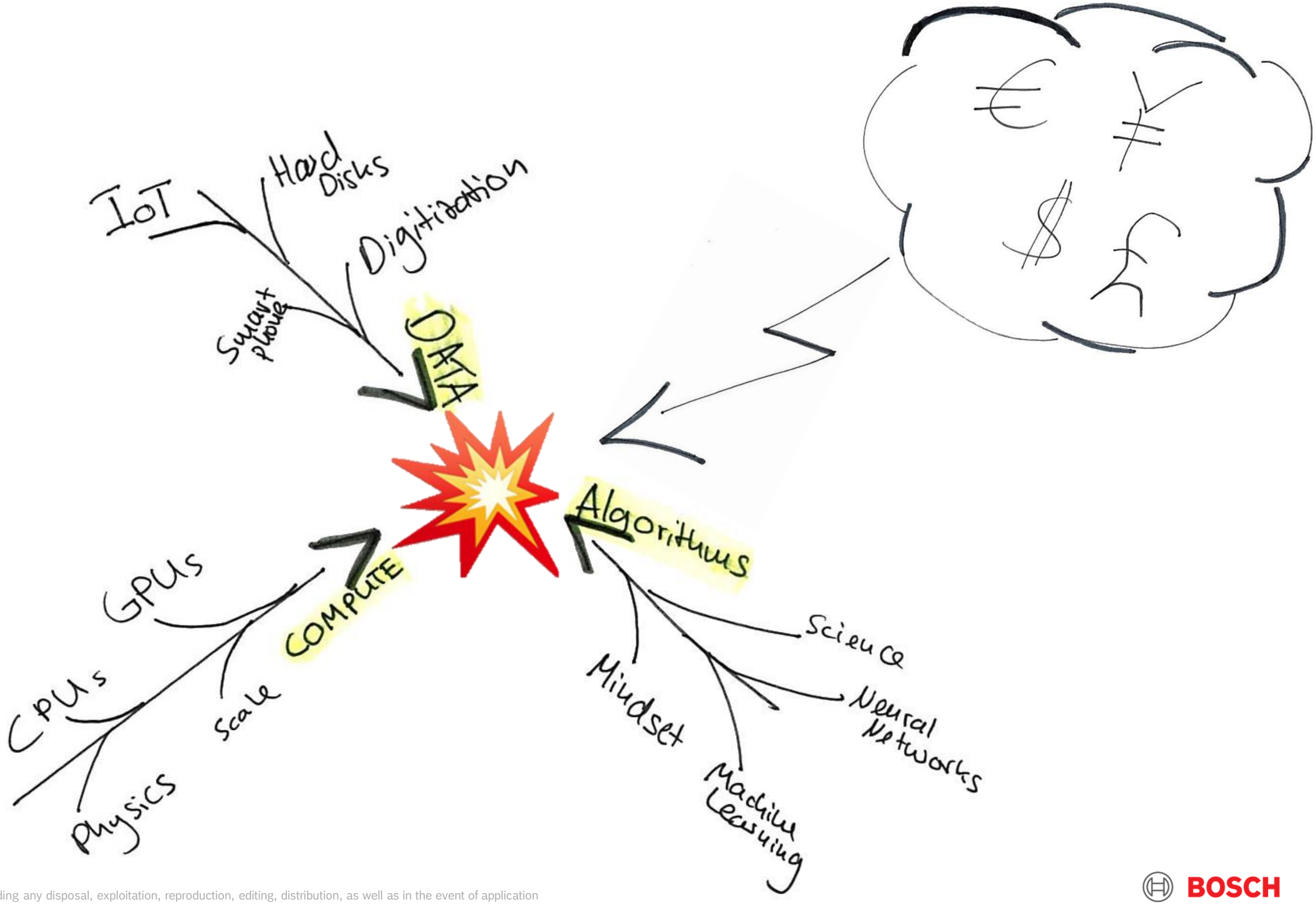
BOSCH

AI today
Rush on academic conferences

Sakyasingha Dasgupta (@DSakya)
05.09.18, 02:31
It's ridiculous that an academic conference (or atleast use to be) #NIPS2018 registrations sold out in < 12 min this year!

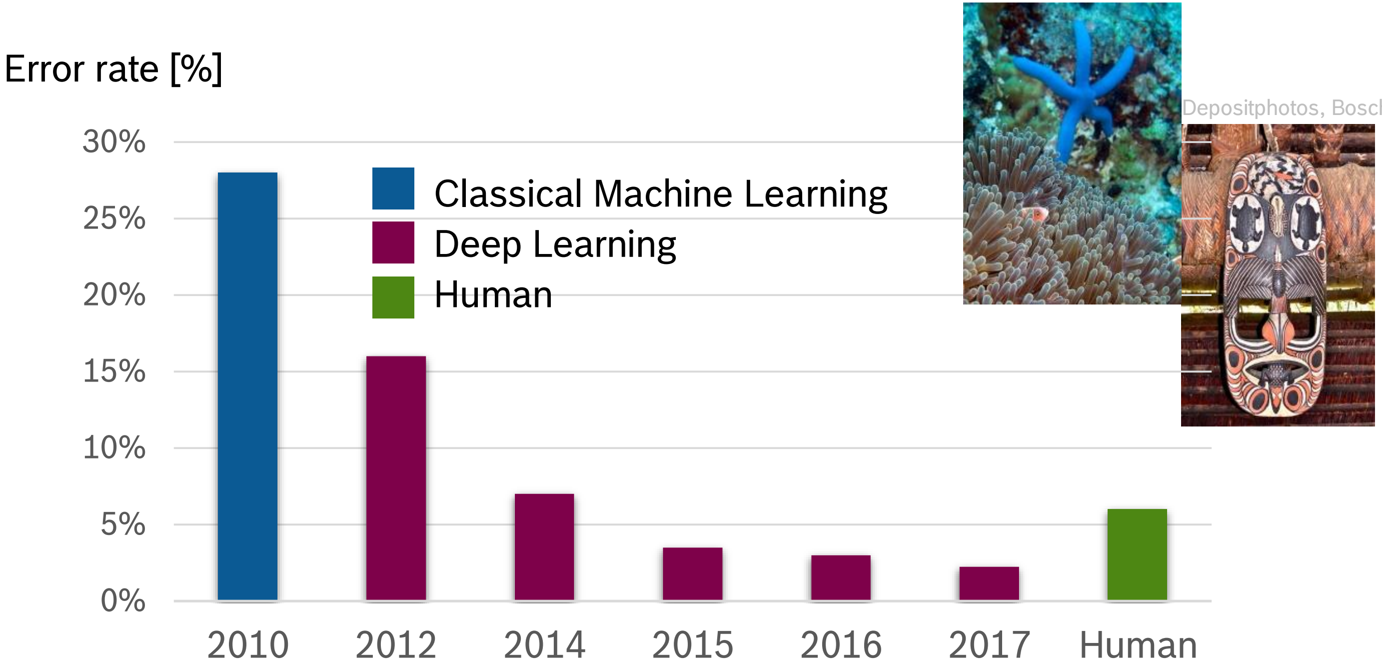
NIPS: Conference on Neural Information Processing Systems

AI today
Why now?



AI today
Why now?

1. Algorithmic advances, especially in the area of deep learning



Depositphotos, Bosch License

➢ Error rate of image recognition “imagenet challenge”. In 2015, algorithms surpassed human image recognition error rate.

BOSCH

AI today
Why now?

1. Algorithmic advances, especially in the area of deep learning

From https://medium.freecodecamp.org/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d

AI today
Why now?

2. Tremendous growth of data

Luca Bruno, AP

2005

Michael Sohn, AP

2013

BOSCH

AI today
Why now?

3. Technological advancements lead to cheaper computing power, storage and memory

Wikipedia.org, CC BY-SA 3.0 fr

=

=

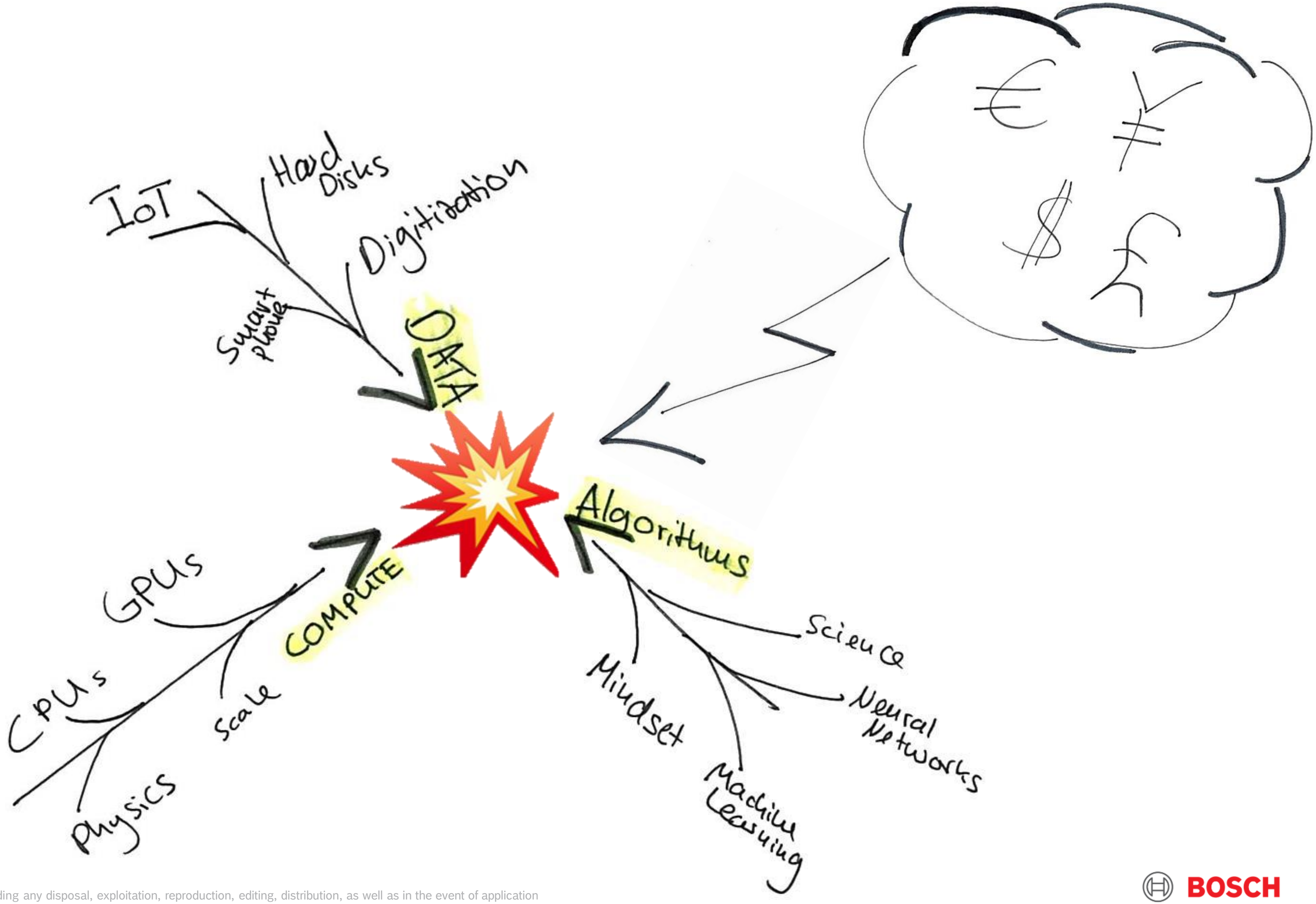
1985, Cray-2
~ 32.000.000 $

2010, iPhone 4
~ 600 $

2016, Apple Watch
~ 300 $

BOSCH

AI today
Why now?



DATA

Data
Why do we need data?

“Big data is to AI what fuel is to engines: If the tank is empty, the engine comes to a halt.”

Max Welling (Amsterdam Machine Learning Lab)

https://www.bosch.com/de/explore-and-experience/thought-leader-max-welling/

Data
Types of data acquisition

no
data acquisition

passive
data acquisition
(digitization)

active
data acquisition
(business case driven)

strategic
data acquisition
(foresight based)

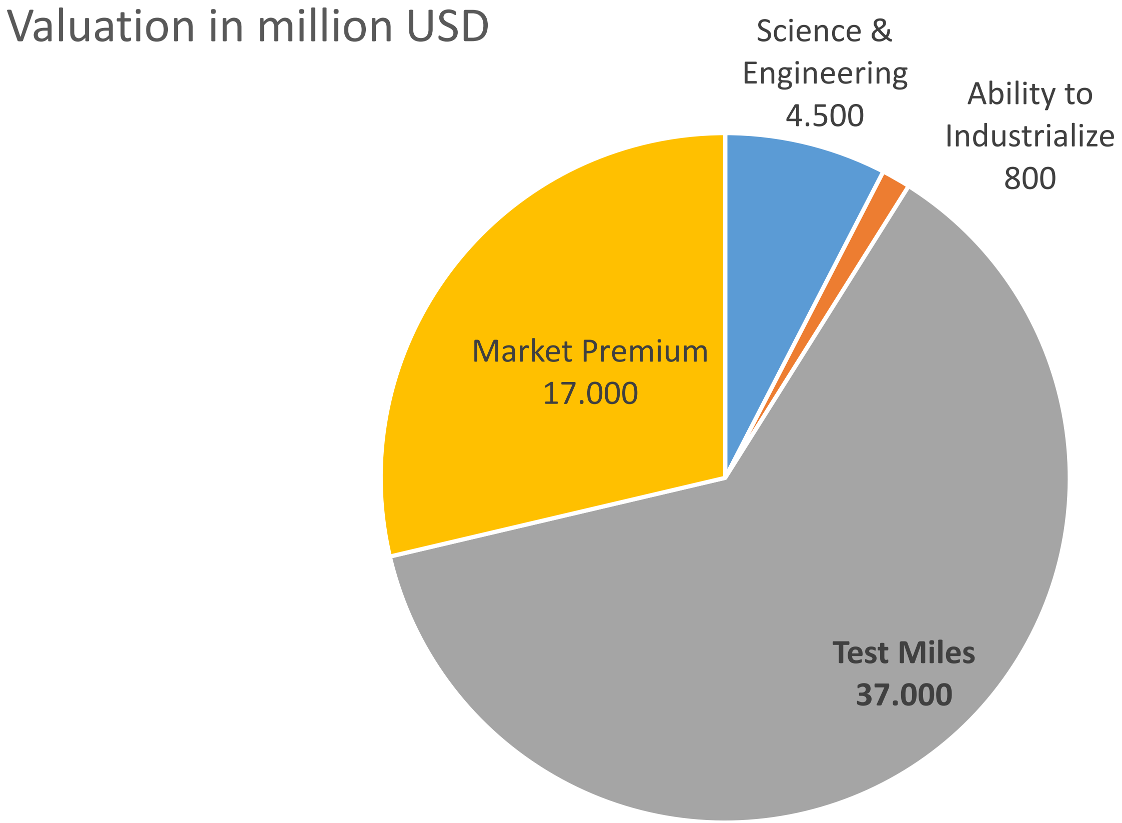
Data
Data at Bosch

Which data do you already or could you use for your business?

Data
Example: Big Data at Waymo

- 10M miles driven (14M km)
- 7B miles in simulator
- Currently 40.000 km of test drives daily
- Hand-over to driver every 8.000 km
- 82.000 vehicles ordered for robo cab fleet
  (Total number of taxi drivers in the USA: 233.000)

Data
Example: Big Data at Waymo



Data Based Value Creation
Data Mining

Depositphotos, Bosch License

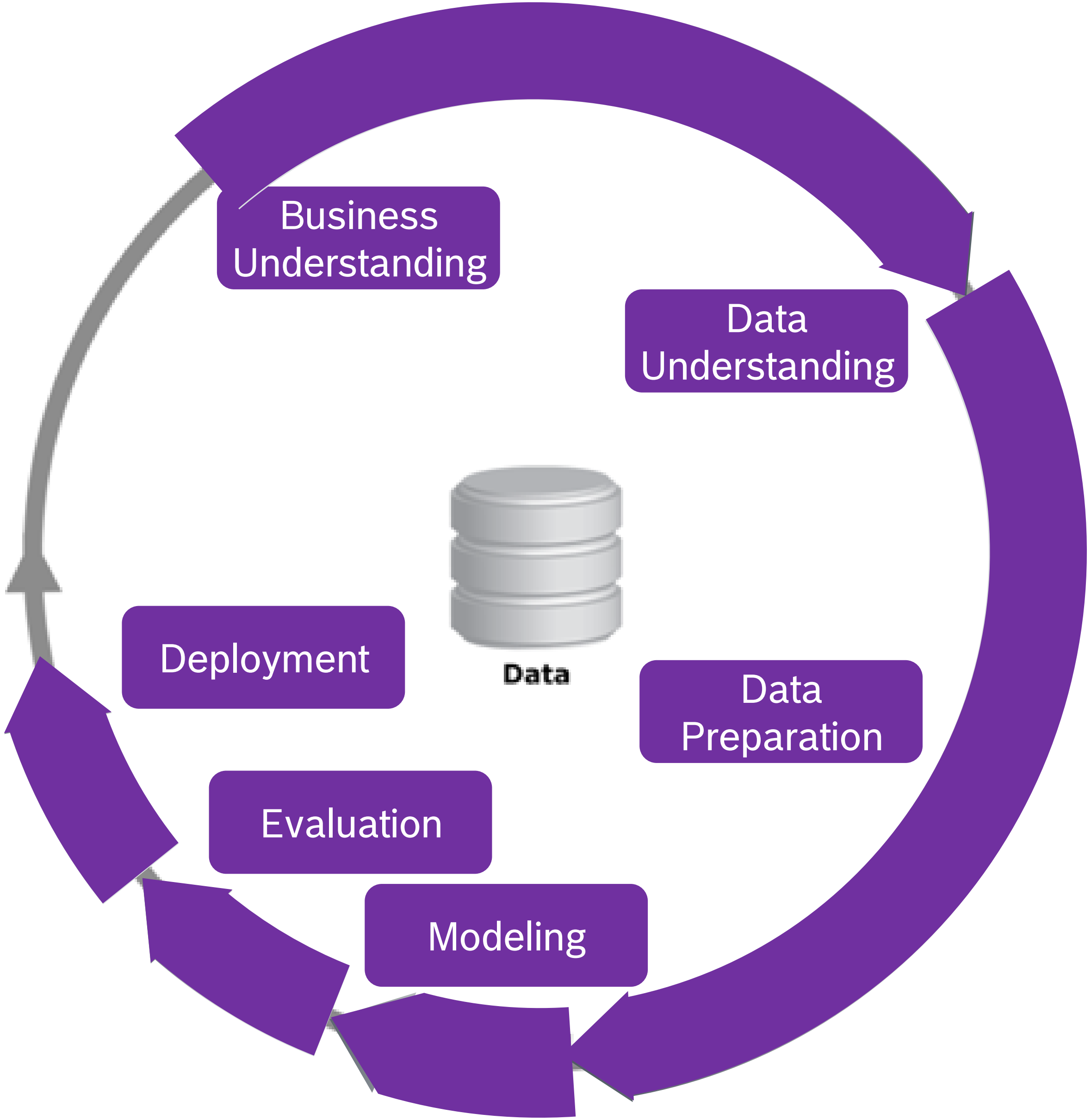
Data mining is the process of discovering hidden, previously unknown and usable information from a large amount of data.

(ISO/IEC JTC1/SC32 WG4 SQL/MM Part 6 WD, 2000)

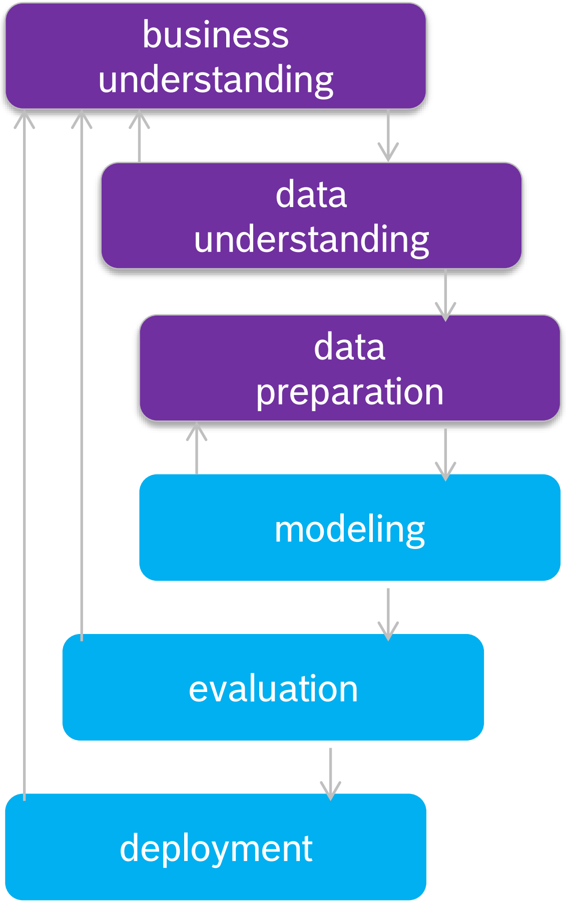
Data Based Value Creation
CRISP-DM

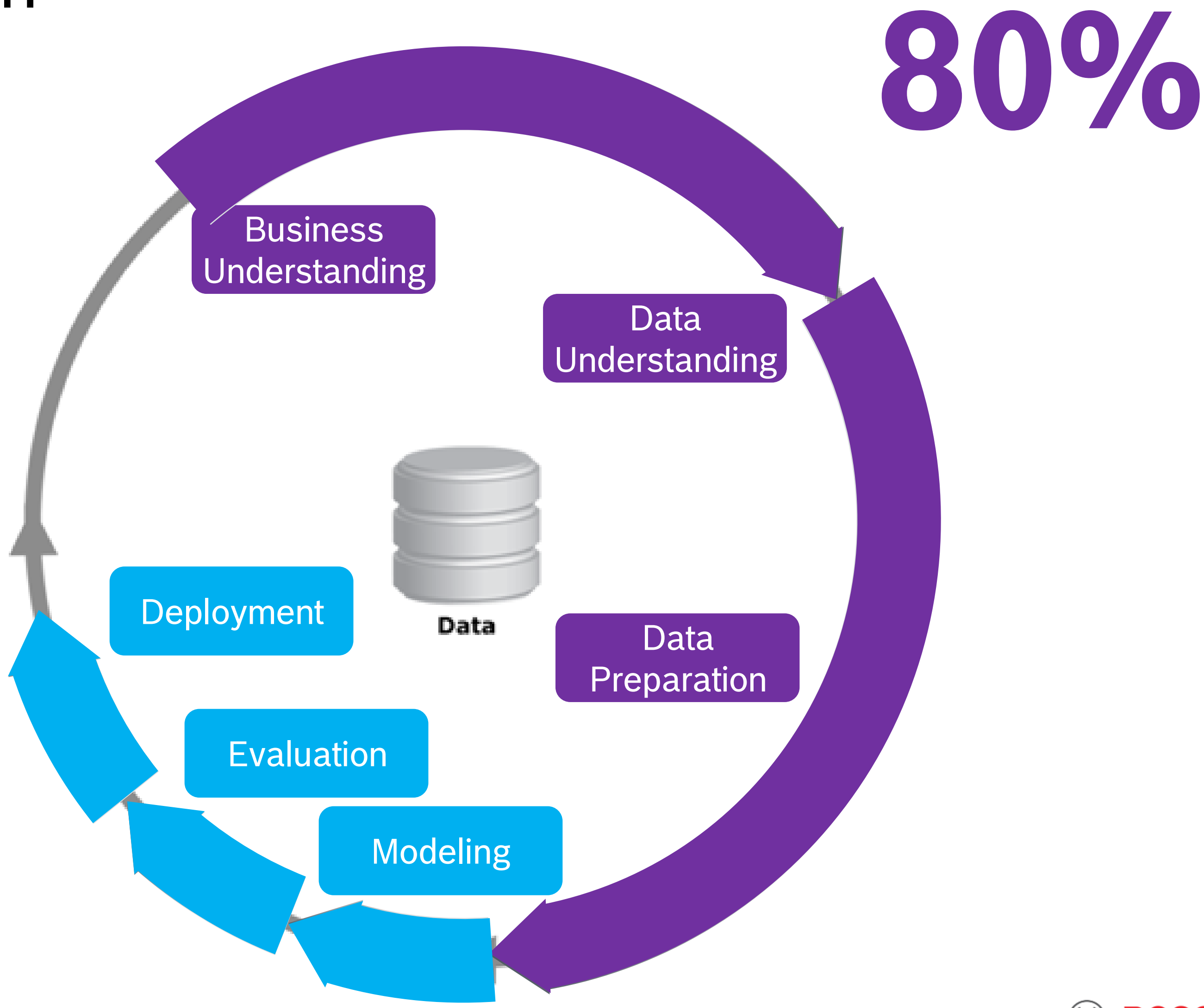
The standard data based value creation process is called CRISP-DM:

Cross-Industry
Standard Process
for Data Mining

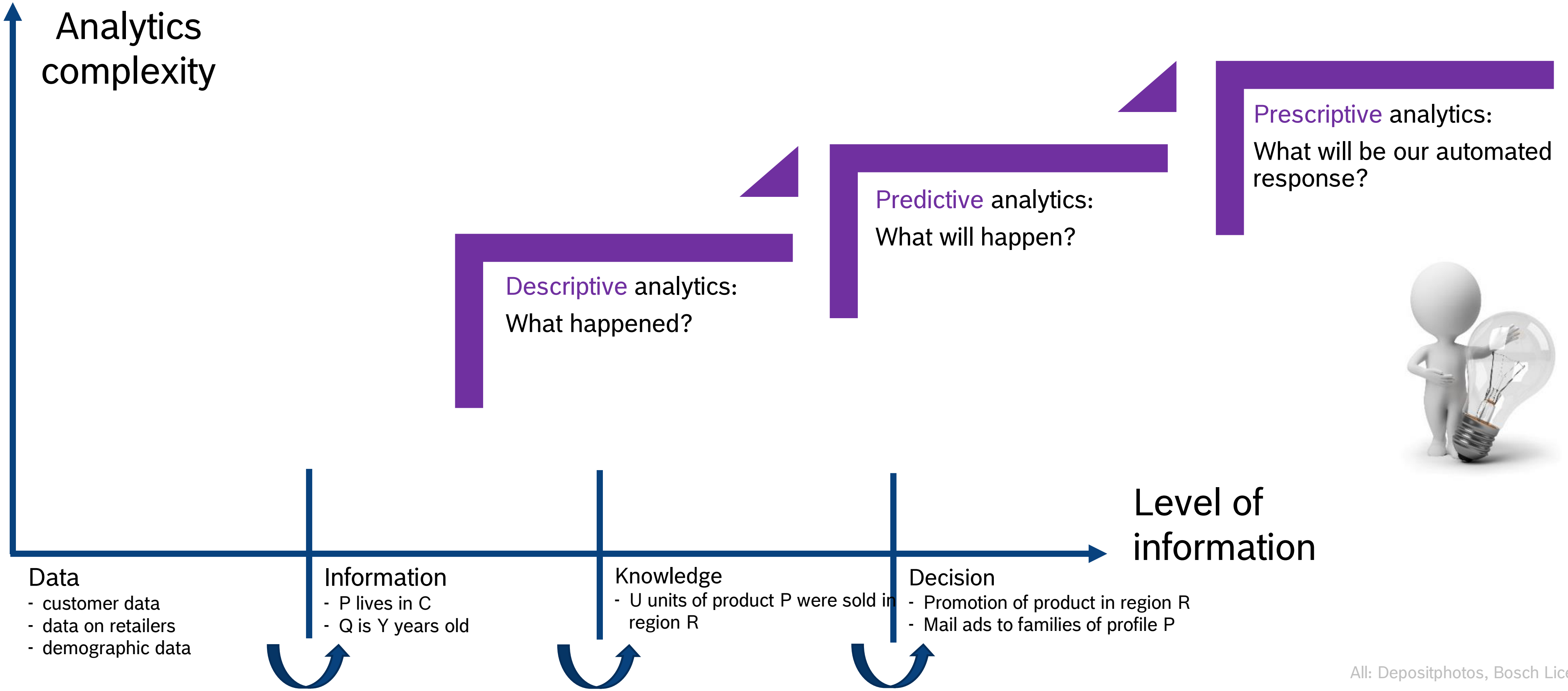


Data Based Value Creation
CRISP-DM





Data Based Value Creation
Levels of Data Mining

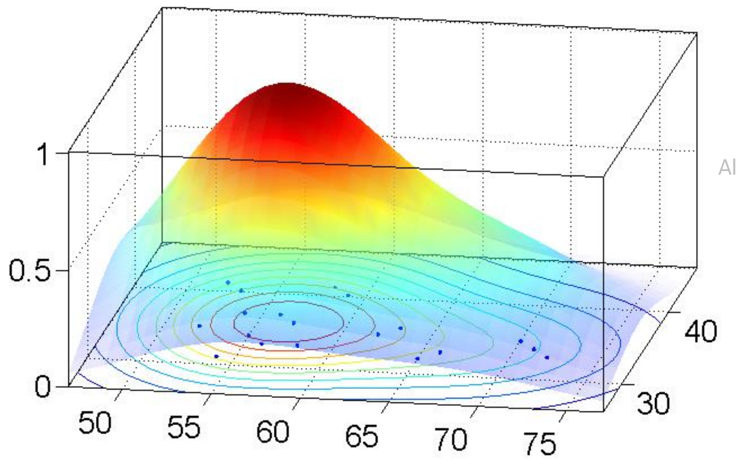


All: Depositphotos, Bosch License

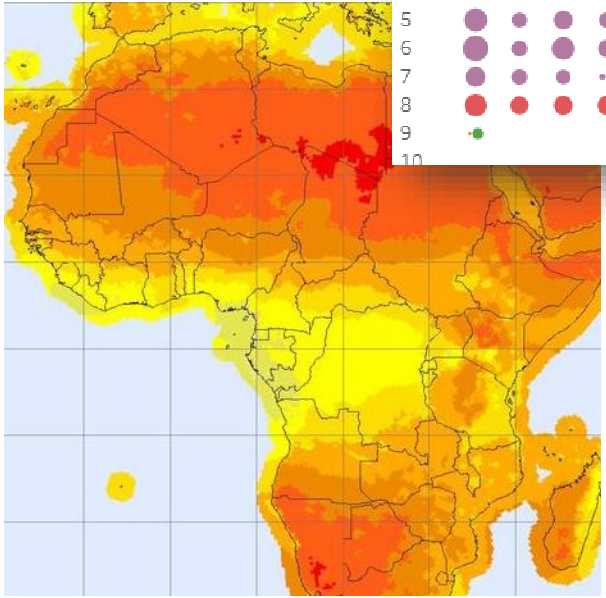
BOSCH

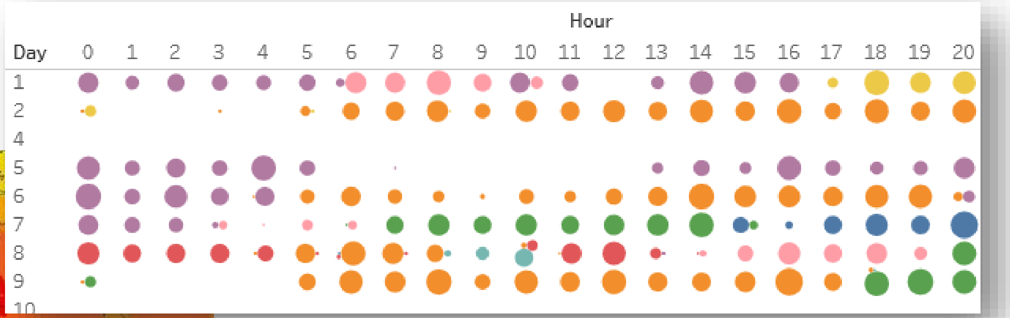
Data Based Value Creation
Descriptive Analytics

We can gain data understanding by the use of descriptive analytics:



All: CC0





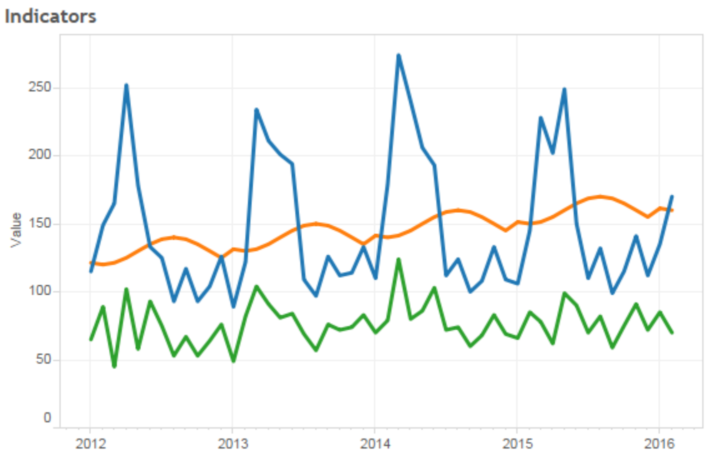
Data Based Value Creation
Predictive Analytics

If we really understand the mechanisms behind our data,

Depositphotos, Bosch License

we can make predictions for the future.

Example: Sales forecasting



Data Based Value Creation
Use Case: Predictive Analytics Toolbox for Finance

A **forecasting toolbox** for Bosch which can answer questions such as the following:

- How much sales is a BU expected to generate in the short and mid-term future?
- Are there **regular sales patterns** in the past from which we are able to **predict future sales** developments?
- What drives **irregularities** and what can i learn for future predictions?

https://pixabay.com/de/wahrsagen-zukunft-magie-astrologie-1989579/, CC0

- Part of **G2/PJ-FCE** in close collaboration with BCAI and CI

Video: Predictive Analytics Toolbox for Finance

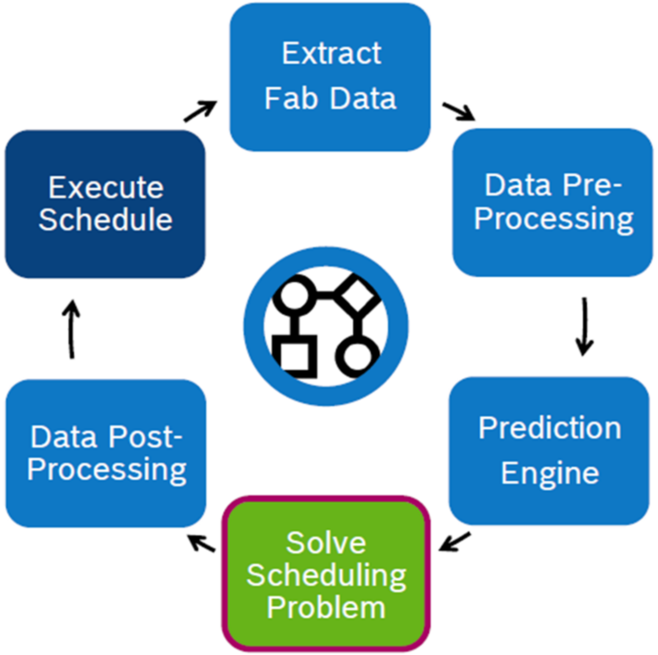
Data Based Value Creation
Prescriptive Analytics

We can expand predictive analytics

to prescriptive analytics by

automated reactions.

Example: Production Scheduling



Data Based Value Creation
Use Case Production Scheduling

Target:

Optimal production plan

also in varying conditions!

RtP1 project requirement:

New optimum production plan

within 5 minutes!

Changing job deadlines/priorities

Bicycle production line by clement127, CC0

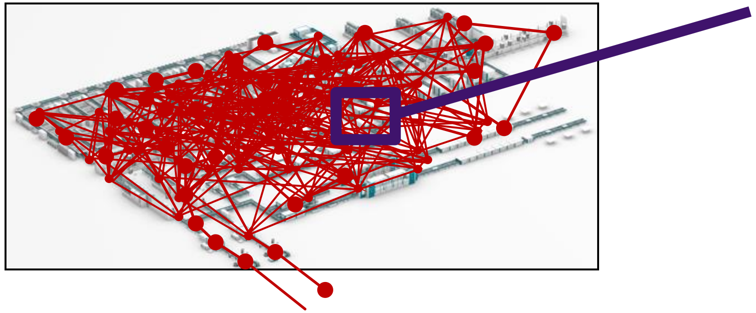
Changing input provision

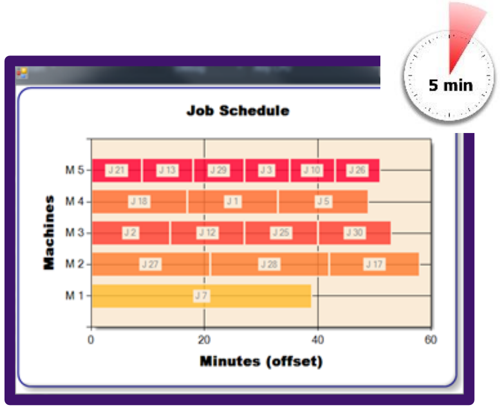
Changing output requirements

Machine stalls

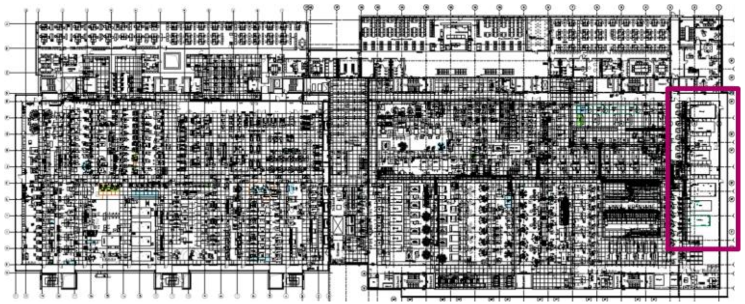
Machine/job-dependent process times

Data Based Value Creation
Use Case Production Scheduling





200mm Implant Workcenter



8 tools + 2 robots for (un-)loading

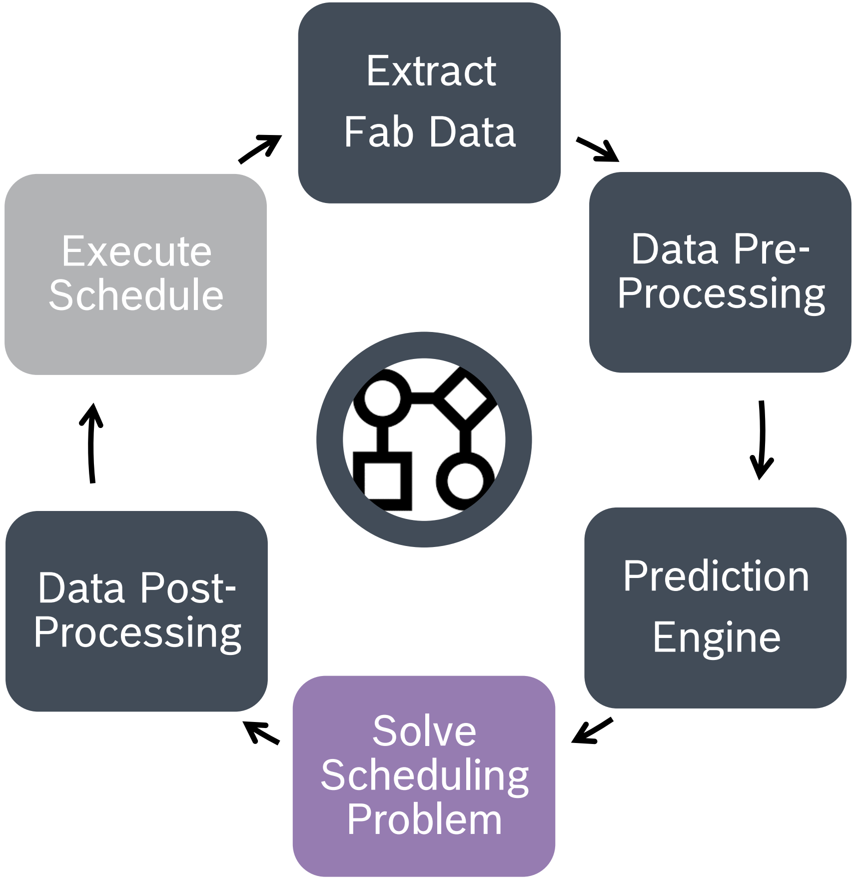
Rule-Based Dispatching

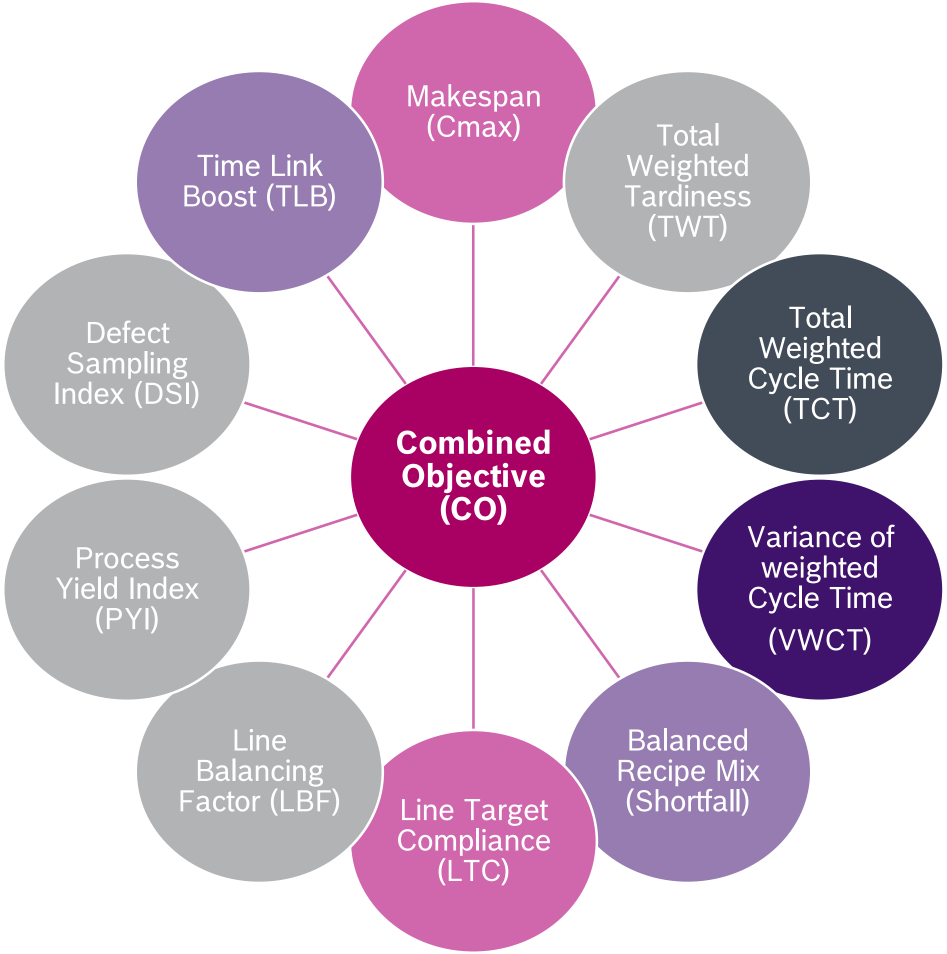
Status quo is production control using rule-based dispatching

Scheduling with Optimization (Solver)

First scheduling solution deployed in RtP1 production after 6 months (BCAI cooperation)

Data Based Value Creation
Use Case Production Scheduling





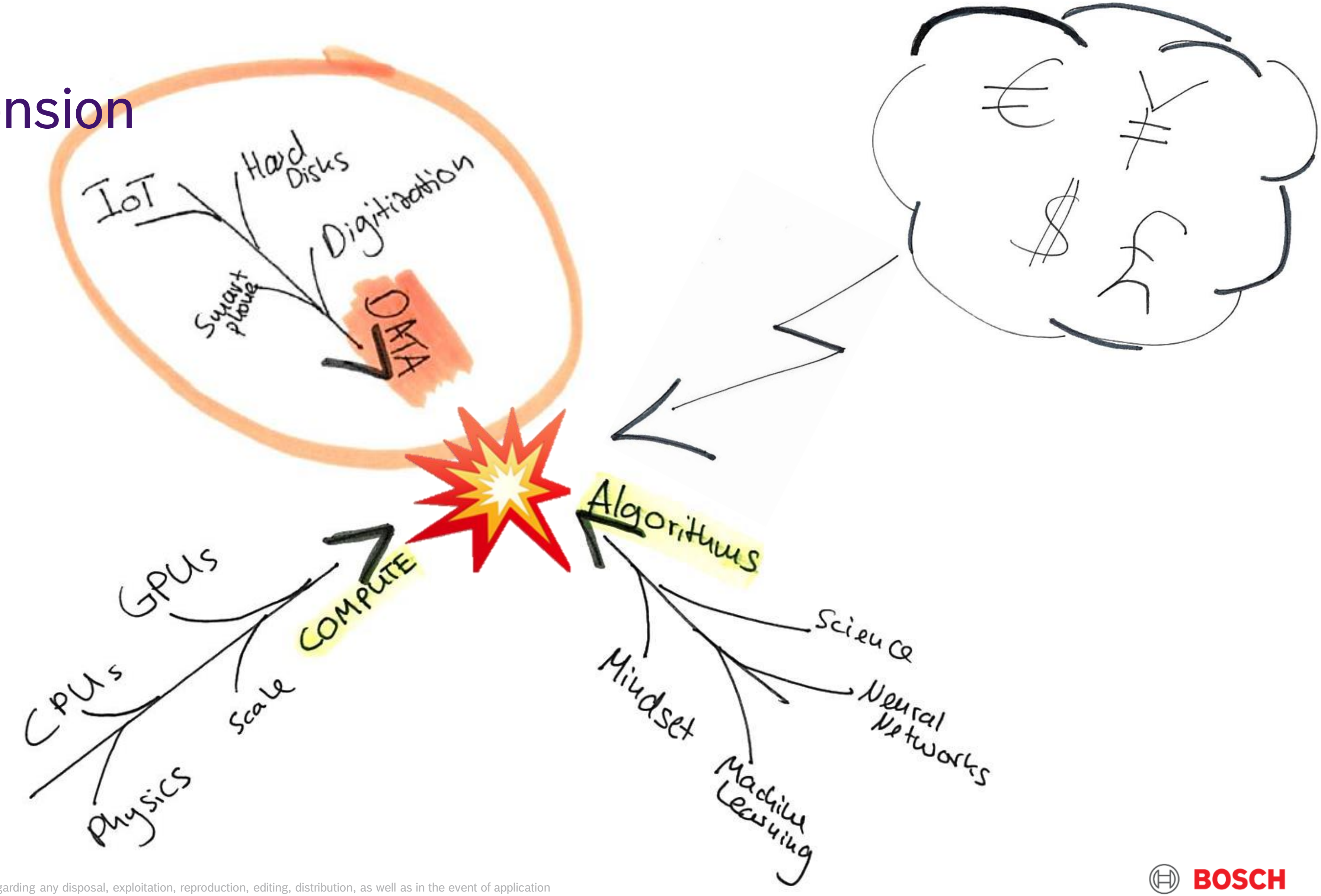
Data Based Value Creation
Use Case Production Scheduling

“Nobody would intuitively do it like that, but it is actually better!”

Minus ~45% StandBy-With-WIP

>5% more throughput with all-times high record

Big Data
The data dimension

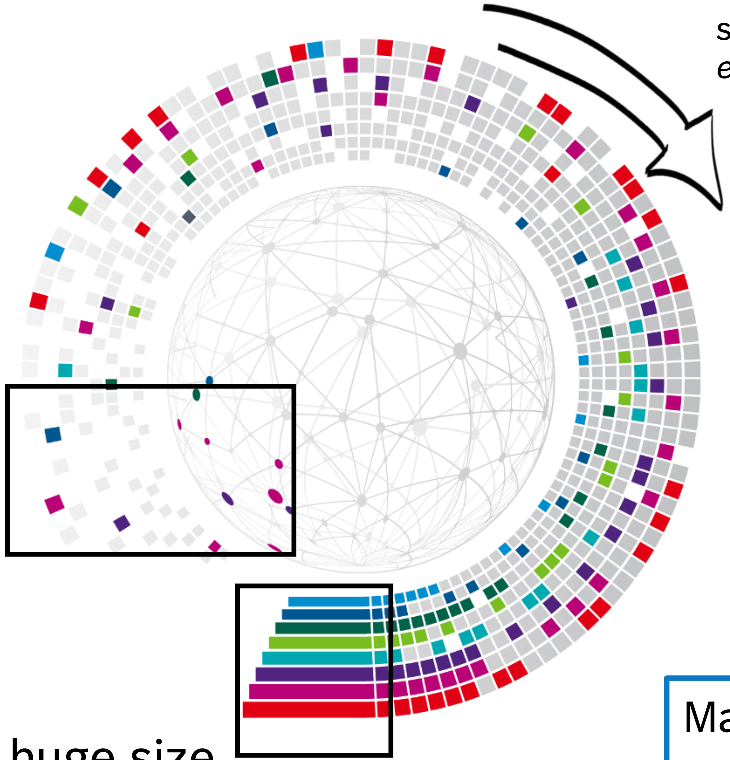


BOSCH

Big Data
The 3 V's

Velocity – data in motion

streaming data (msec to sec to respond)
e.g.: near real-time business data



Variety – data in many forms

e.g. text, video, images;
structured and unstructured
For example: self-driving car's cameras and sensors

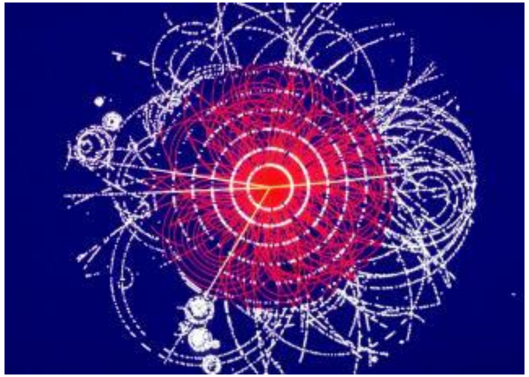
Volume – data of huge size

large amounts of historical data (TBs or more)
e.g.: Connected Industry production lines

Matching **one of these criteria** already indicates that Big Data technologies may be beneficial.

Big Data
What is Big Data? Volume.

- Amount of data so large that we need new methods to process



CERN LHC data totaling 200 PB in 2017

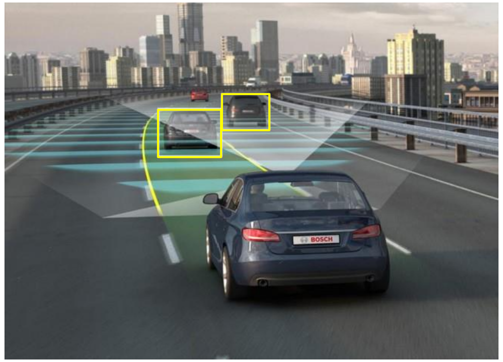
Each day 4 PB are newly generated on Facebook

Cycle Time Reduction needs a large corpus of historical data to reveal patterns

Big Data
What is Big Data? Velocity.

- Data needs to be processed quickly for fast decision-making

High-frequency trading requires **seconds to milliseconds** reaction times



Autonomous driving needs to **process sensor, image, map data** rapidly to steer the car safely

Big Data
What is Big Data? Variety.

- Beyond tables: all sorts of data can now be analyzed

Voice data is analyzed for speech recognition

Pictures/Videos can be used for image recognition tasks

Natural language / text data can be analyzed to extract information

Big Data
When do we speak about big volume data?

- ❖ When we have > 100 GB of data which need to be processed continuously
- ❖ When the working memory of our PC is not sufficient anymore
- ❖ When the amount of data is rising in GBs on a monthly basis
- ❖ When relational databases, Excel etc. do not work efficiently anymore

Big Data
The compute dimension

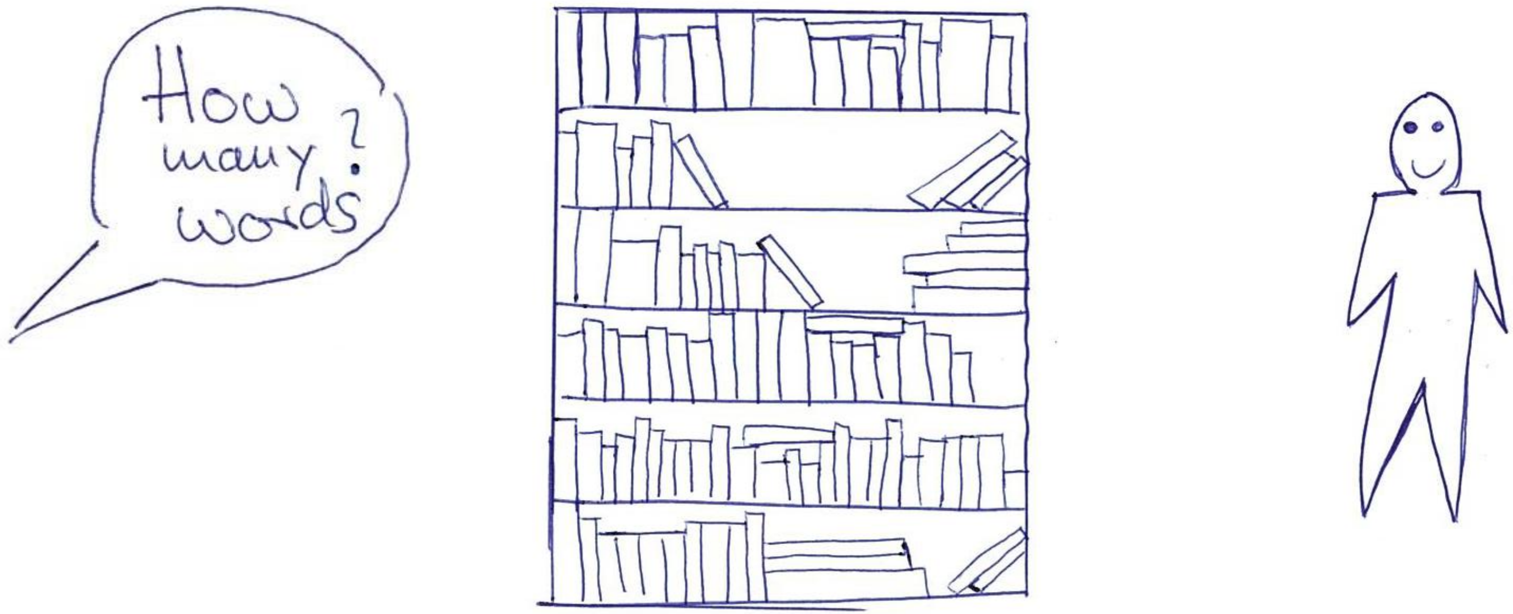
Big Data
Big Data computation – the idea of batch processing

- Goal: process very large amounts of (historic) data

- Idea: Divide and conquer algorithm

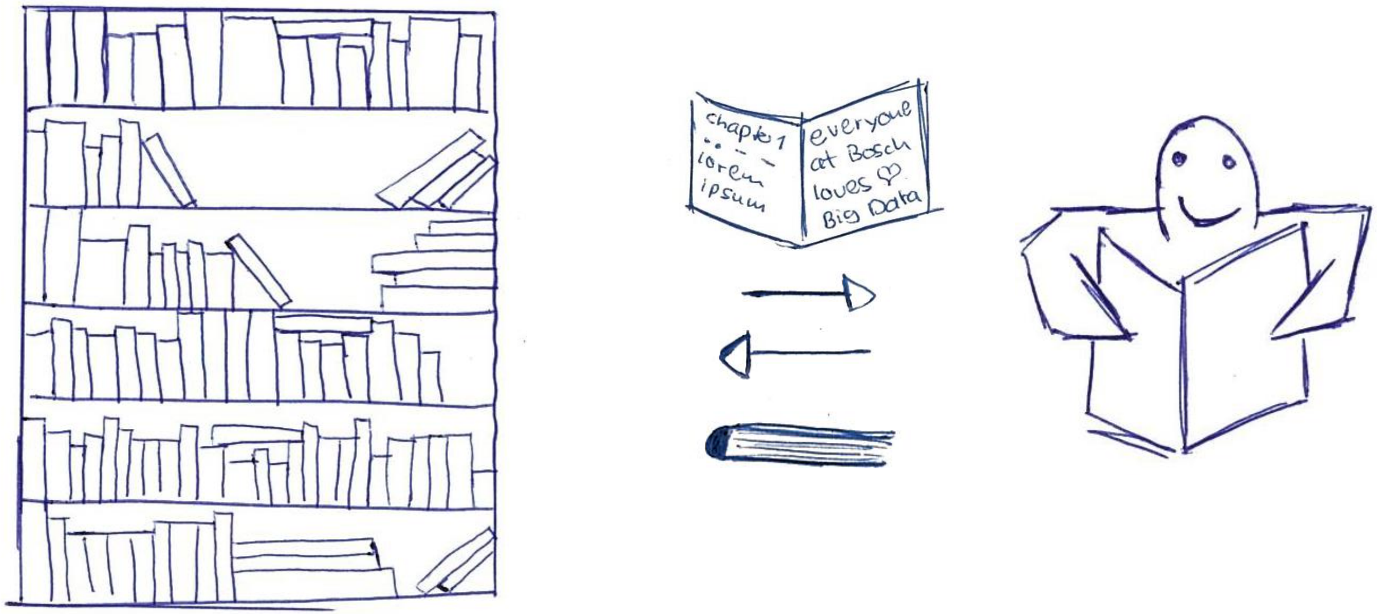
- How-to: Bring “processing to the data”

Big Data
How Big Data computation works – Counting Words



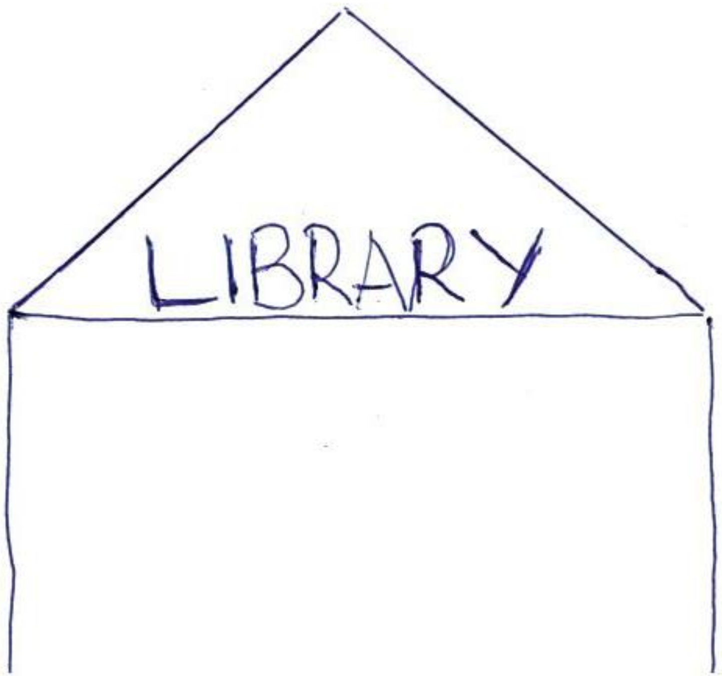
Challenging Task: count number of words in the whole bookshelf.

Big Data
How Big Data computation works – Counting Words



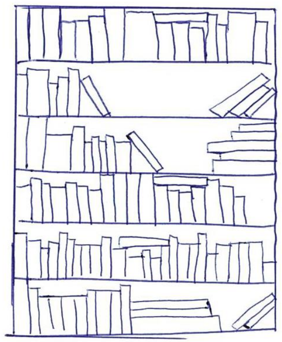
Big Data

How Big Data computation works – Counting Words



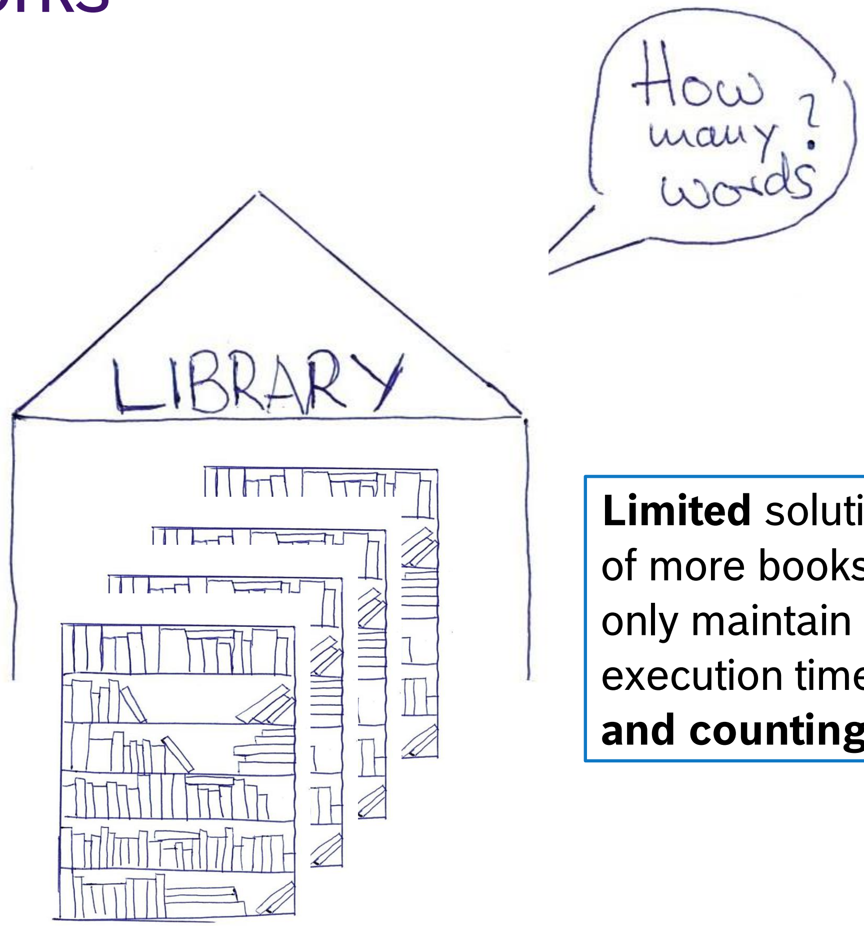
Sometimes you have a library to process …

Big Data
How Big Data computation works



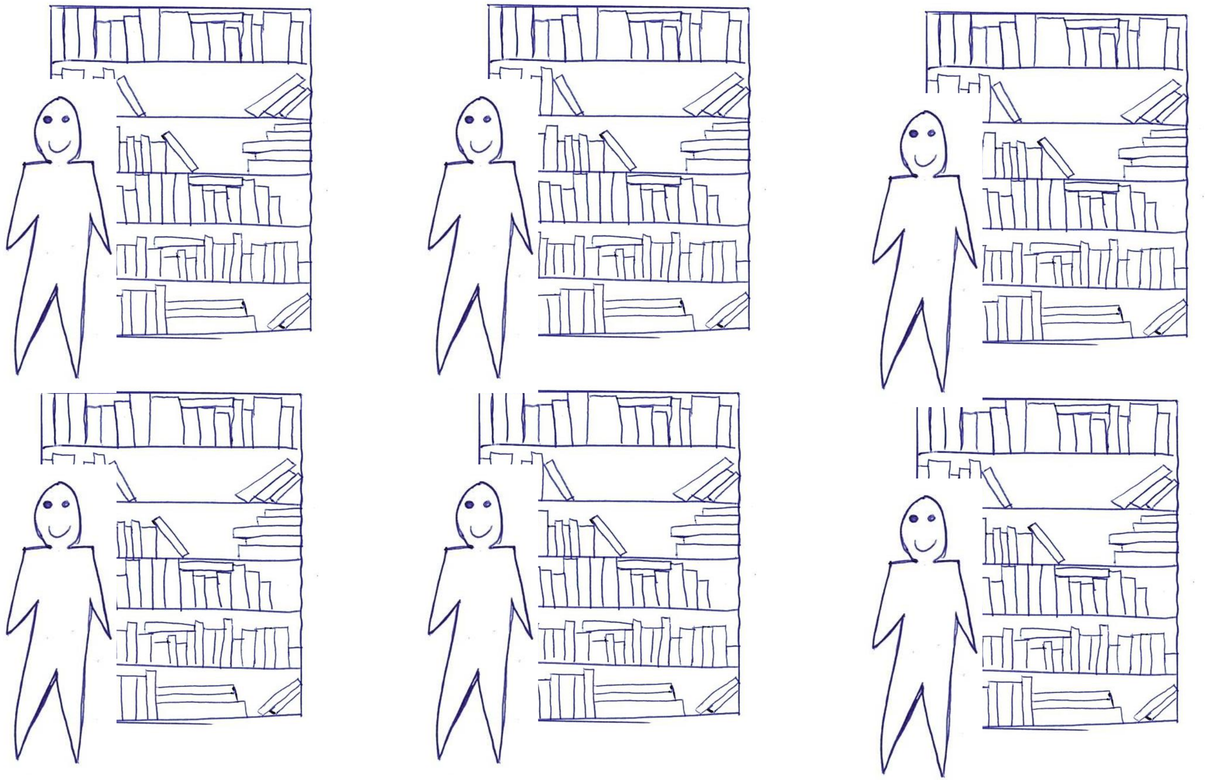
=

same time



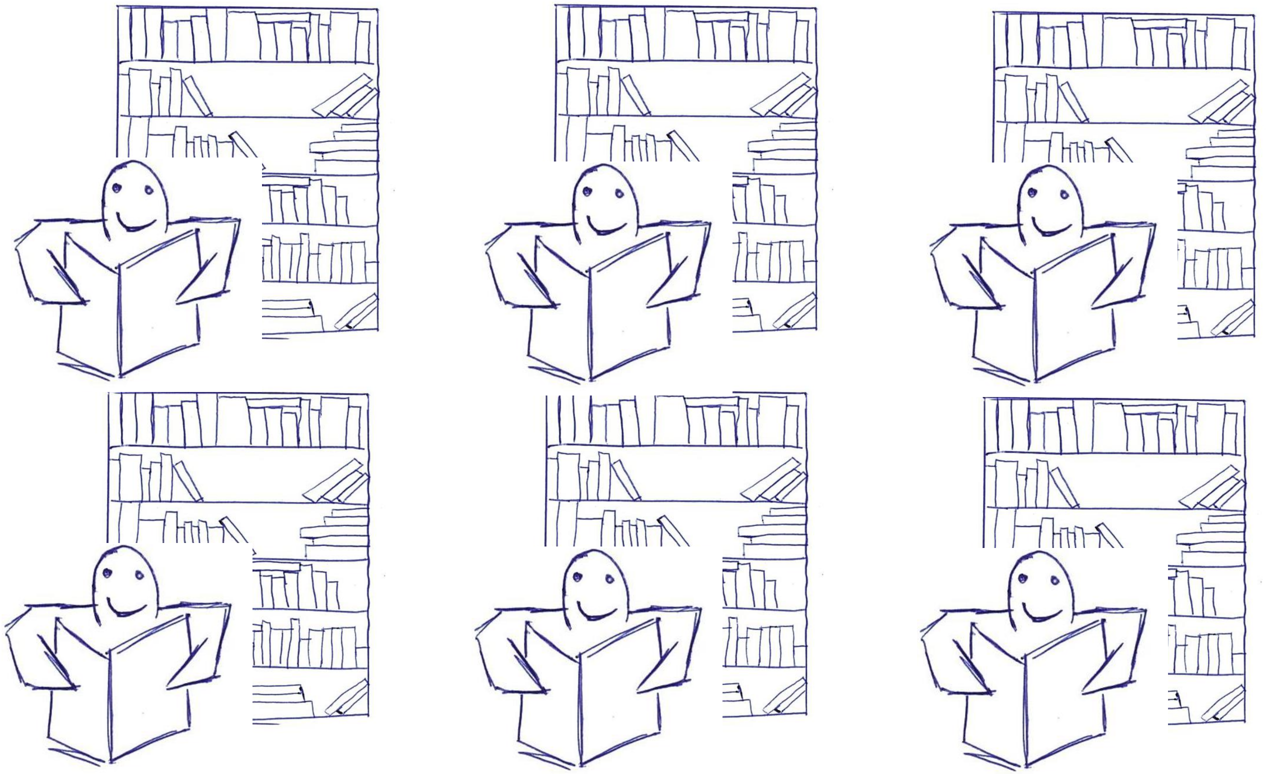
Limited solution: in the case of more bookshelves we can only maintain overall execution time by **reading and counting faster**!

Big Data
How Big Data computation works



We move the workers close to data to minimize overhead: The **computation comes to the data**.

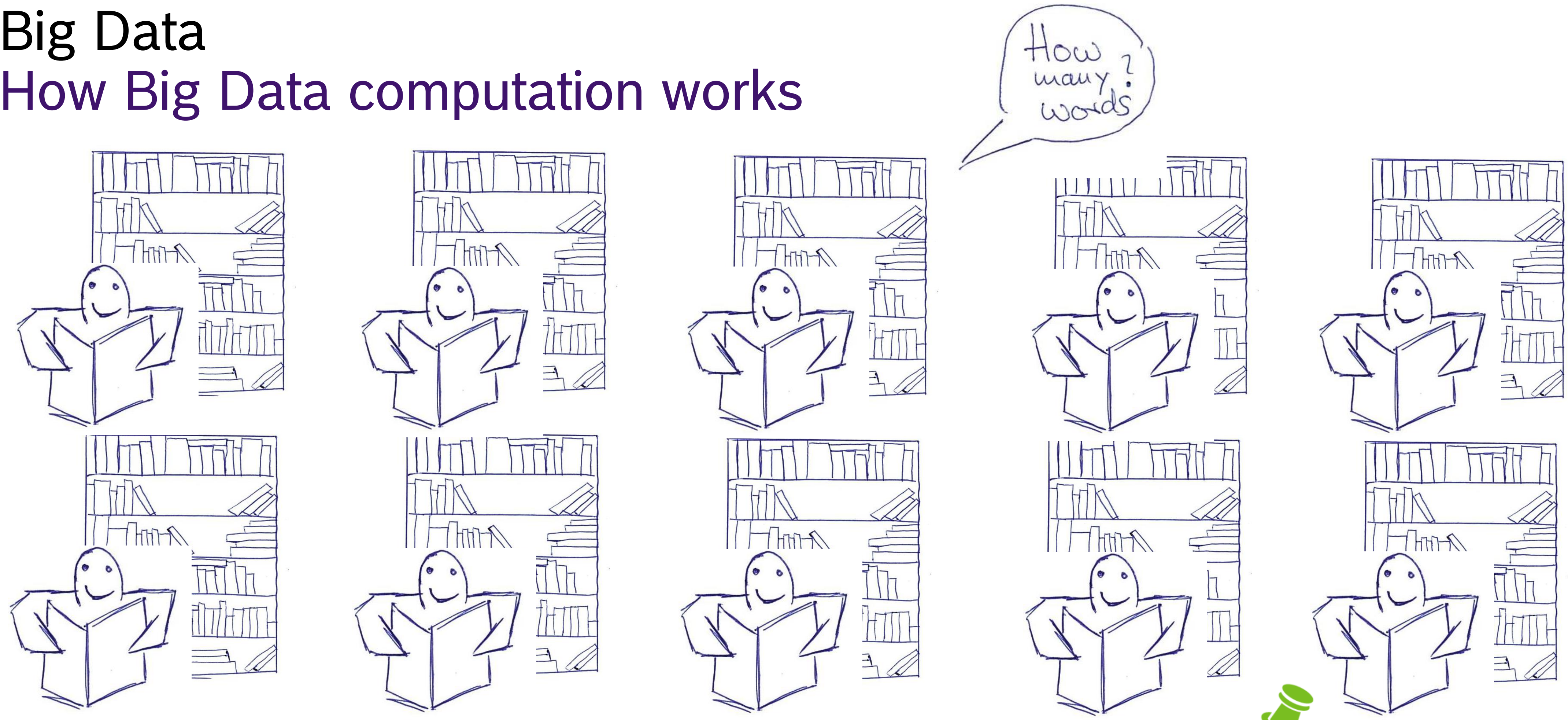
Big Data
How Big Data computation works – Counting Words



The **Big Data** approach is to find a **parallel description** of the task to **process more data in the time**.

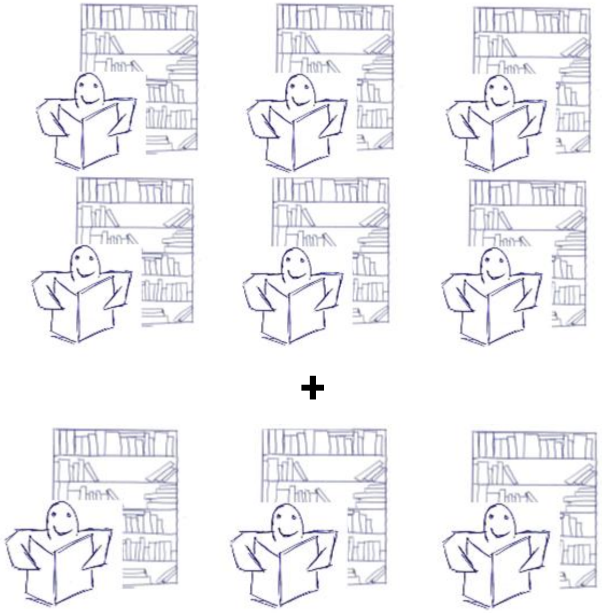
Big Data

How Big Data computation works

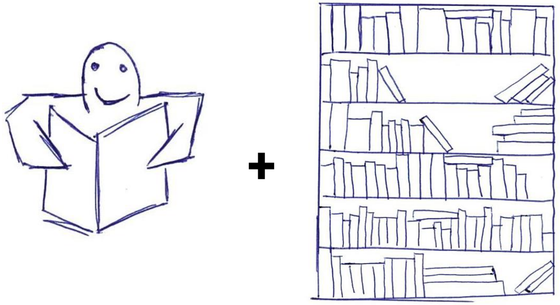


As **data volume grows**, workers can be added simply. The solution is highly **scalable.**

Big Data
Big Data computation – key benefits

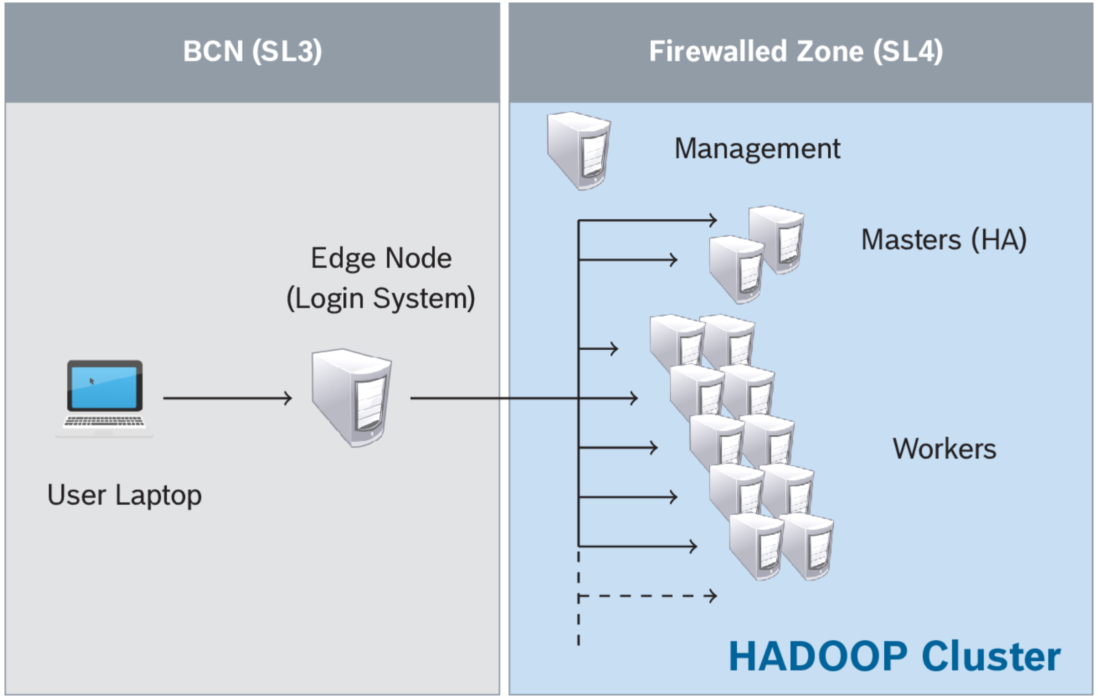


Scalable solution that can grow dynamically with volume



Computation and data are always **bundled** in order make data processable

Big Data
A typical Hadoop cluster



Example configuration:

- Total net storage (10 Worker Nodes): **160 TB**
- Total working memory: **2560 GB**
- Total compute: **240 cores (480 with HT)**

BOSCH

Big Data
Hadoop ecosystem – tools for many situations

Search framework for large amounts of data

In-memory analytics for immediate access

Processing of **quickly incoming data** (velocity)

Big Data
Data quality

Schema Level

- Illegal values 1
- Violated attribute dependencies
- Uniqueness
- Referential integrity 2
- ...

Instance Level

- Missing values 3
- Misspellings
- Redundancy/duplicates
- Contradictory values
- Wrong references 4
- ...

Schema Level

- Structural conflicts 5
- Naming conflicts
- ...

Instance Level

- Inconsistent representations 6
- Inconsistent timing
- ...

1 Not allowed

2 Department not defined

Name	First Name	SSN	Date of Birth	Department
Müller	Maria	12345	01.13.1988	127
Müller	Maria	12345	01.03.1988	

4 contradiction

3 missing

Source 1

Name	Street	City	Sex
Müller	Amalienstraße	Munich	F

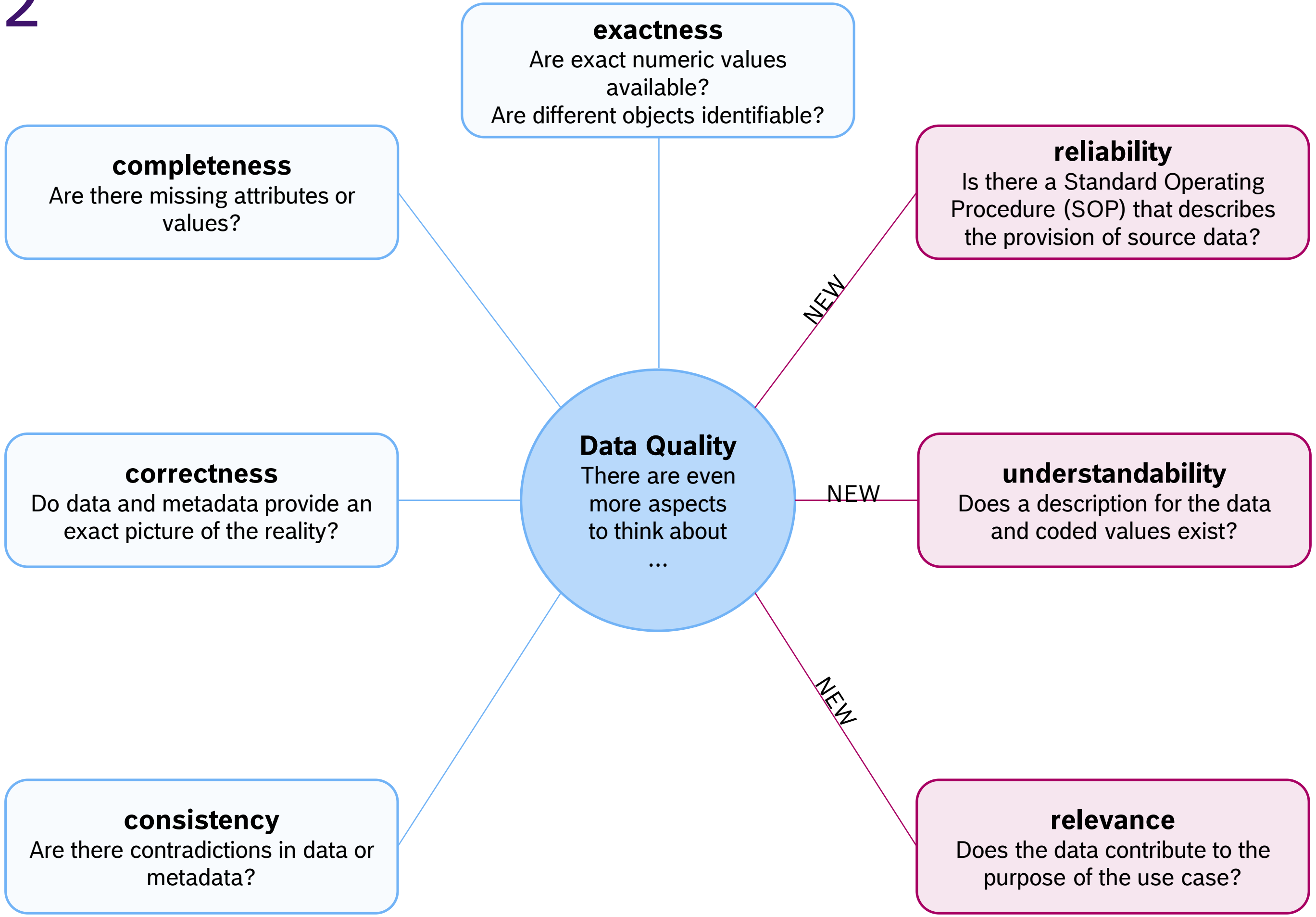
5 Structure differs

6 Representations differ

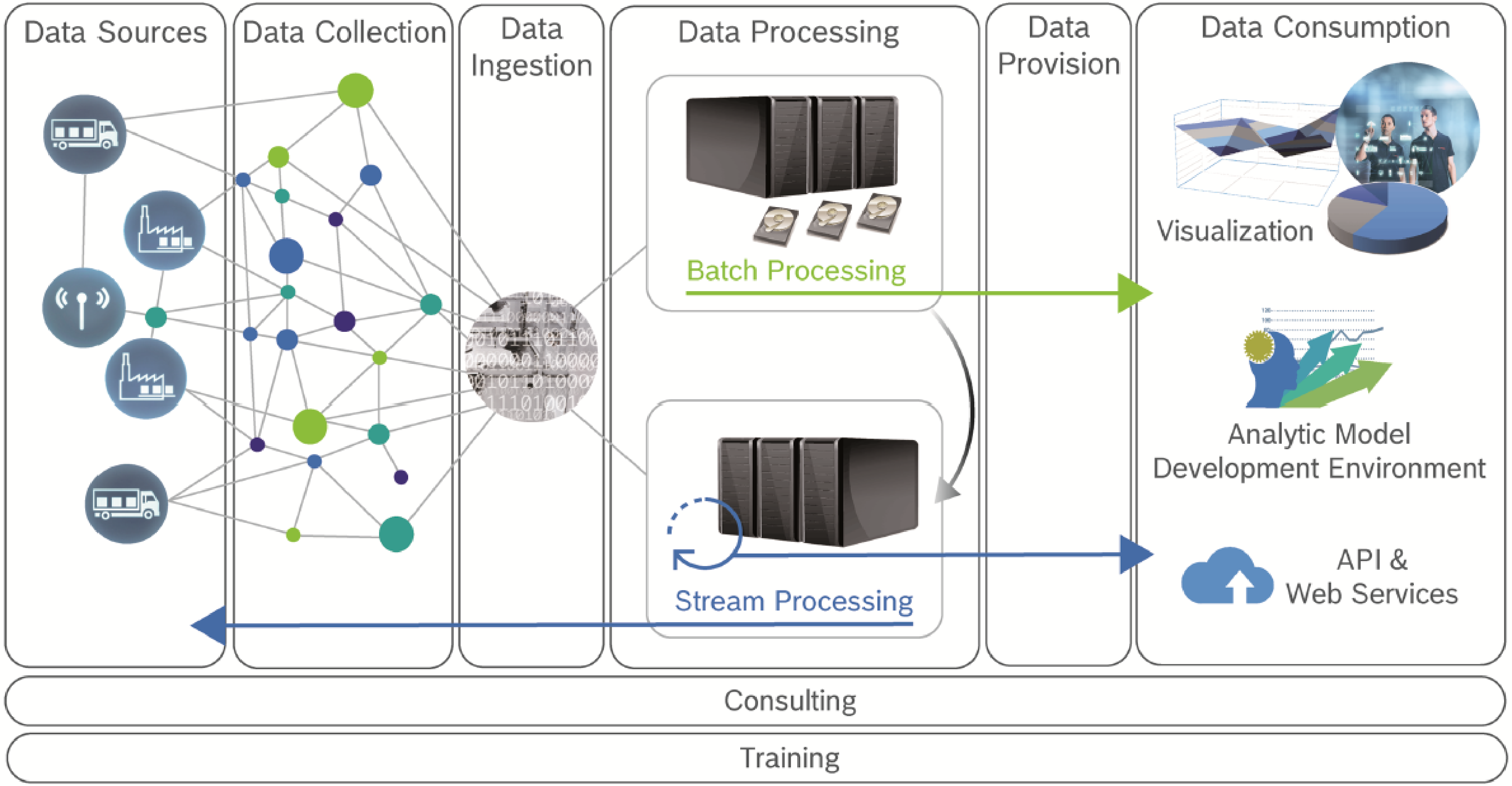
Name	Address	Sex
Müller	Amalienstraße, Munich	1

Source 2

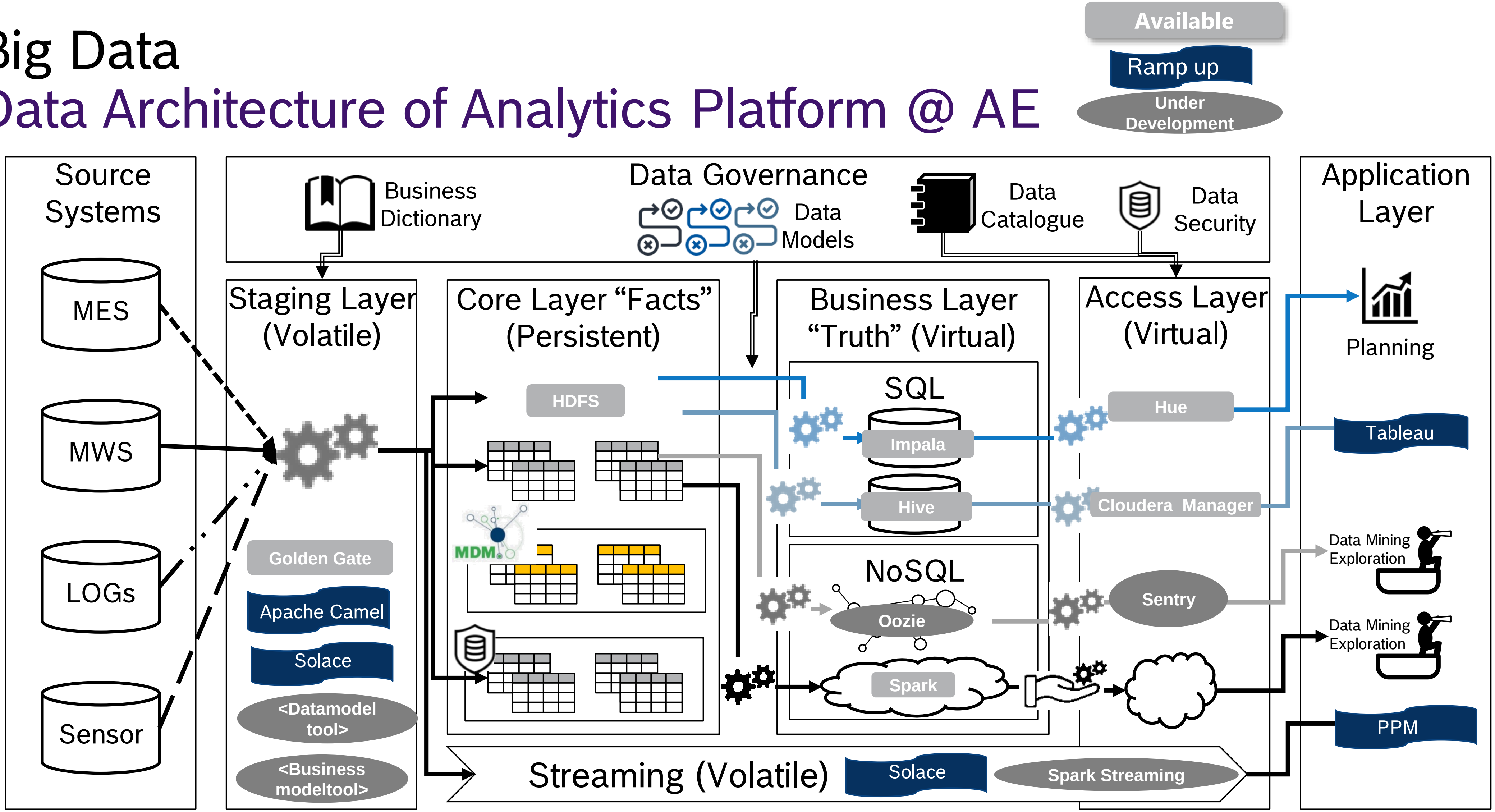
Big Data
Data quality 2



Big Data
RB i4.0 Analytics Platform



Big Data
Data Architecture of Analytics Platform @ AE



BOSCH

MACHINE LEARNING

Machine Learning
Definition

Definition by Prof. Tom M. Mitchell

A computer program is said to learn

from experience E

w.r.t. a task T and measure P,

if its performance P on T,

improves with experience E.

http://bzo.bosch.com/bzo/de/article_page_37325.html (Bosch Zünder Online)

Example

Your e-mail program improves its spam filter mechanism by watching you marking e-mails as spam or not.

Machine Learning
Learning from experience

Depositphotos, Bosch License

Learning...

❖ To say “car”

❖ To distinguish cars from trucks

❖ To drive a car ...

all requires experience!

Machine Learning
Do machines learn from experience?

Traditional software development

- ❖ Task described by formal specification
- ❖ Implemented with an unambiguous set of instructions

```
def bar(n):
        if n < 0:
                n = 0

        result = 0
        for i in range(n+1):
                result +=i
        return result

def foo(m):
        l = lambda x: (x+1)*x /2
        return l(m)

if __name__ == "__main__":
        main(sys.argv)
```

Job of the programmer:

- ❖ Find out the logics to solve the task
- ❖ Implement path to solution step by step

Job of the program:

- ❖ Execute step by step what the programmer implemented
- ❖ No changes from experience

BOSCH

Machine Learning

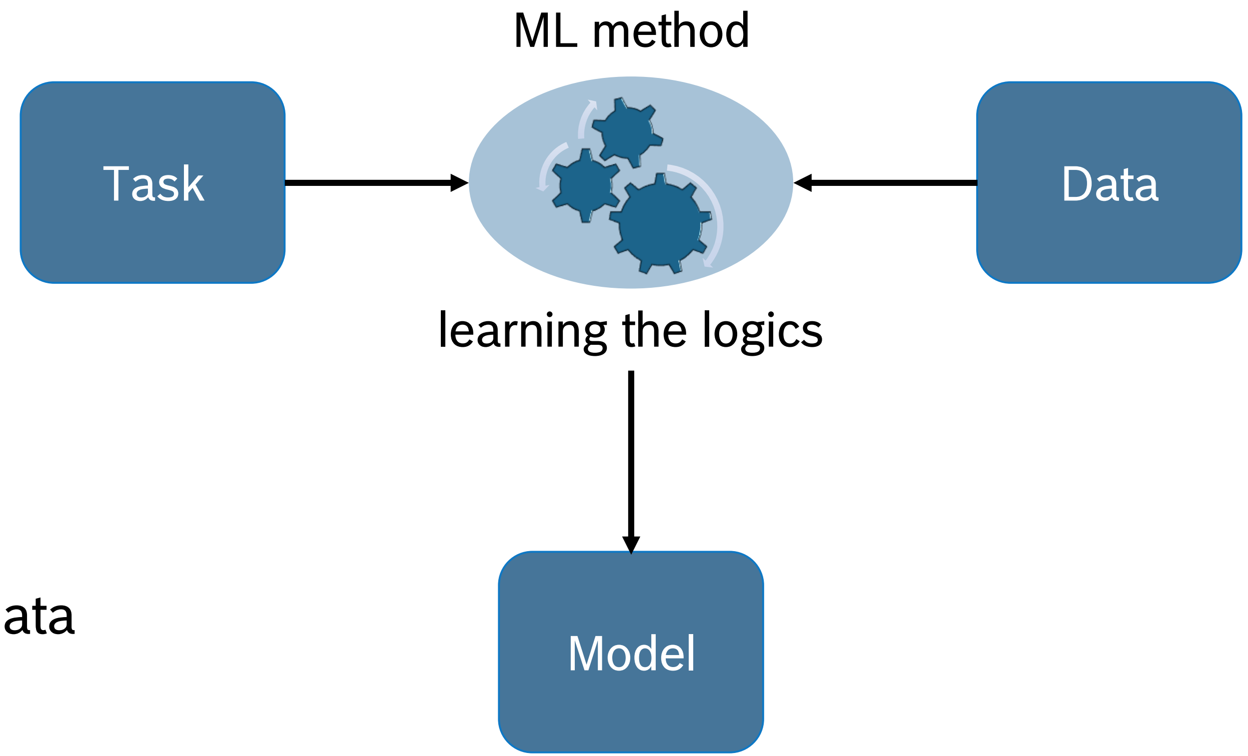
Machine Learning (ML) as a new software development paradigm

Job of the ML engineer:

- ❖ Find the right ML method
- ❖ Feed it with the right data according to the task to solve

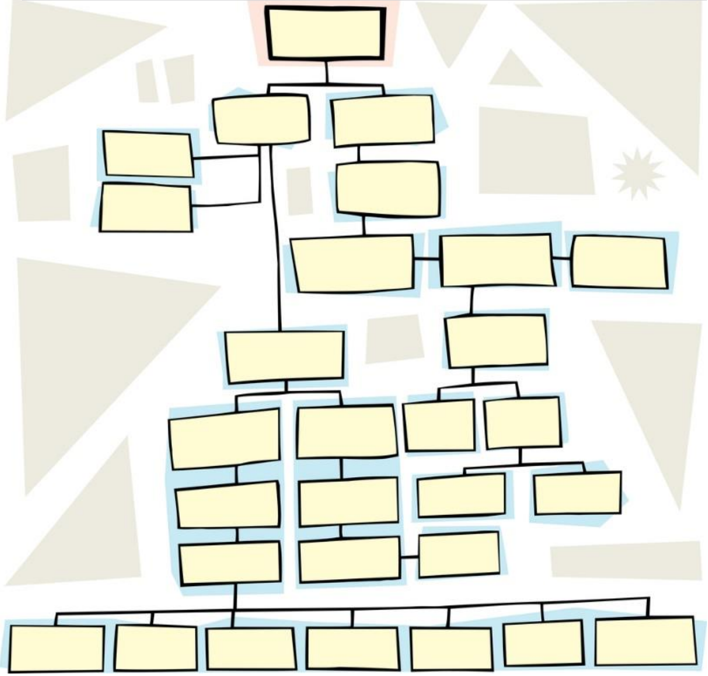
Job of the ML method:

- ❖ Learn the logics to solve the task
- ❖ Optimize iteratively by learning from data describing the right behavior



Machine Learning
Explicit vs. Implicit Knowledge

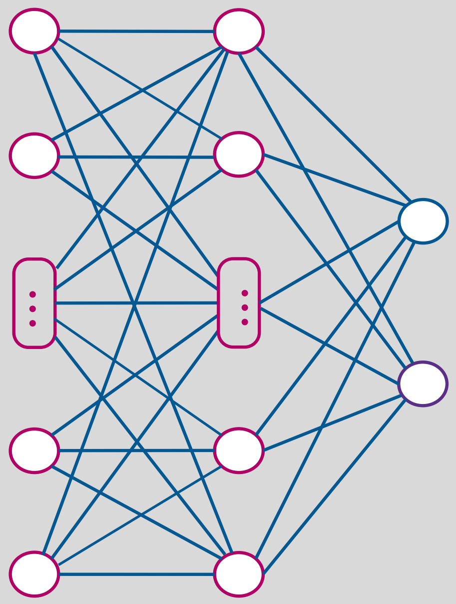
Traditional Programming



Explicit (rule based) knowledge

Implicit (data based) knowledge

Machine Learning Model



Machine Learning

Example: a daisy vs tulip model

In your garden, there are daisies and tulips. Your mom tells you to pick daisies for her.

?

We want to solve your daisy picking task

by building a model,

i.e. a system that helps you to

distinguish daisies and tulips!

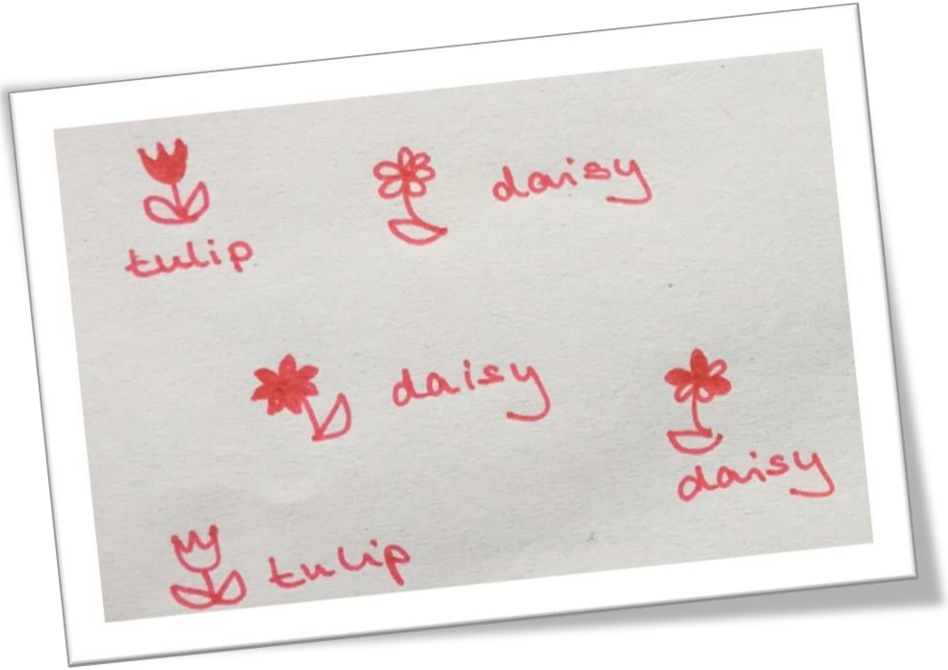
BOSCH

Machine Learning
Features as a basis for the model

Now choose features,

i.e. criteria that help you and the model

to decide if a flower is a daisy or a tulip.



Examples features:

- Number of leaves
- Height
- Color

BOSCH

Machine Learning
Model building

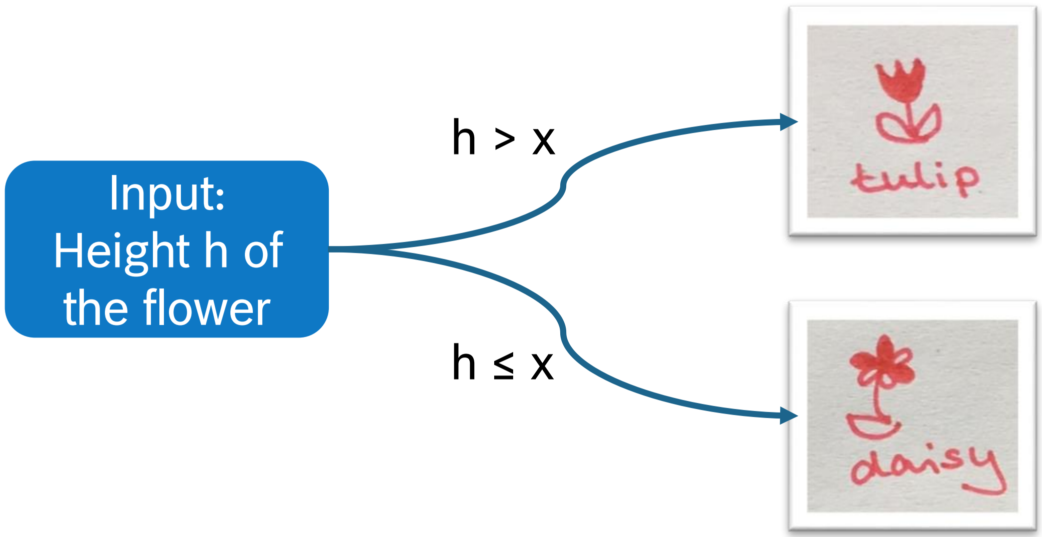
1. Choose a feature (height)

2. Take measurements

Type	Height (cm)
Daisy	8
Tulip	16
Tulip	20
Daisy	4
Tulip	9
Daisy	5
Daisy	10

3. Build a model

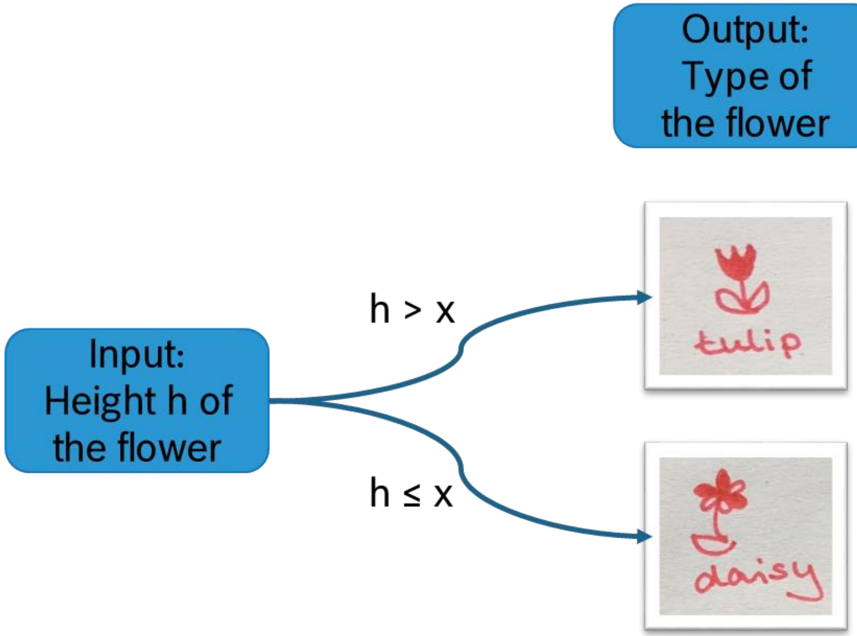
Output: Type of the flower



This model has a parameter x.

Machine Learning
Model training

Type	Height (cm)
Daisy	8
Tulip	16
Tulip	20
Daisy	4
Tulip	9
Daisy	5
Daisy	10



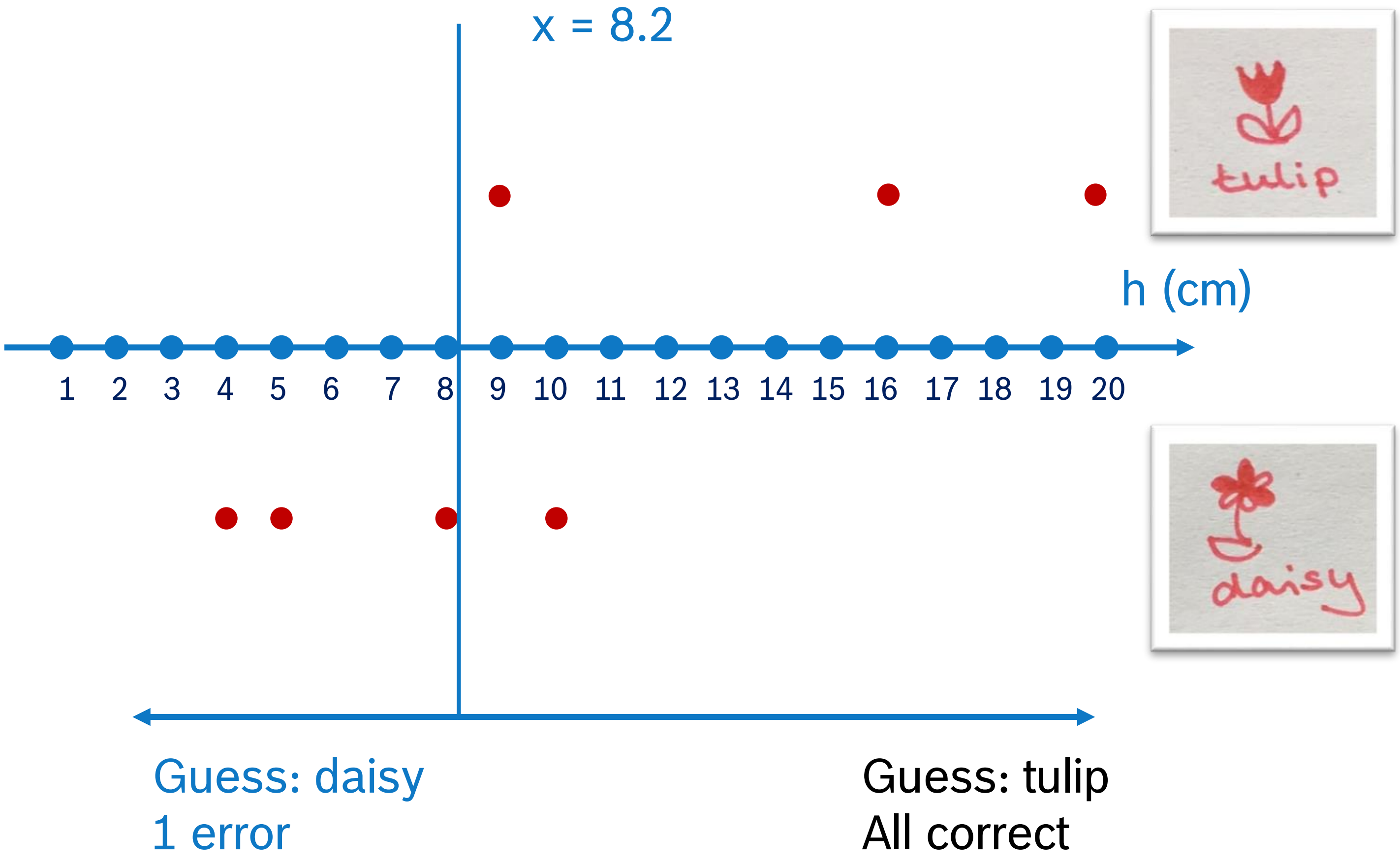
Model training means to optimize the parameter x such that it makes the least errors.

We learn the parameter of the model from data!

Exercise: Optimize x!

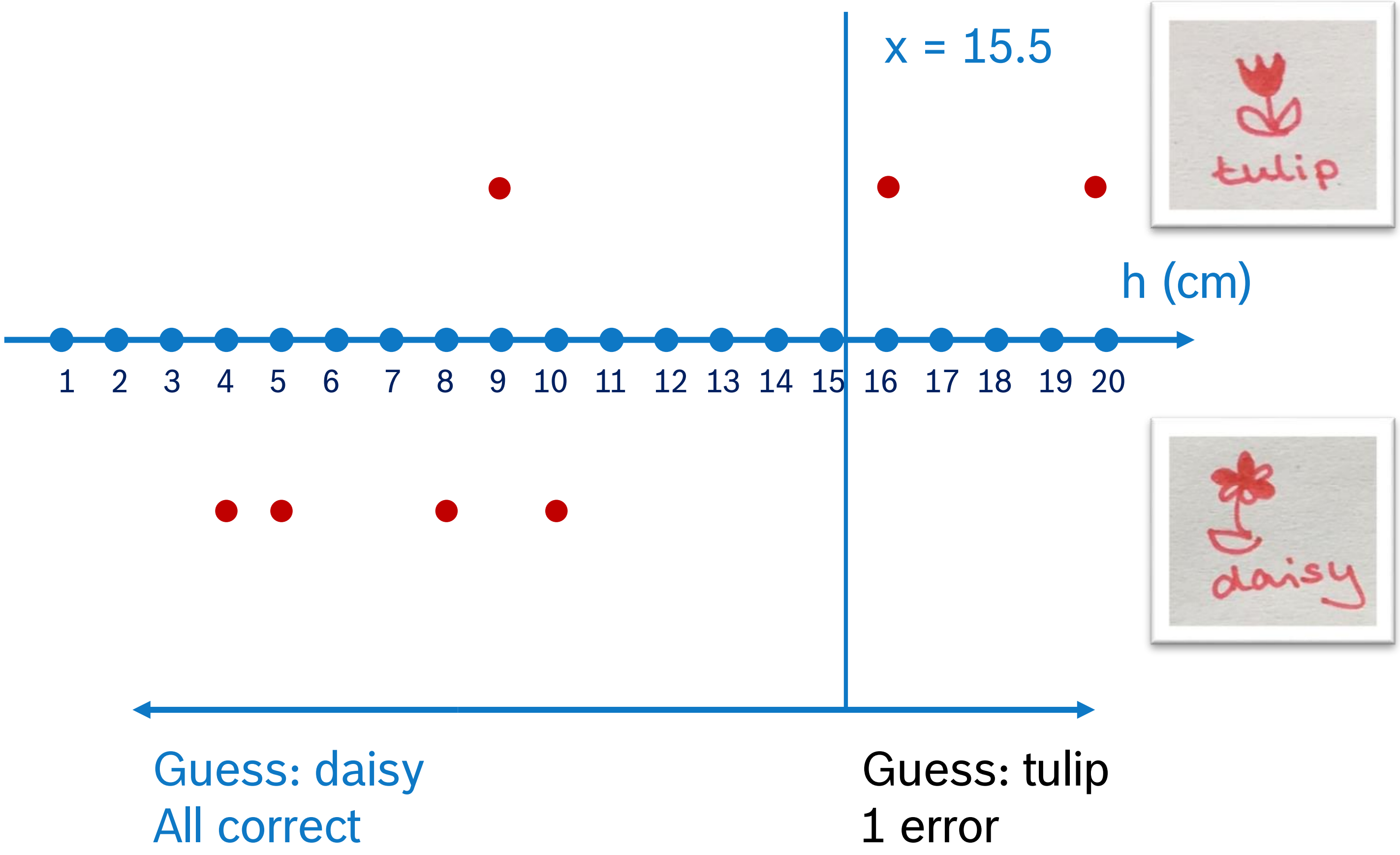
Machine Learning
Optimized model 1

Type	Height (cm)
Daisy	8
Tulip	16
Tulip	20
Daisy	4
Tulip	9
Daisy	5
Daisy	10

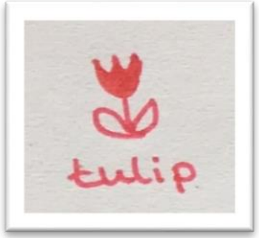


Machine Learning
Optimized Model 2

Type	Height (cm)
Daisy	8
Tulip	16
Tulip	20
Daisy	4
Tulip	9
Daisy	5
Daisy	10



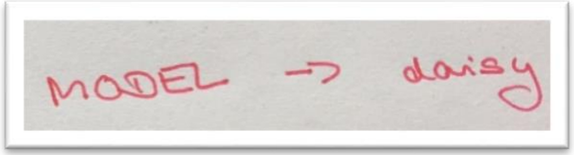
Machine Learning
Training Strategy: Supervised Learning



Model training in 3 steps:

Predict

Estimate output of your model based on new input



Error check

Evaluate the error between prediction and target output

Learn

Update model parameters to mininize error

modify x

Machine Learning
Model Accuracy

How good is our model and our parameter choice?

Different ways to measure accuracy:

$$\frac{\#\ \text{correct predictions}}{\#\ \text{predictions}}$$

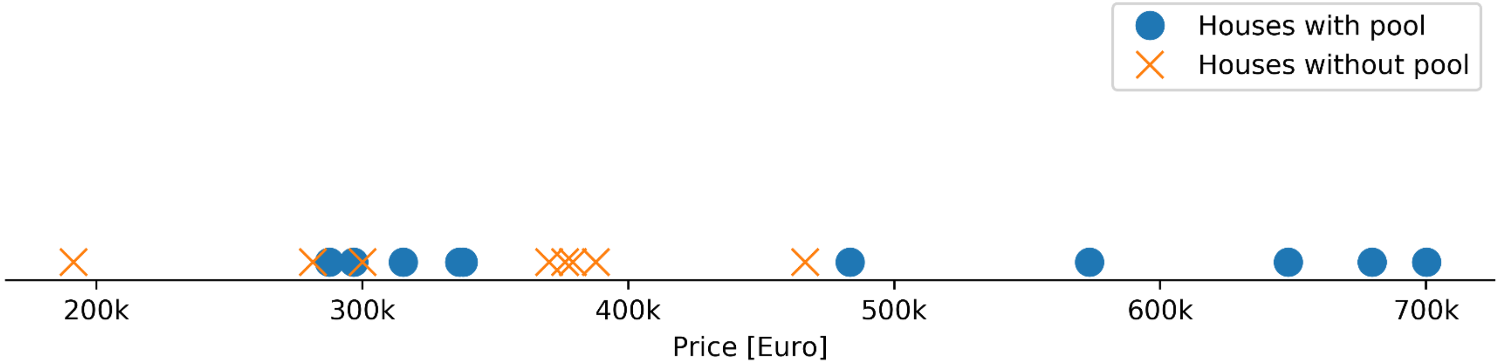
<table>
<tr><td>Prediction
Reality</td><td></td><td></td></tr>
<tr><td></td><td>OK</td><td>Error 1</td></tr>
<tr><td></td><td>Error 2</td><td>OK</td></tr>
</table>

Machine Learning
Another Example

Which house has a pool?

Price	Pool
200k	0
405k	1
230k	0
357k	1
189k	0
171k	1
…	…

Flickr designmilk, CC BY-SA 2.0



Machine Learning
Model Accuracy

Input:
Price of the house

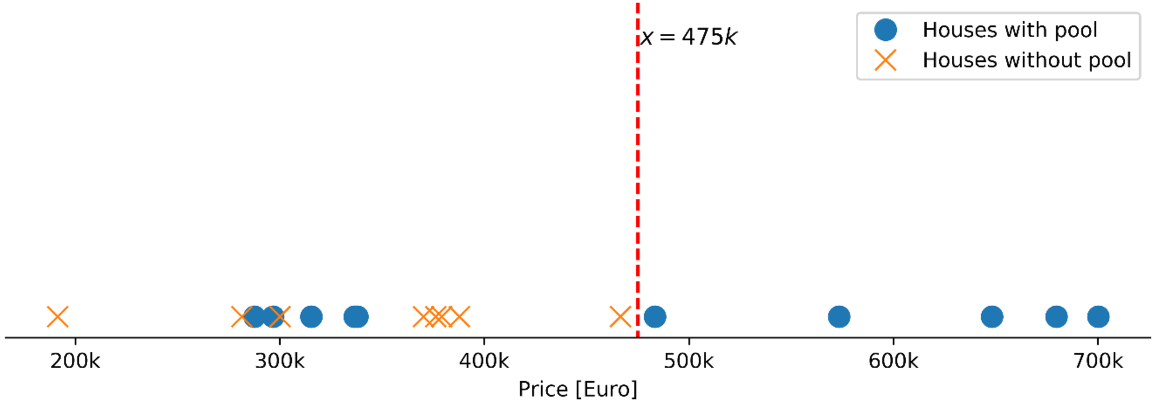
price ≤ x

price > x

Output:
Pool or no pool

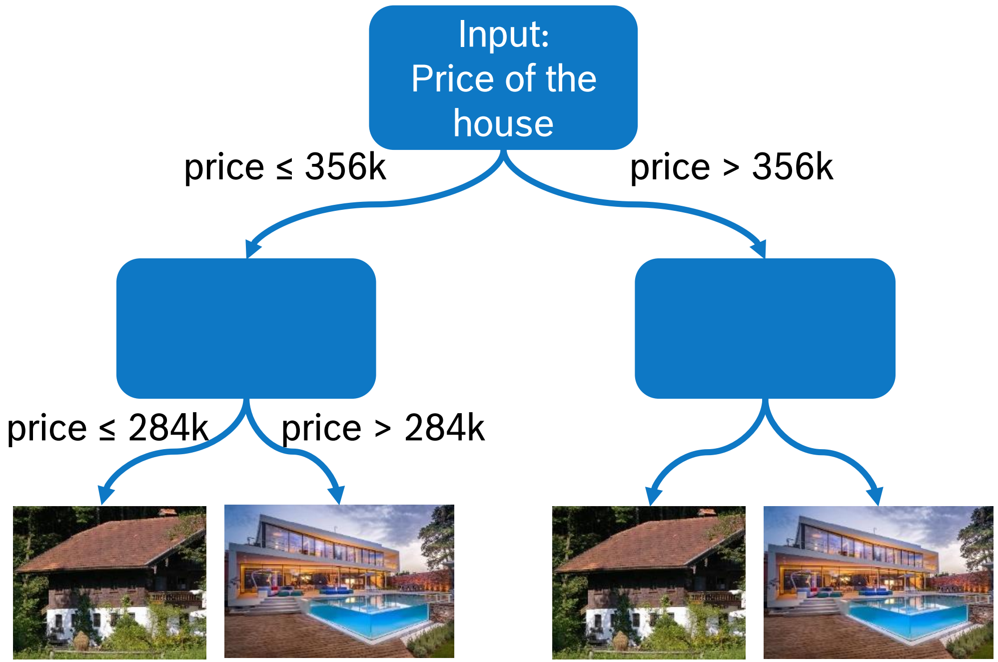
Bauernhaus im Innviertel by Werner100359, CC BY-SA 3.0

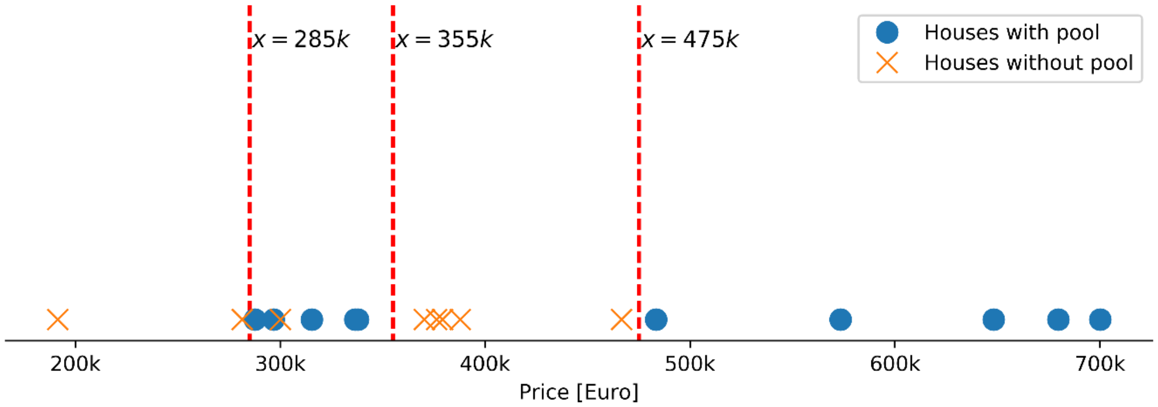
Flickr designmilk, CC BY-SA 2.0



Accuracy: 77.8%

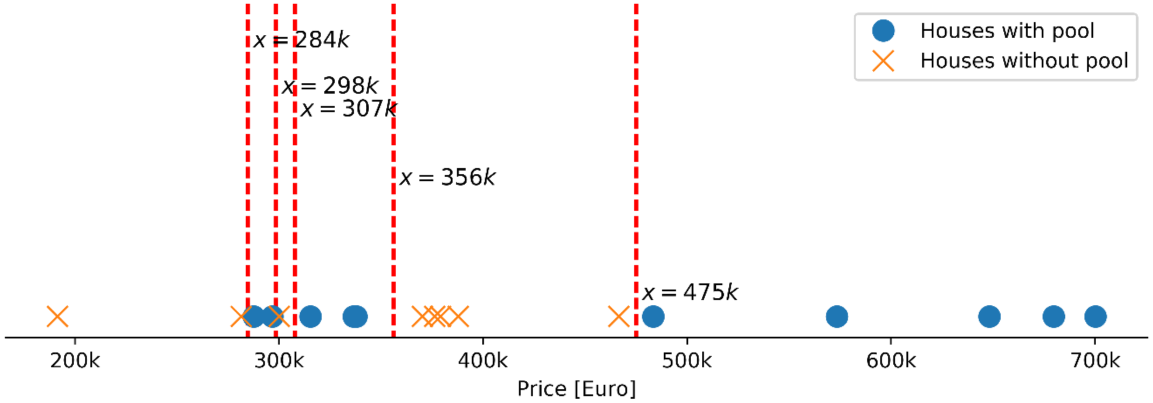
Machine Learning
Model Accuracy





Accuracy: 94.4%

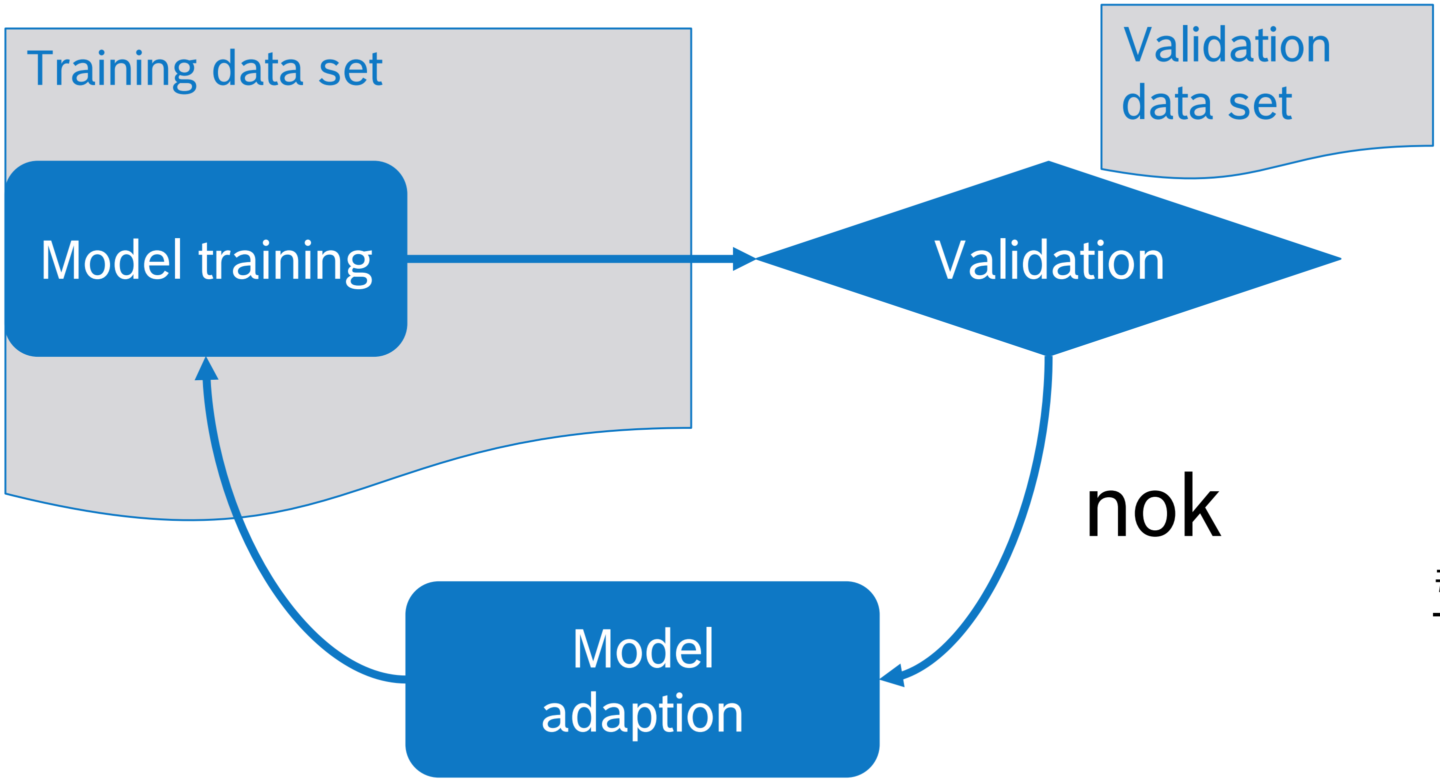
Machine Learning
Model Accuracy



Accuracy: 100%

BOSCH

Machine Learning
Model Validation & Adaption



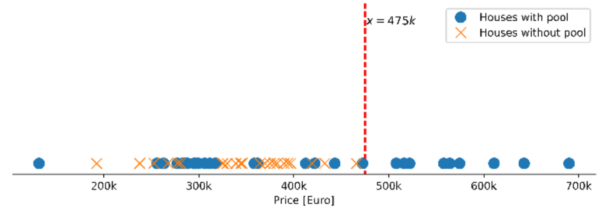
Adapt model structure and/or complexity

Accuracy:

$$\frac{\#\ \text{correct predictions}}{\#\ \text{predictions}}$$

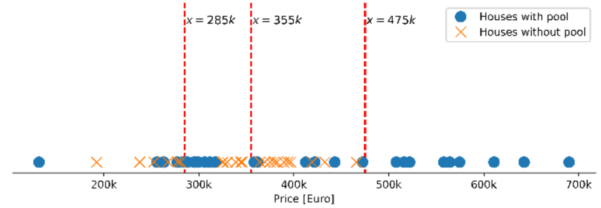
Machine Learning
Model Validation

Apply trained model to houses not used for training (validation data)



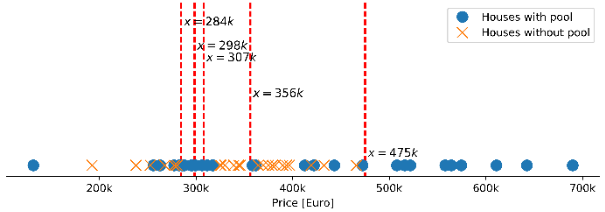
Training accuracy: 77.8%

Validation accuracy: 61.1%



Training accuracy: 94.4%

Validation accuracy: 66.7%

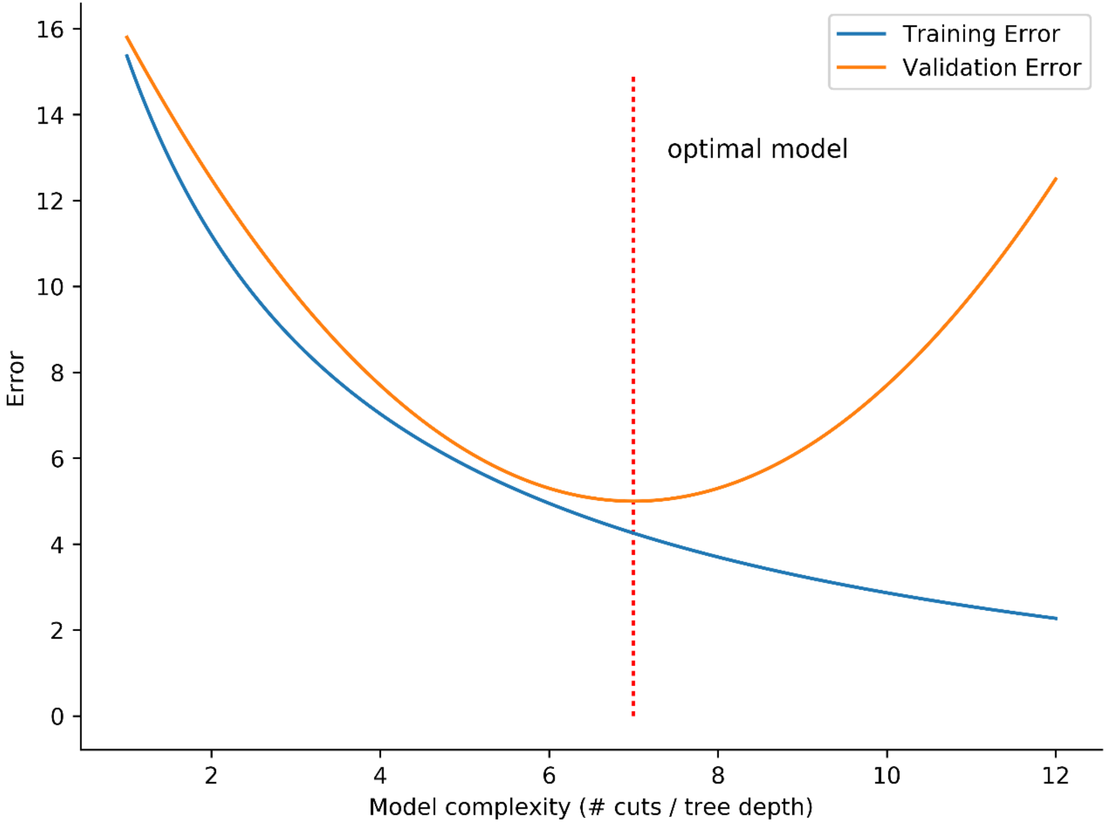


Training accuracy: 100%

Validation accuracy: 62.9%

BOSCH

Machine Learning
Training vs Validation Error



Machine Learning
Task: Classification

The task in the examples is called classification.

Swedish Fish, gummi bears and gummi worms by ChildofMidnight, CC BY-SA 3.0

Target: sort data into predefined classes

Examples: scrap (yes/no)

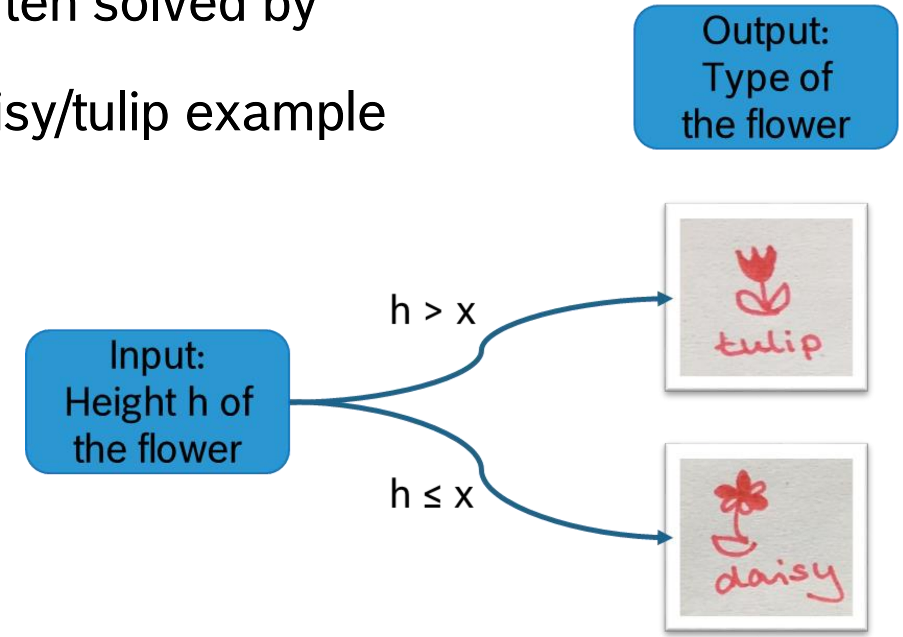
driver types (sporty, economic, ...)

image recognition (car, tree, pedestrian)

Machine Learning
Classification methods

The task of classification is often solved by

❖ Decision trees as in the daisy/tulip example



❖ Deep learning (later)

Machine Learning
Task: Regression

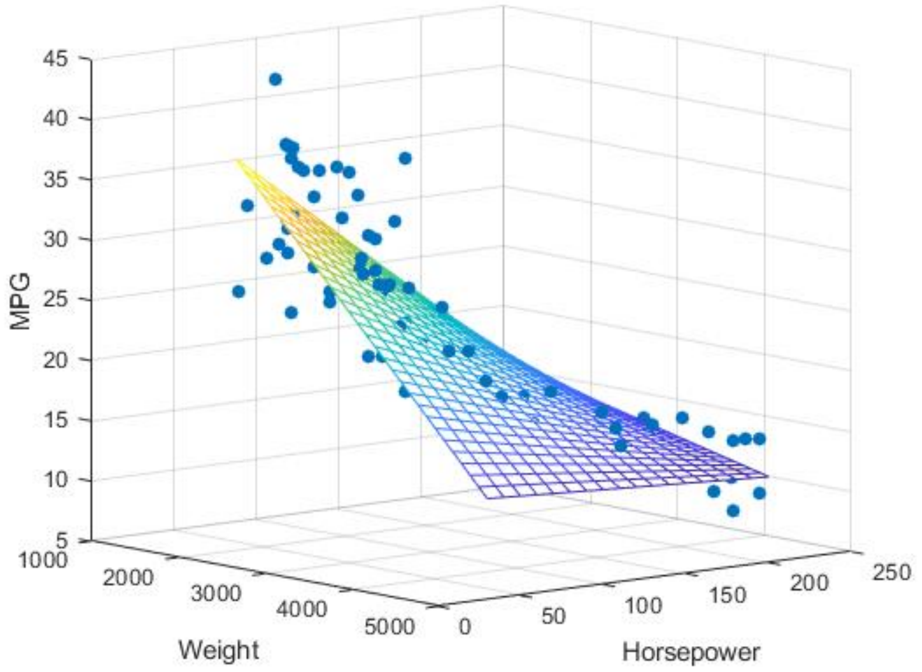
Target: numerical estimation of a variable due to certain input variables

Multiple linear regression graphic by MathWorks

Example: vehicle range (MGP)

depending on vehicle

weight and horsepower

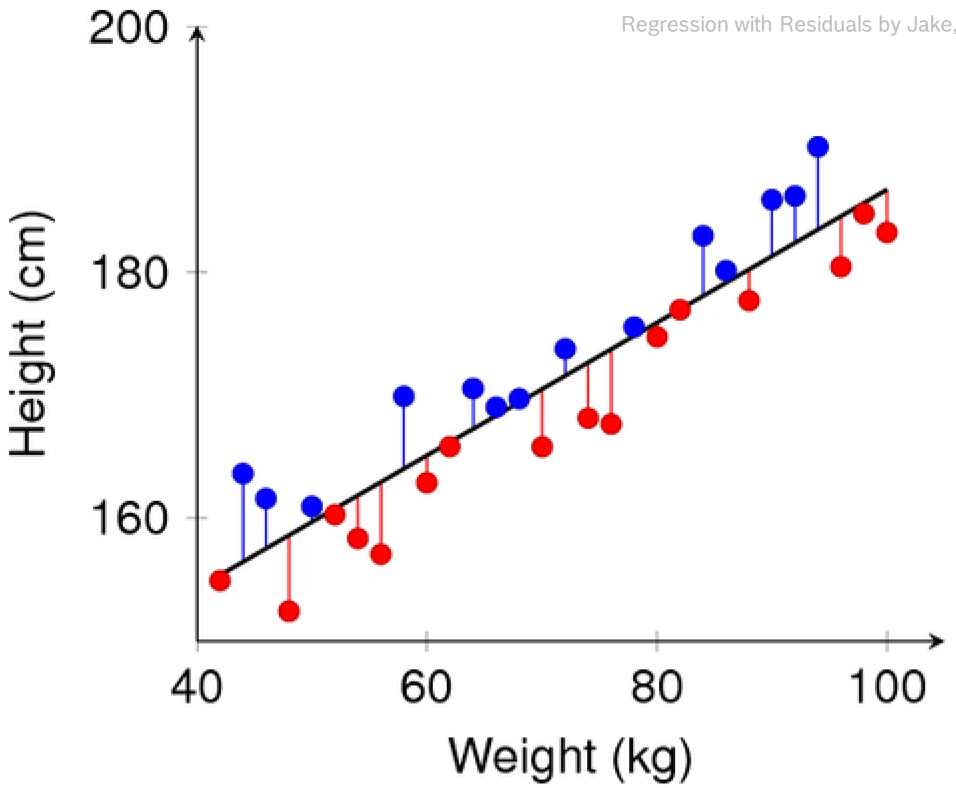


BOSCH

Artificial Intelligence
Regression methods

❖ Deep learning (later)

❖ Polynomial regression



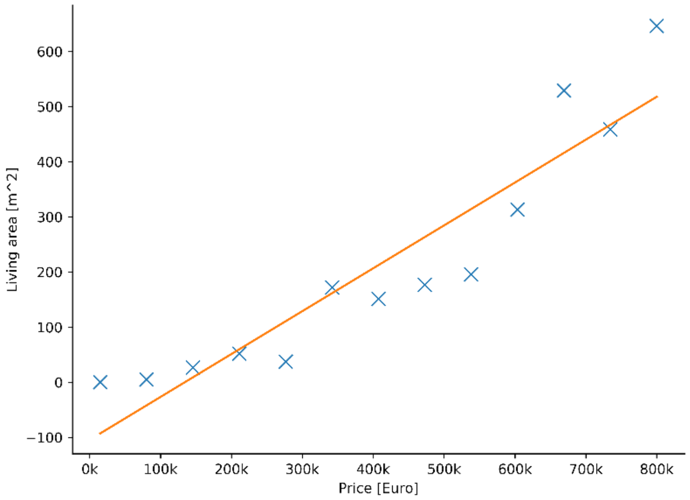
Regression with Residuals by Jake, CC BY 2.5

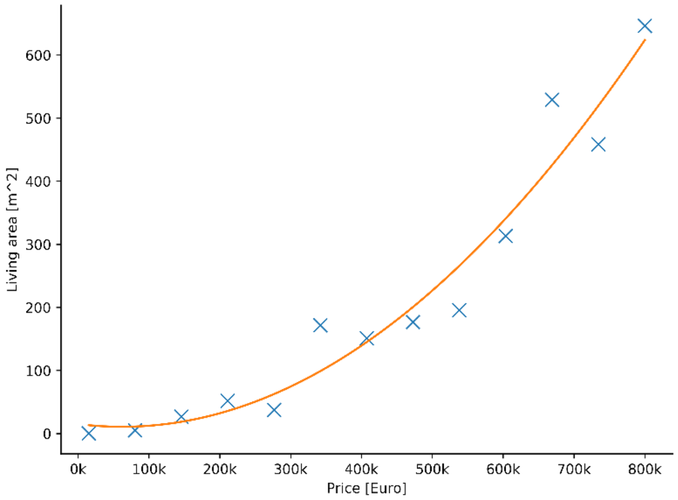
Machine Learning Concepts
Training vs Validation Error for Polynomial Regression

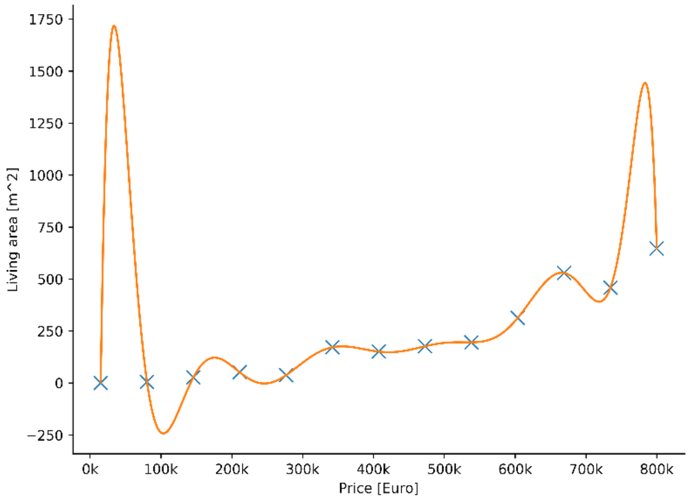
Linear regression
→ High training error
→ High validation error

Quadratic regression
→ Lower training error
→ Low validation error

Polynomial regression ^10
→ Low training error
→ High validation error





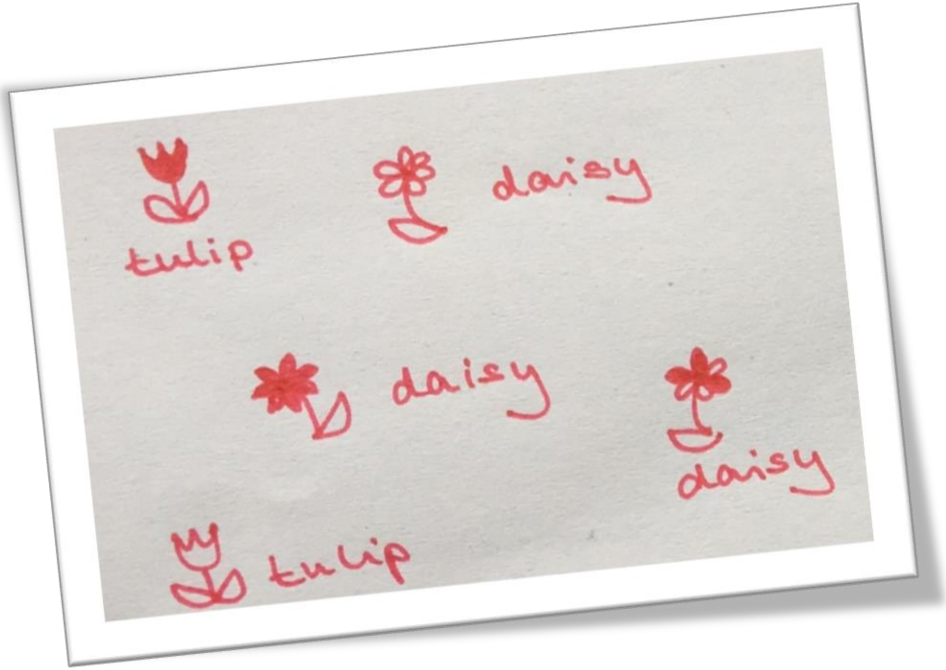


Machine Learning
Labeled Data

Supervised learning bases on labeled or tagged data

which provide the piece of information

you are searching for and the model shall provide.



BOSCH

Machine Learning
Labeled Data – a Billion $ Business


The Future of AI Depends on a Huge Workforce of Human Teachers

● VCs are investing heavily in startups that tap people to make artificial intelligence better at speaking, seeing, and driving.

By Matthew Hutson

For an autonomous car to recognize pedestrians and stop signs, it's typically fed thousands or millions of photos, all hand-labeled. To nail a conversation, a digital assistant needs to be told over and over when it's failed. And so Rubin spends 10 to 30 hours a week on her phone or computer evaluating search results and chat retorts through a site called Clickworker. Her income, generally $10 to $14 an hour, pays for part of her college commute from New Jersey and some of her ...

Bloomberg

Jacques Bughin, a director of the McKinsey Global Institute, speculates that the nine-figure market could hit $5 billion in five years… A radiologist tagging a medical image can cost up to $2000.

Providers of data labelling service

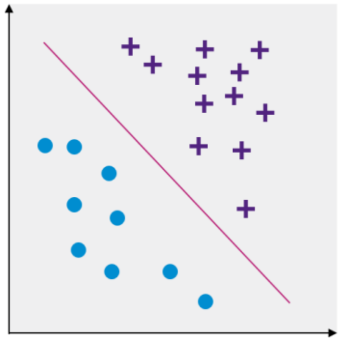
amazon mechanical turk beta, ALEGION, Mighty Ai, defined crowd, cloudfactory, CrowdFlower, clickworker

Reducing labelling efforts significantly is an ongoing research topic.

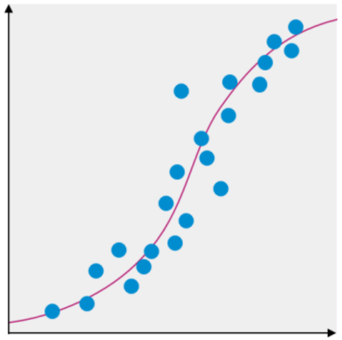
Machine Learning
Supervised vs. Unsupervised Learning

Supervised Learning

Build a predictive model using training data with known output (labels)



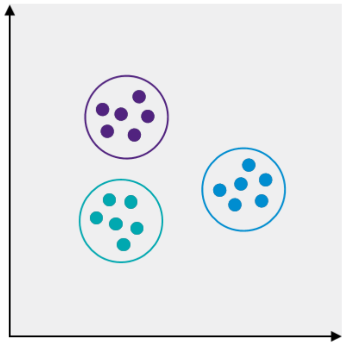
Classification



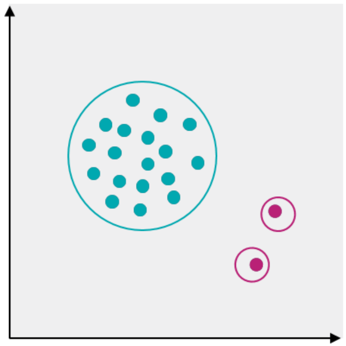
Regression

Unsupervised Learning

Describe “hidden” structure in unlabeled data



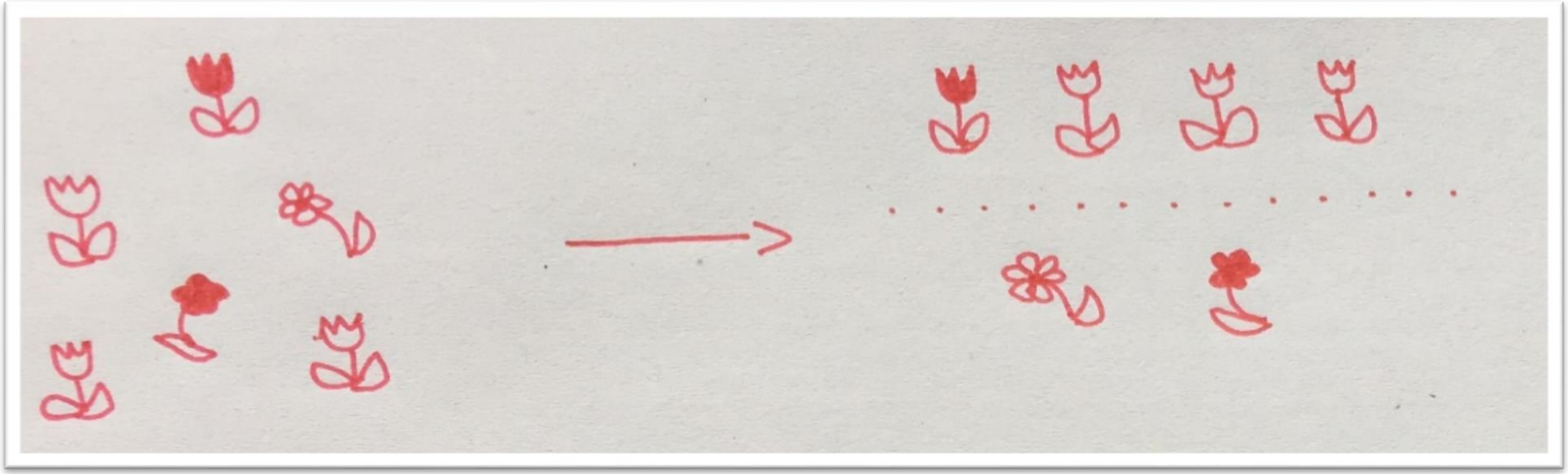
Clustering



Anomaly Detection

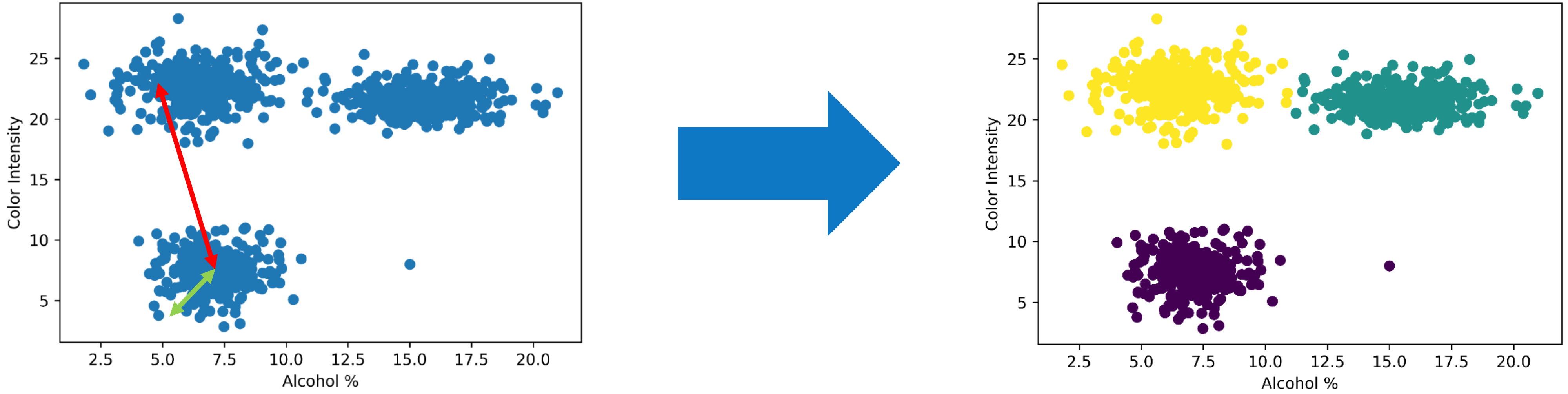
Machine Learning
Unsupervised Learning Task: Clustering

Target: find meaningful groups in your data and create “classes”



Machine Learning
Distance Based Clustering

Strategy: data in the same group are “closer” to each other than to data in other groups



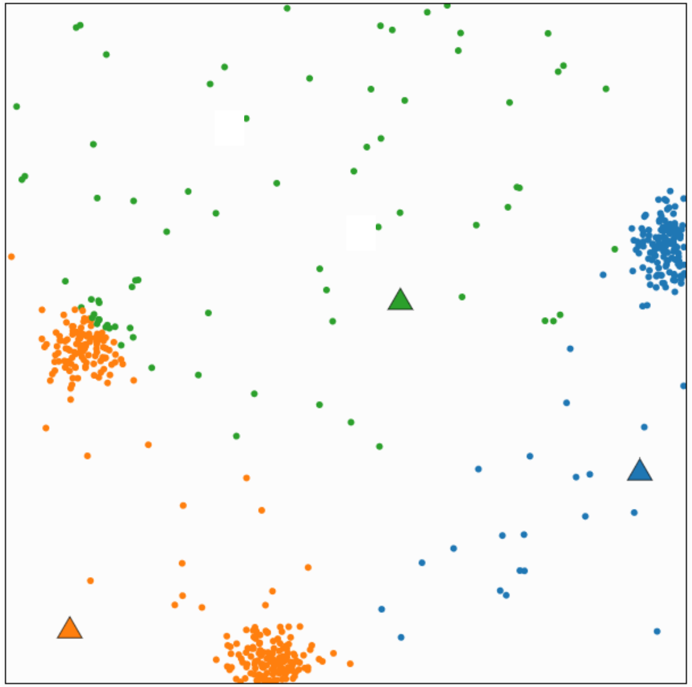
Machine Learning
K-means Algorithm

1. **Input:** *N* data points
2. **Initialize k centroids randomly**
3. **Repeat until converged:**
   - **Assign** each data point to the closest centroid:

$$k = \arg\min_{k} ||\vec{x}_n - \vec{c}_k||^2$$

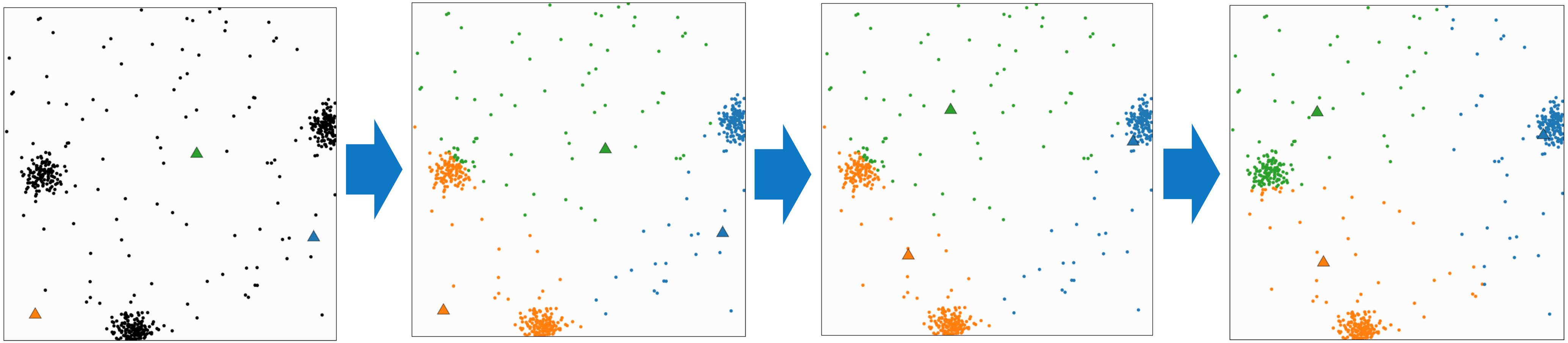
   - **Recompute** new centroids:

$$c_k = \frac{1}{|\mathcal{C}_k|} \sum_{n \in \mathcal{C}_k} \vec{x}_n$$



Karanveer Mohan, MIT License

Machine Learning
K-means Algorithm

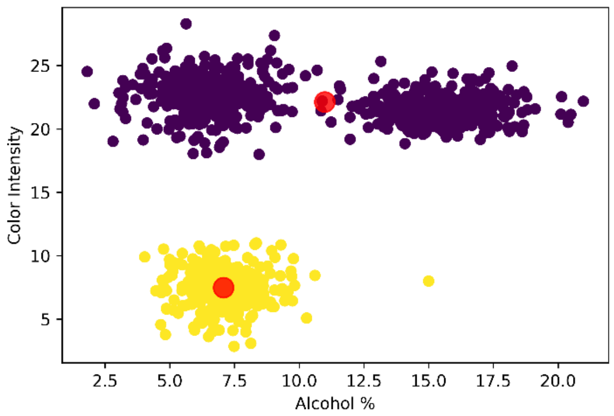


Karanveer Mohan, MIT License

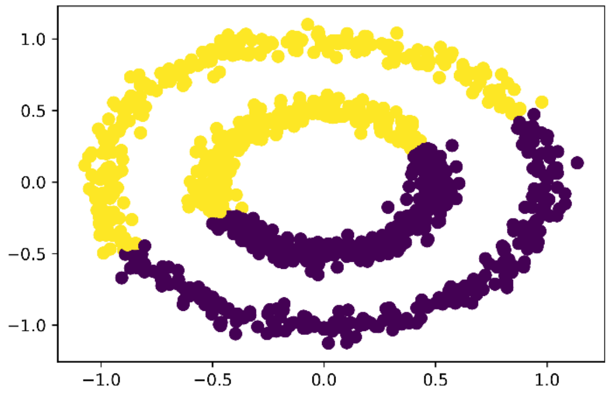
http://stanford.edu/class/ee103/visualizations/kmeans/kmeans.html

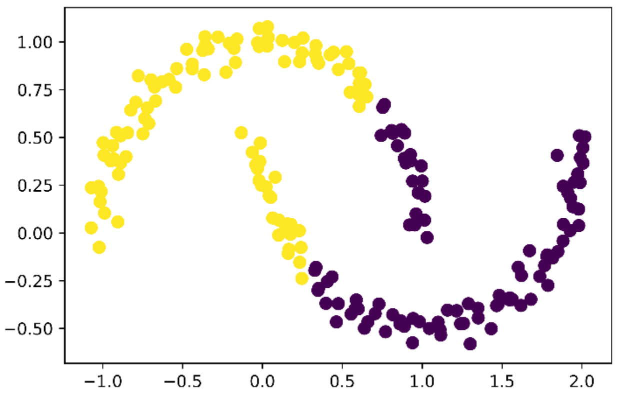
Machine Learning
Problems with K-means

Wrong number of clusters



Non-spherical distribution



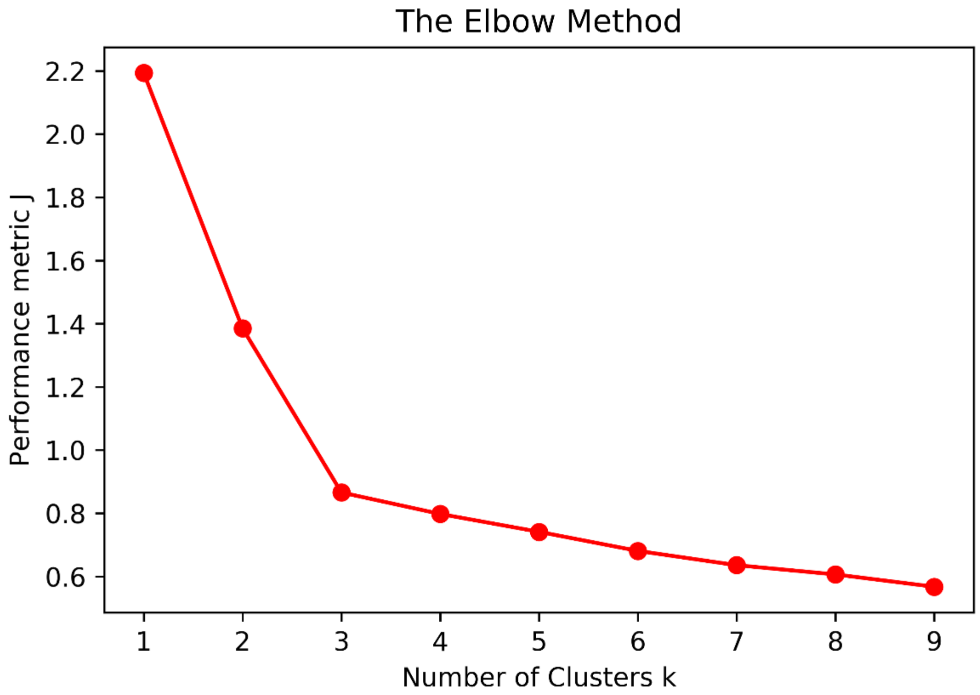


k-means is very sensitive to outliers

 Internal | CR/PJ-AI-E & CI/OST | 2019

Machine Learning
Elbow Method

How many clusters to choose?



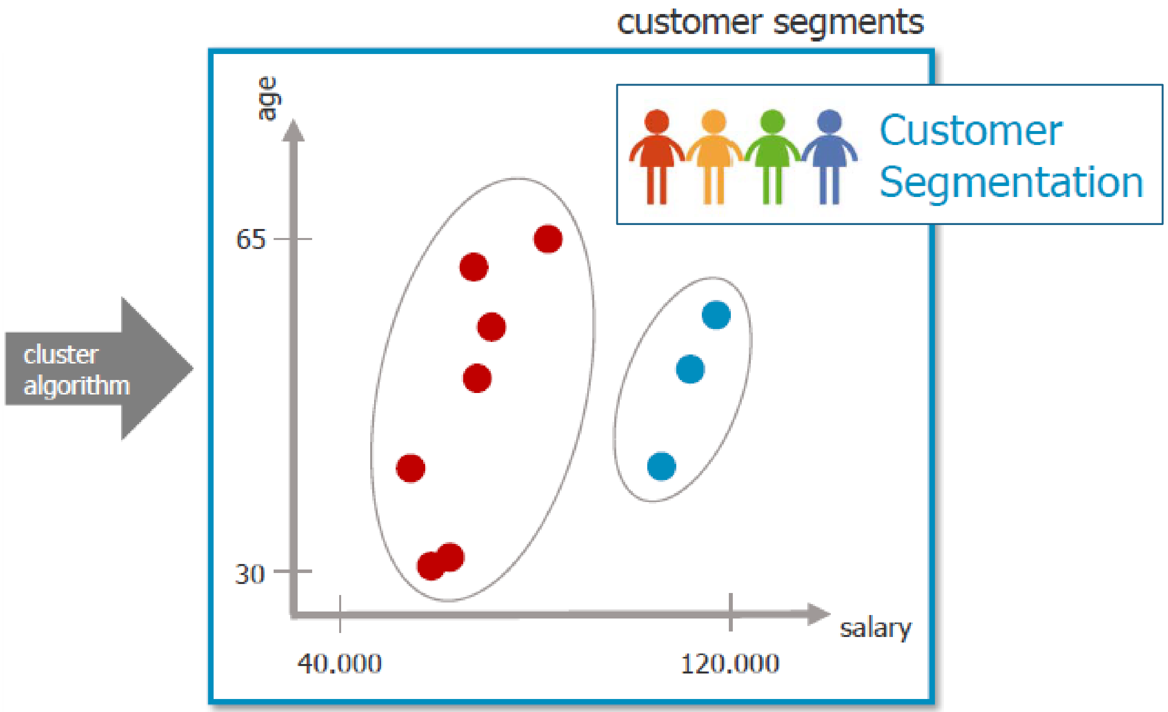
$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||\vec{x}_n - \vec{c}_k||^2$$

Machine Learning
Clustering Application

Example: customer segmentation

customer data

ID	Age	Salary	Address (Town)	
1	30	55.000	Stuttgart	...
2	31	57.000	Frankfurt	...
3	56	120.000	Munich	...
4	55	65.000	Stuttgart	...
5	47	66.000	Hamburg	...
6	65	76.000	Hamburg	...
7	48	105.000	Berlin	...
8	42	49.000	Köln	...
9	42	99.000	Munich	...
10	63	62.000	Berlin	...
...	...		...	...



Customer Segmentation by Prof. Mitschnang, Universität Stuttgart, IVPS

Machine Learning
Unsupervised Learning Task: Anomaly Detection

Depositphotos, Bosch License

Target: detect outliers in your data

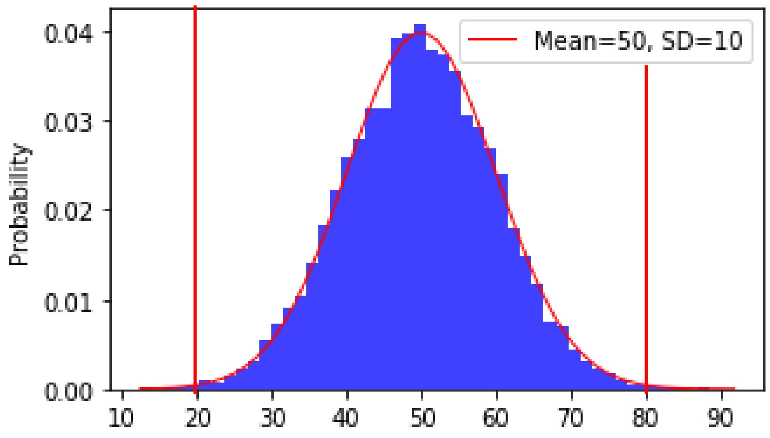
Examples:

- unrealistic sensor data
- technical problems in production
- fraud detection
- cyber security

Machine Learning
Anomaly Detection Methods

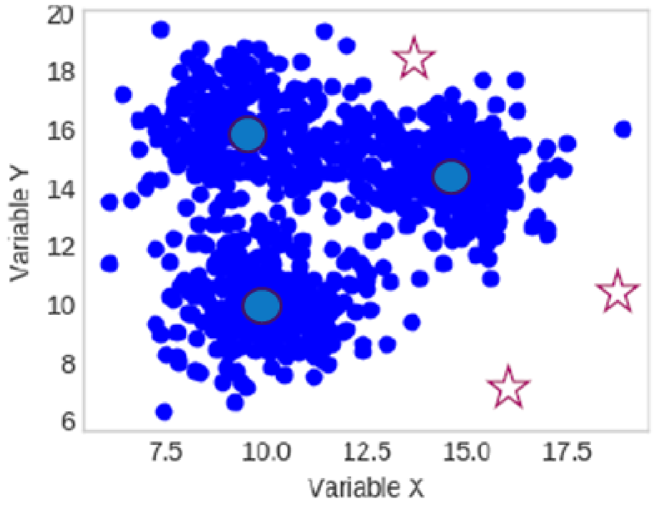
Statistical Methods

Outliers are points with a certain deviation from the mean value, often 3σ (standard deviation)



K-Means

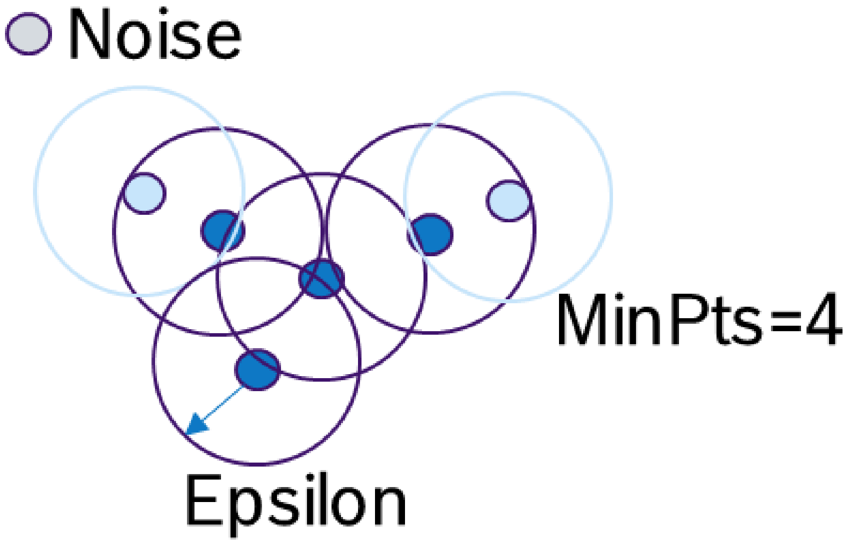
Outliers are clusters with very few points and points far away from the cluster centers



Machine Learning
Density-Based Clustering

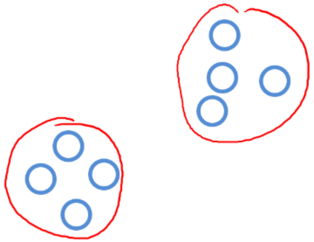
- *p* is a core point if
  ≥ minPts points are within distance ε
- *q* is directly reachable from a core point *p* if point *q* is within distance ε from *p*
- *q* is reachable from *p* if there is a “chain” of directly reachable points connecting p and q

- Reachable points form a cluster
- All other points are outliers



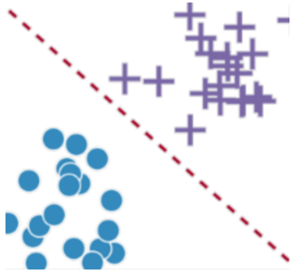
Machine Learning
Overview of 4 important ML Tasks

Clustering



... find meaningful clusters in your data to create “classes”...

Classification

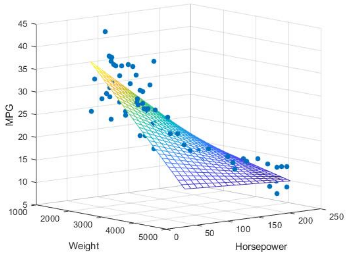


... sort your data into predefined classes ...

Anomaly detection

... detect outliers in your data...

Regression



... calculate numerical estimations from numerical input data...

BOSCH

Machine Learning
ML Tasks: Exercise

Bosch Media Space

Robert is a line manager.

In order to detect scrap earlier, he wants to predict whether a part will pass or fail at the end of line test based on measurements in previous parts of the production cycle.

Which task will he solve?

Classification, Clustering, Regression, or Anomaly detection?

Machine Learning
ML Tasks: Exercise

Bosch Media Space

Andrew is a logistics manager.

His plant uses 500 different packaging materials.

He would like to reduce it to 20 standardized boxes based on product type, form and dimensions.

Which task will he solve?

Classification, Clustering, Regression, or Anomaly detection?

BOSCH

Machine Learning
Sweet Spot of Machine Learning

Work on data based machine learning

if you have ...

- ❖ A clear business case
- ❖ Sufficient amounts of data
  carrying the relevant information
- ❖ Sufficient computational resources
- ❖ Knowledge, experience, and/or partners that help you get started
- ❖ Patience to go through trial and error process

Depositphotos, Bosch License

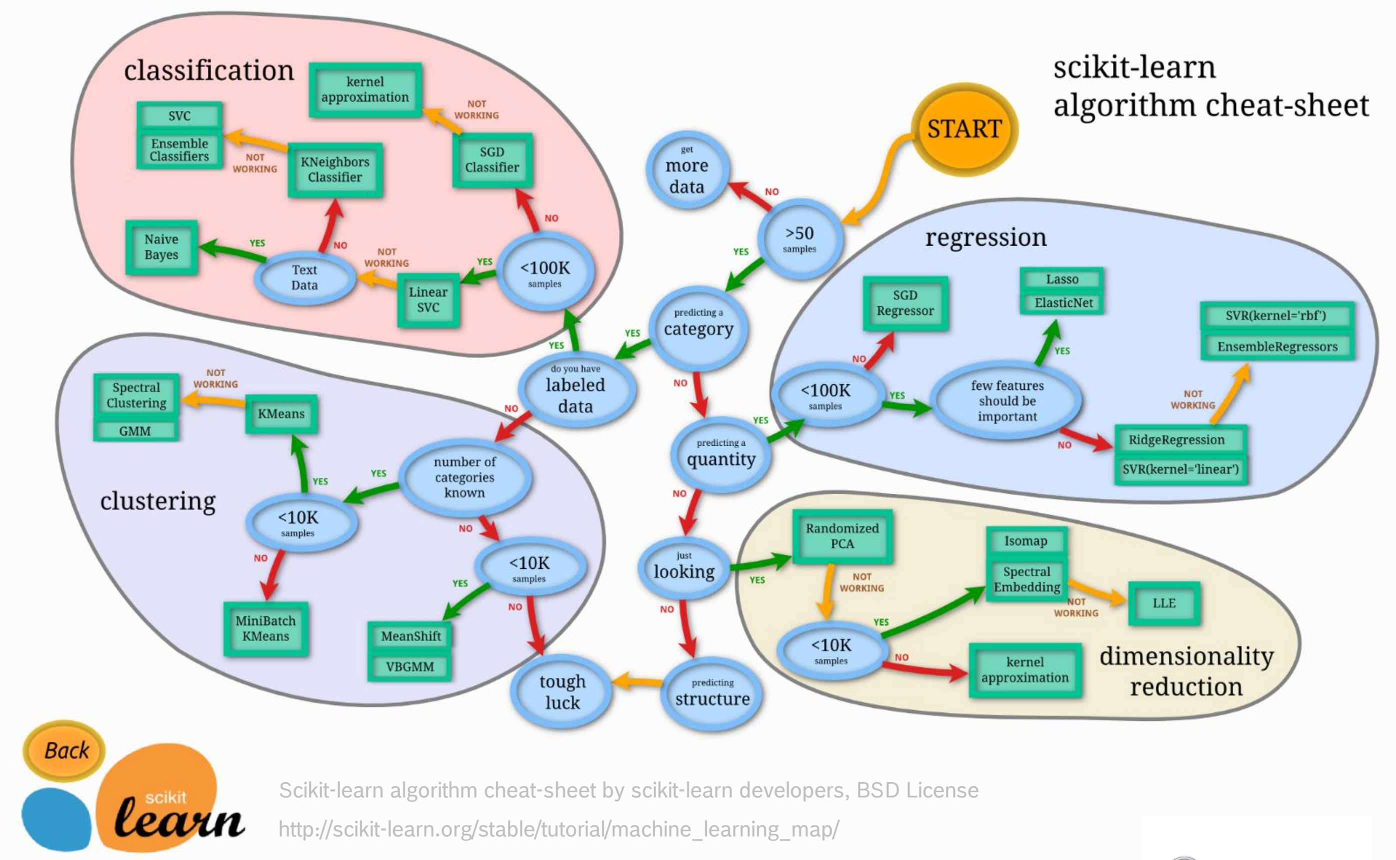
Machine Learning
Helpful resources for Machine Learning

❖ Frameworks, libraries & tutorials

- Scitkit-learn
- H2O
- TensorFlow
- Microsoft CNTK

❖ Guidelines:

- ML decision tree (Scikit-learn)
- Method pros & cons (recast.ai)



Scikit-learn algorithm cheat-sheet by scikit-learn developers, BSD License
http://scikit-learn.org/stable/tutorial/machine_learning_map/

Machine Learning
Machine Learning Strategies

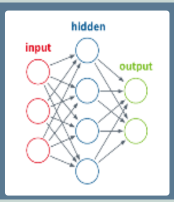
Tasks

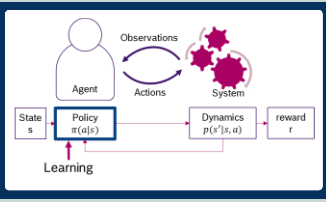
Strategies

Methods

Machine Learning Strategies

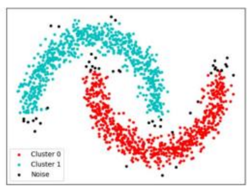
Supervised Learning





Reinforcement Learning

Unsupervised Learning

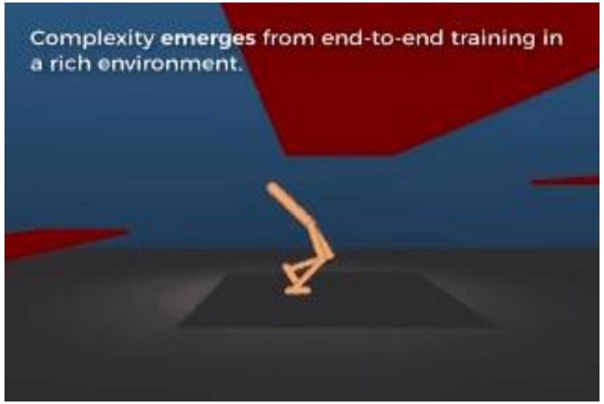


Machine Learning
What is Reinforcement Learning?

- Reinforcement learning is the field of sequential decision making under uncertainty
- Goal: maximize the long-term cumulative reward

Atari video games

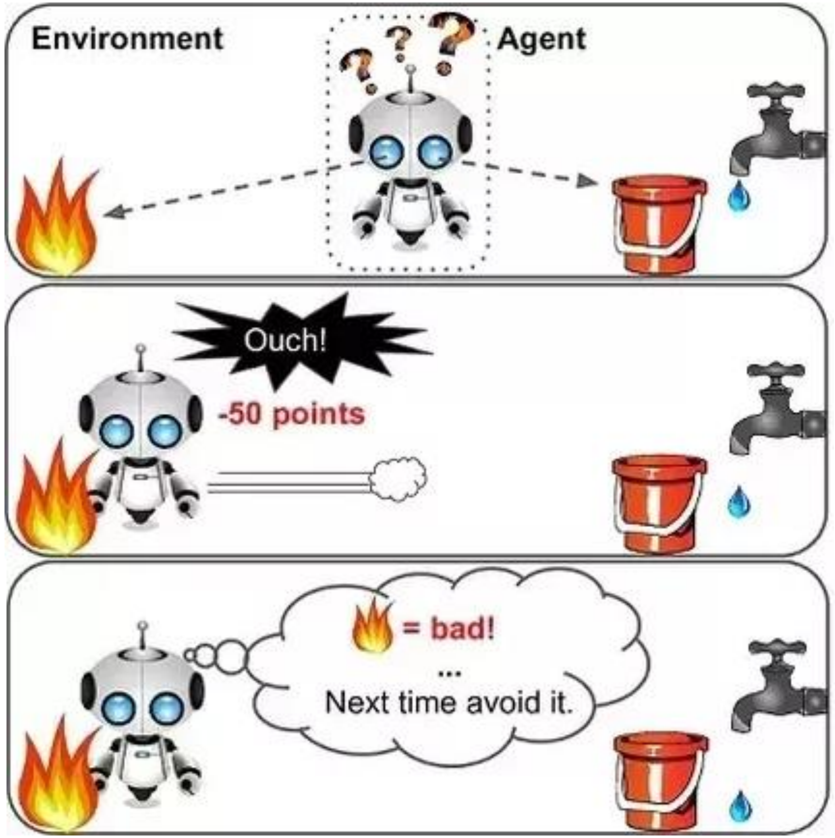
Robotics



Learning to walk

BOSCH

Machine Learning
ML Strategy: Reinforcement Learning



- System perceives environment
- Decides what to do according to its policy

- System interacts with environment
- Gets a penalty or reward

- System learns = changes its policy

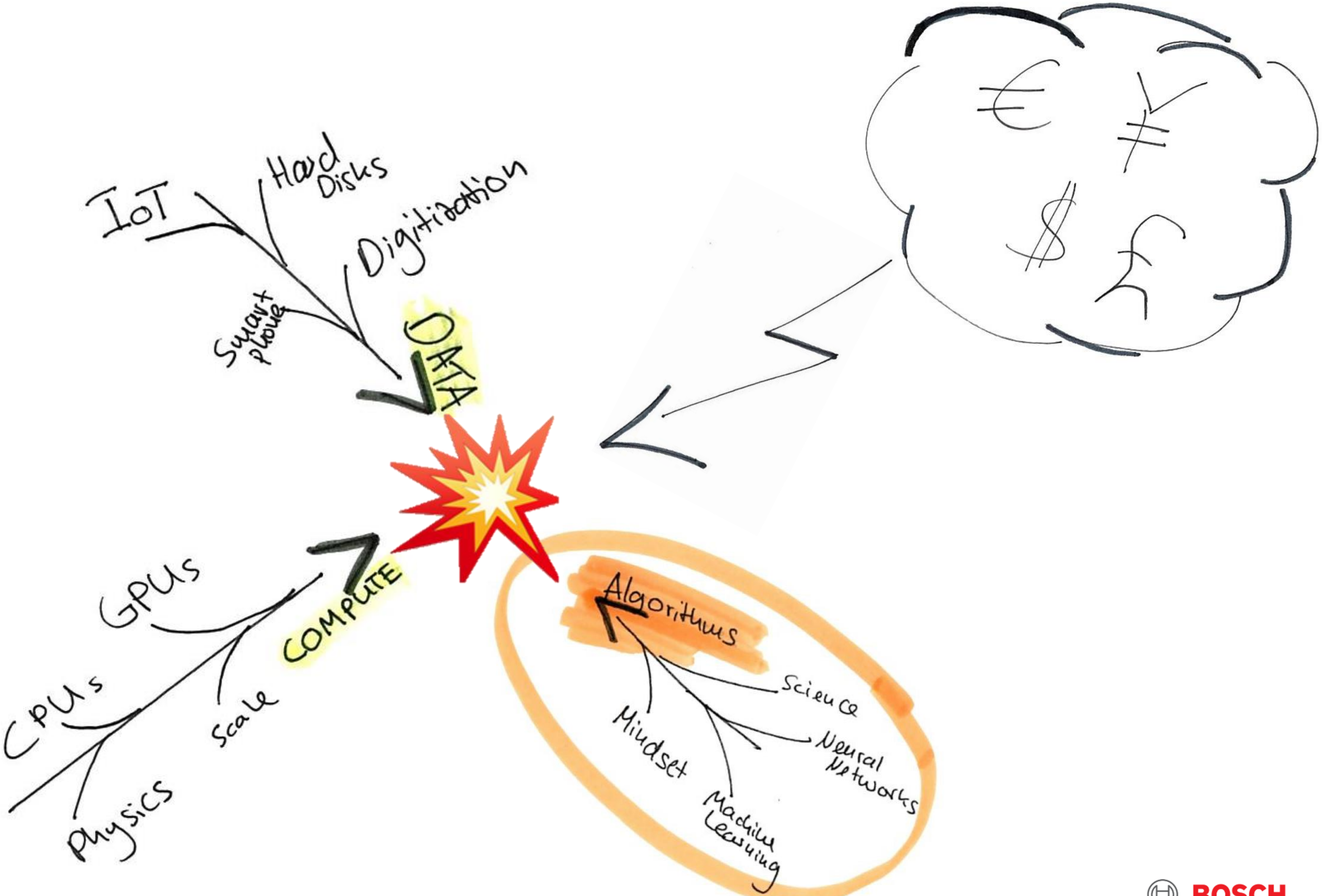
Machine Learning
ML Strategy: Reinforcement Learning

Depositphotos, Bosch License

- Not every action provokes reactions in the environment that provide information about how successful the policy is!
- Example: Games like chess
- Reinforcement learning always needs interaction with the environment and/or simulation!

DEEP LEARNING

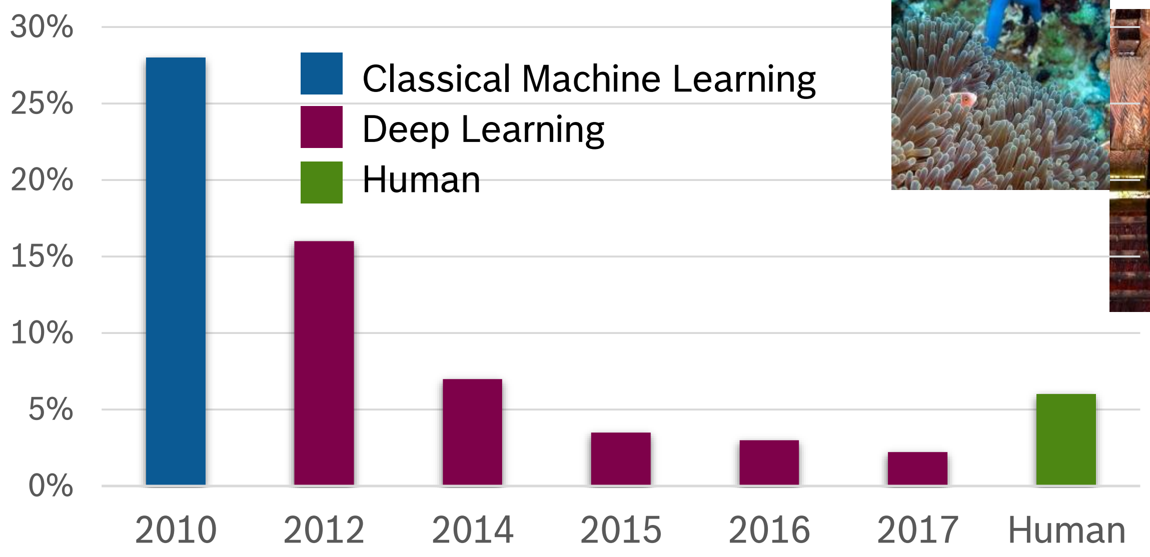
Why AI?
Why now?



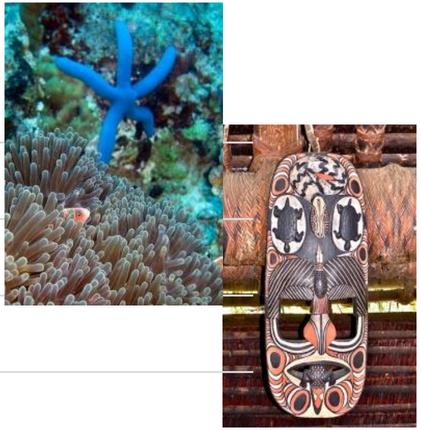
BOSCH

Deep Learning
Deep Learning Breakthroughs

Error rate [%]



Depositphotos, Bosch License

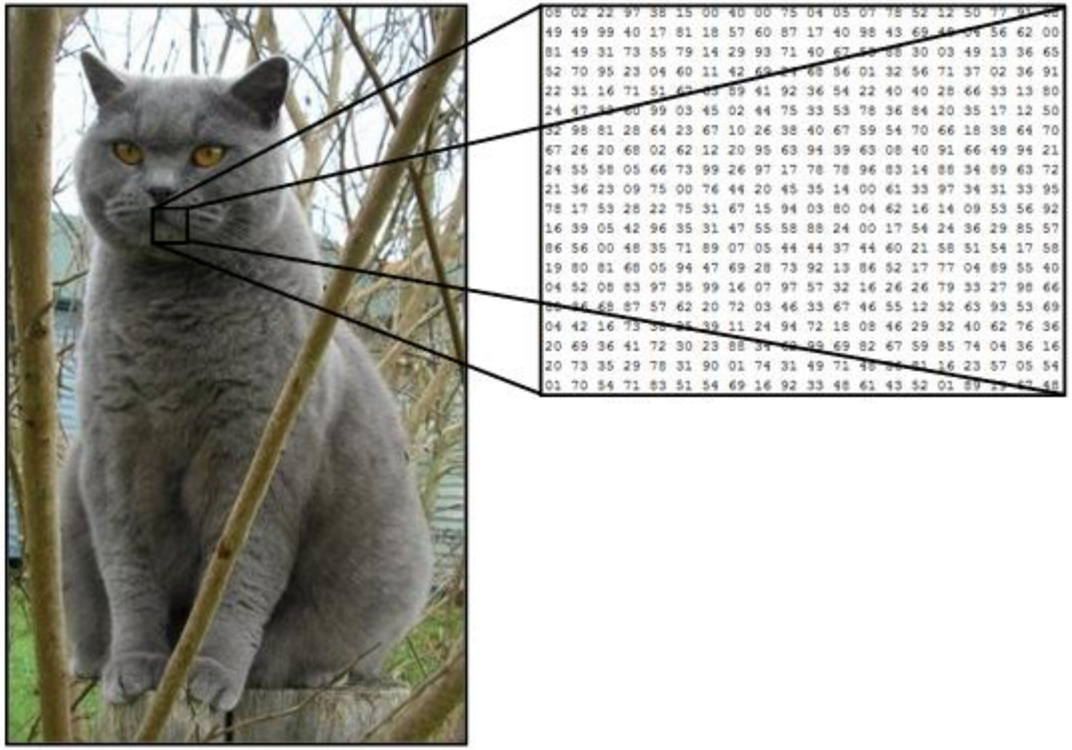


ImageNet Competition for image recognition

- Significant improvement due to deep learning
- Same story in text & speech recognition
- Why? And how?

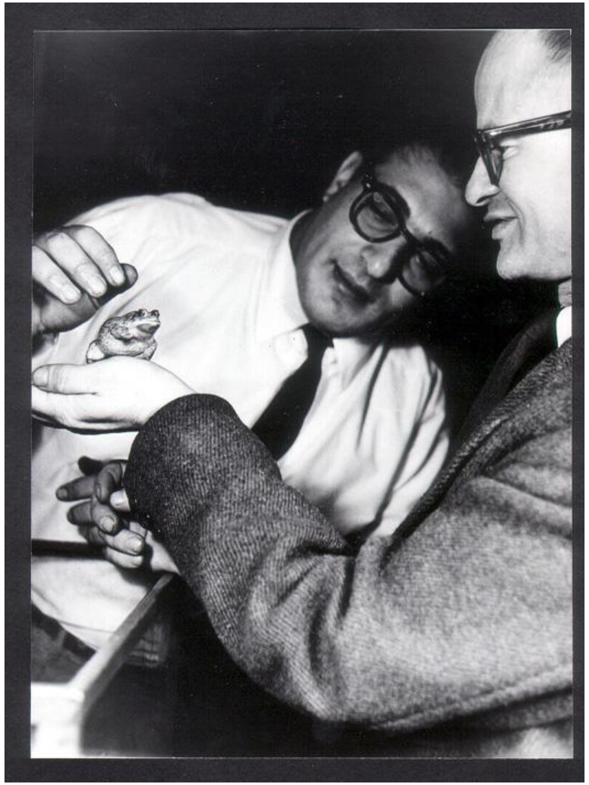
BOSCH

Deep Learning
Challenges of image recognition



Copyright (c) 2015 Andrej Karpathy, The MIT License (MIT)

Deep Learning
The “Magic” of our Human Brain



Lettvin Pitts by lapx86, CC BY-SA 3.0

Image recognition is tough for classical machine learning, while it is natural and easy for humans.

How do humans learn to recognize objects?

In 1943, Walter Pitts and Warren McCulloch proposed the first artificial neural network to imitate the human brain’s process of learning tasks from experience and examples!

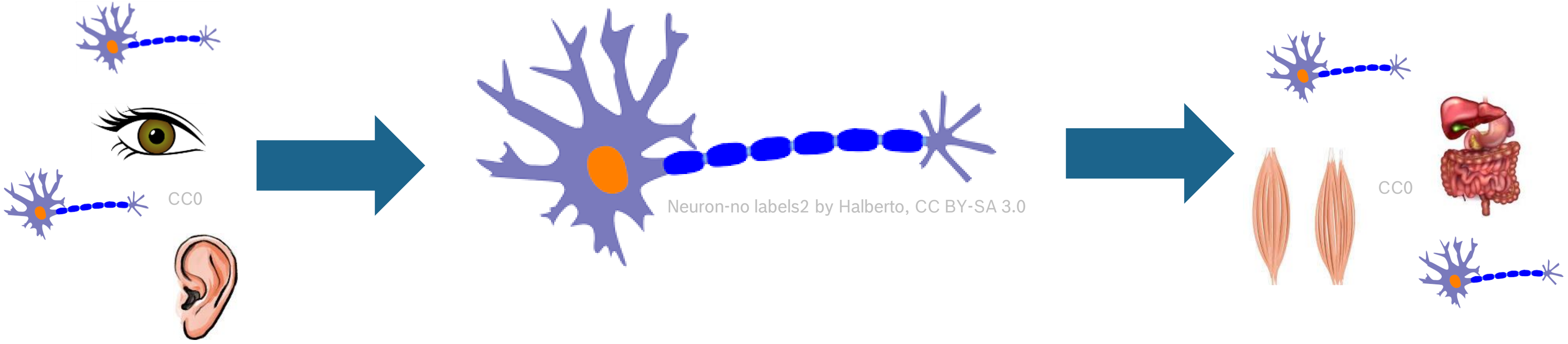
BOSCH

Deep Learning
Neurons

Our brains consist of neurons

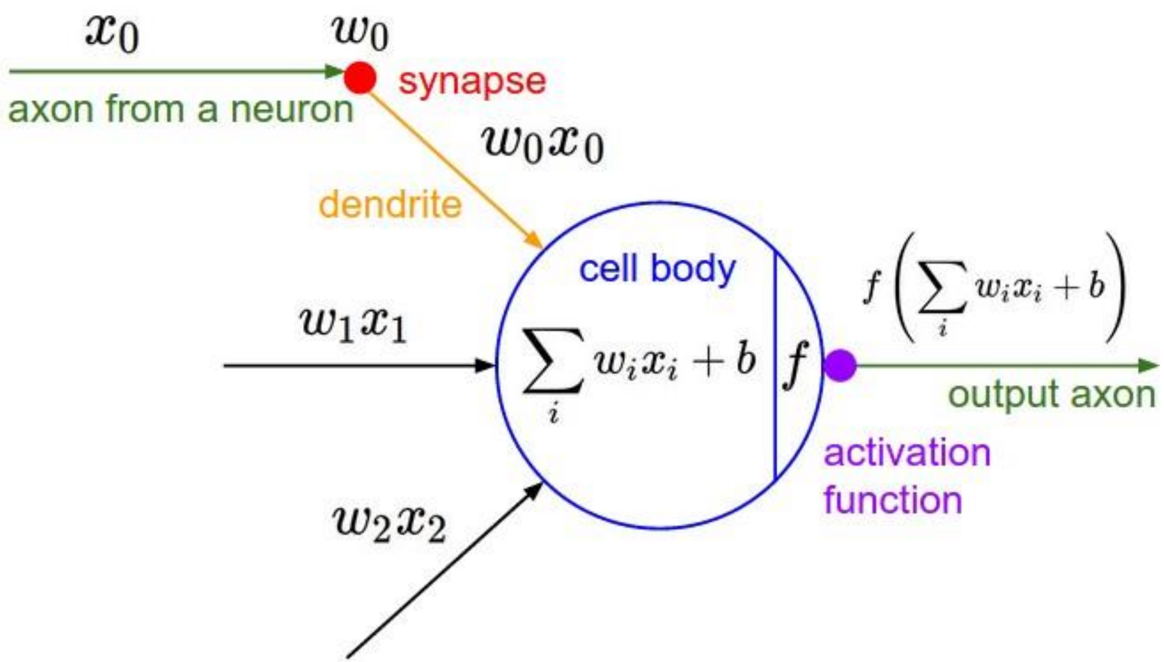
that process input signals from other neurons or “sensors” like our eyes

and pass output signals to other neurons or “actors” like our muscles!

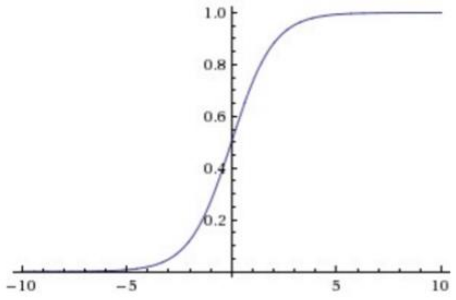


BOSCH

Deep Learning
Artificial Neurons

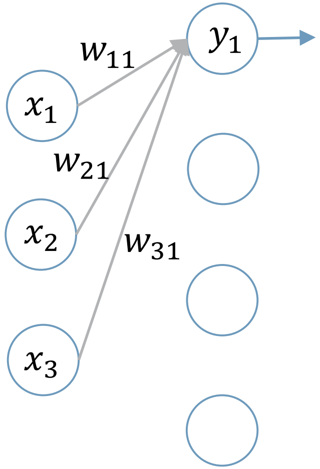


Non-linear activation function f(x):



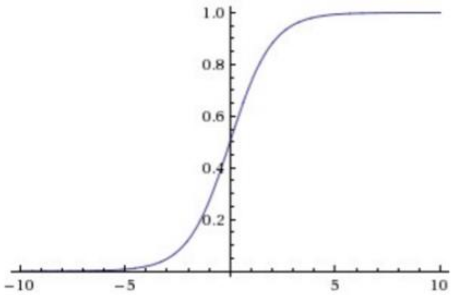
Sigmoid

Deep Learning
Connected Artificial Neurons



$$y_1 = f(w_{11}x_1 + w_{21}x_2 + w_{31}x_3)$$

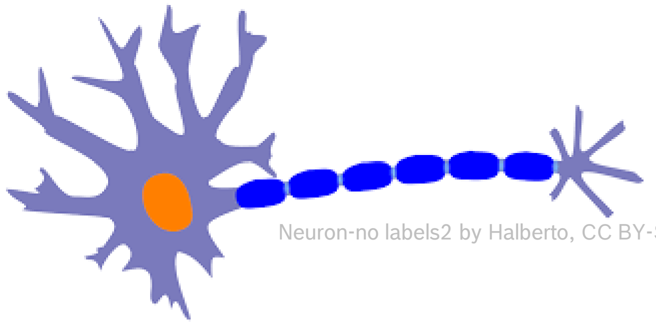
Non-linear activation function f(x):



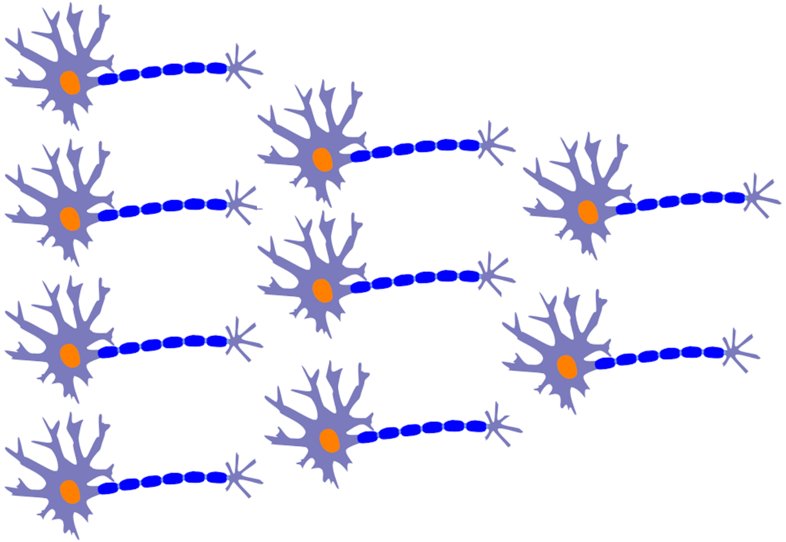
Sigmoid

Deep Learning
Neural Networks

A single neuron can only learn simple relations.



Neuron-no labels2 by Halberto, CC BY-SA 3.0

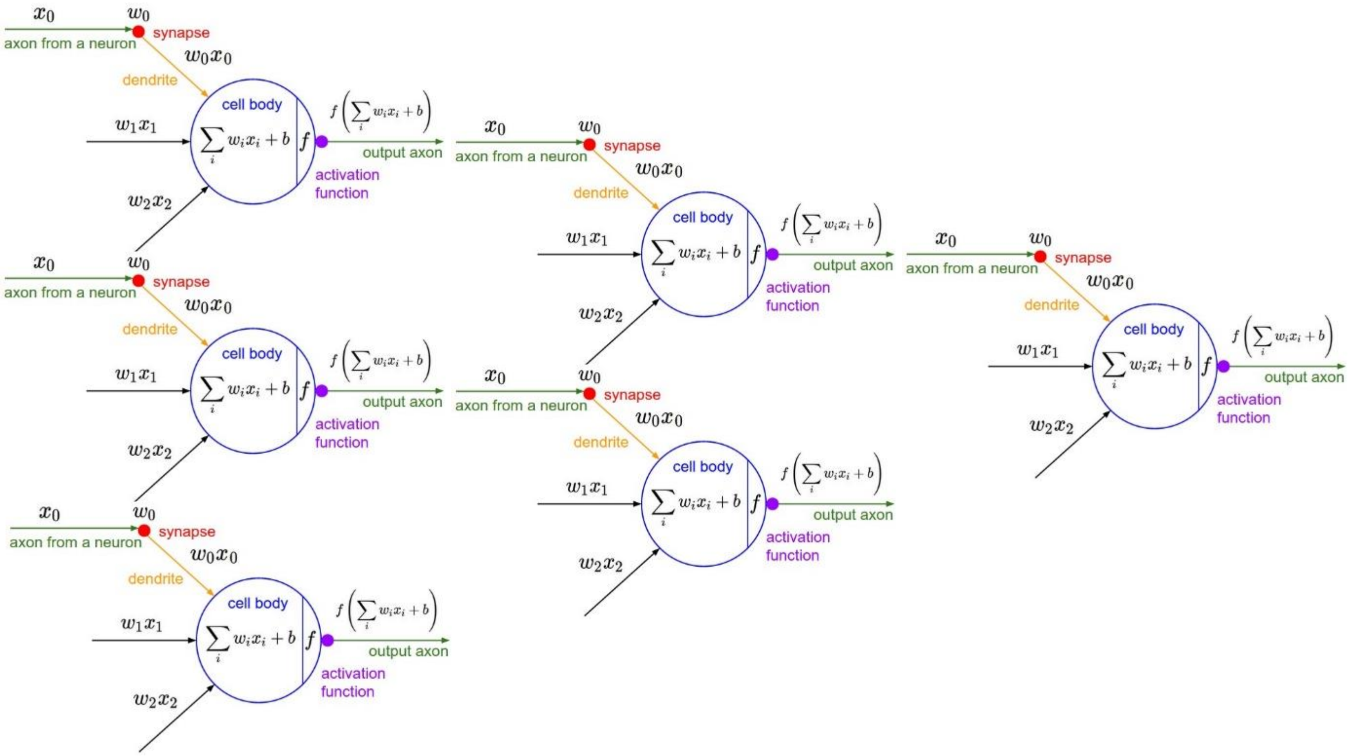


Neural networks of many

connected neurons can

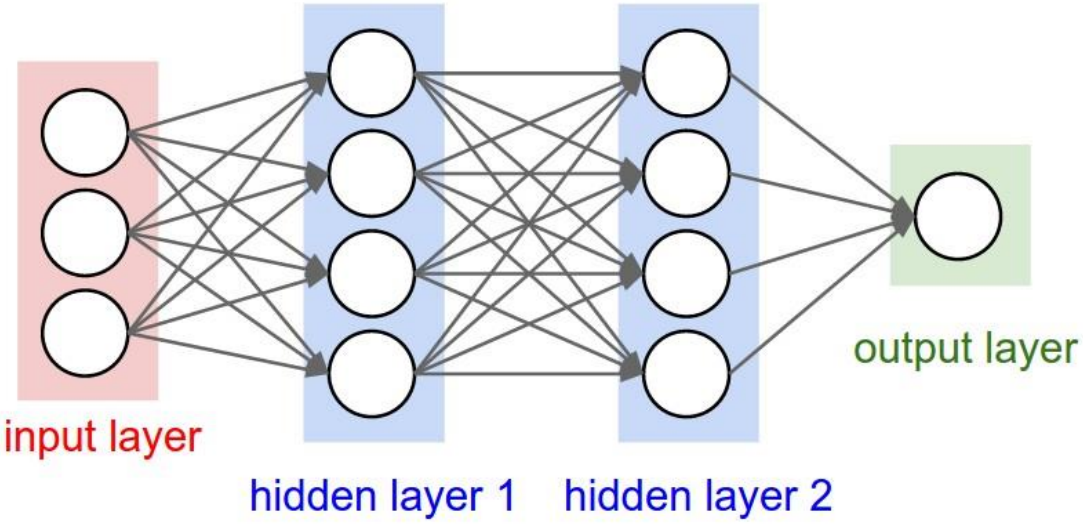
learn arbitrarily complicated relations!

Deep Learning
Artificial Neural Networks



Single neuron in neural network using c++ by Stackoverflow,, CC BY-SA 3.0

Deep Learning
Artificial Neural Networks



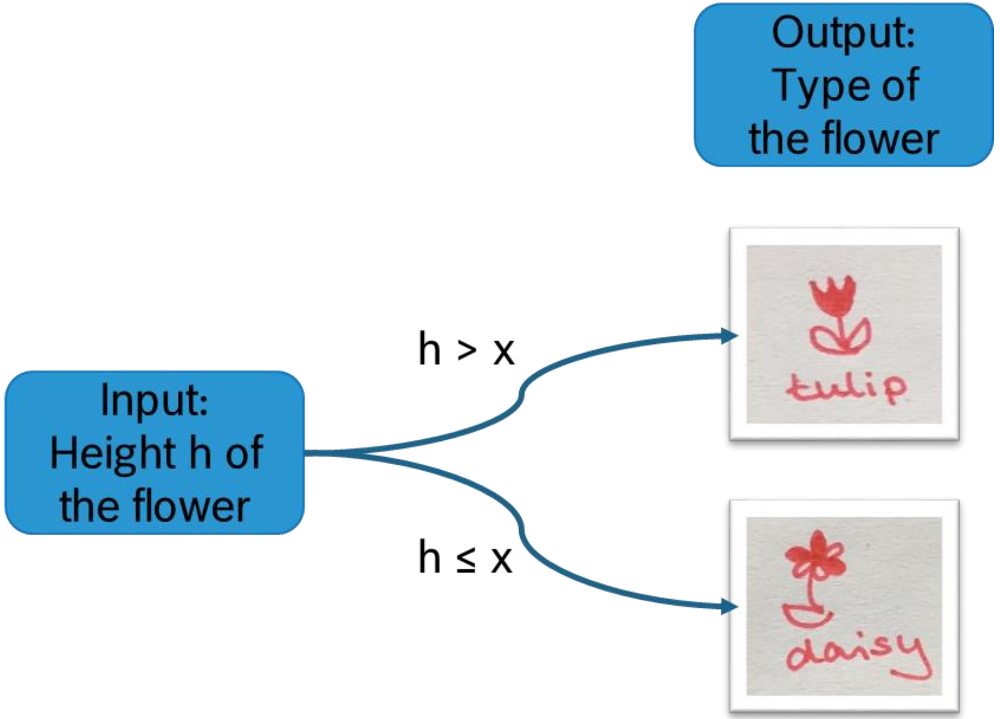
Copyright (c) 2015 Andrej Karpathy, The MIT License (MIT)

- Modern artificial neural networks combine artificial neurons in layers
  - ➢ No connections between neurons of the same layer
  - ➢ No connections skipping a layer

Why are artificial neural networks so powerful?

Deep Learning
Artificial Neural Networks

Remember features as our input variables for classification or regression:



What are relevant features for

- Image recognition?
- Sentiment analysis?

Deep Learning
Image recognition example

Example: Iris data set

Input: flower photo → Hand-designed features → Learning from features → Output: flower type

length & width of sepals & petals

no satisfying result

Kosaciec szczecinkowaty Iris setosa by Radomil, CC BY-SA 3.0

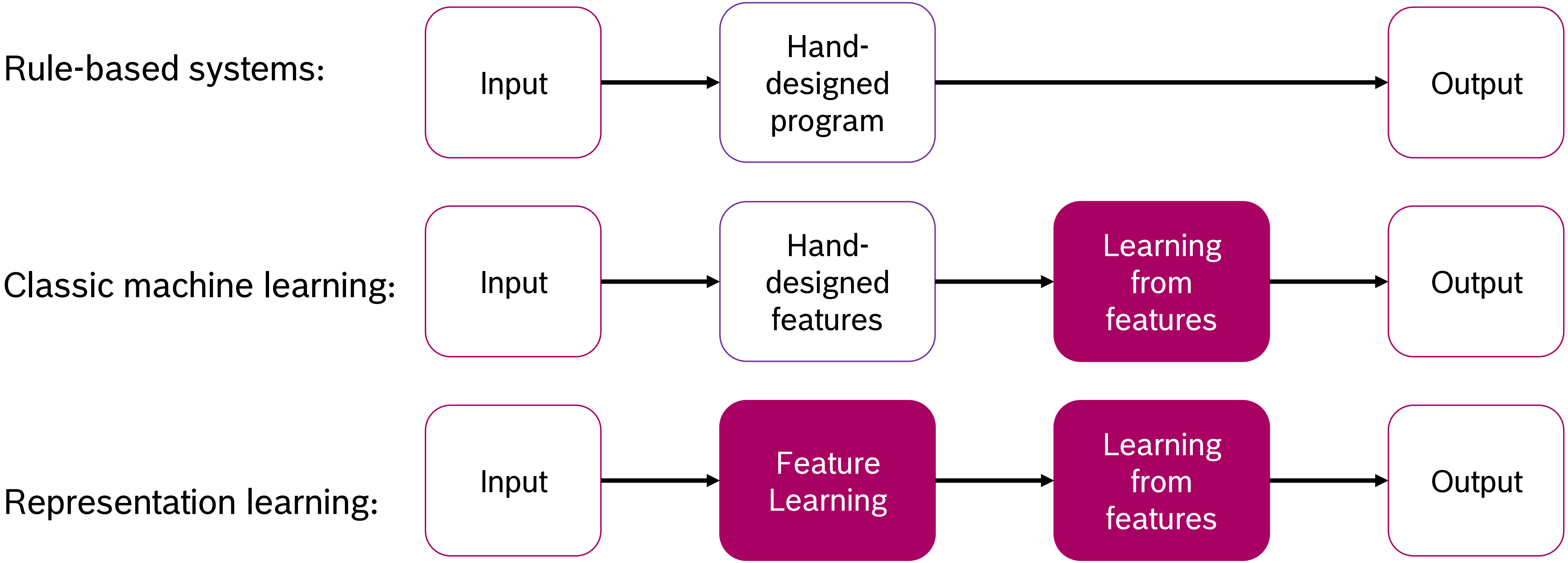
Blue flag flower close-up by Danielle Langlois, CC-BY-SA 3.0

Iris virginica by Frank Mayfield, CC-BY-SA 2.0

Input: flower photo → Feature Learning → Learning from features → Output: flower type

BOSCH

Deep Learning
Representation Learning



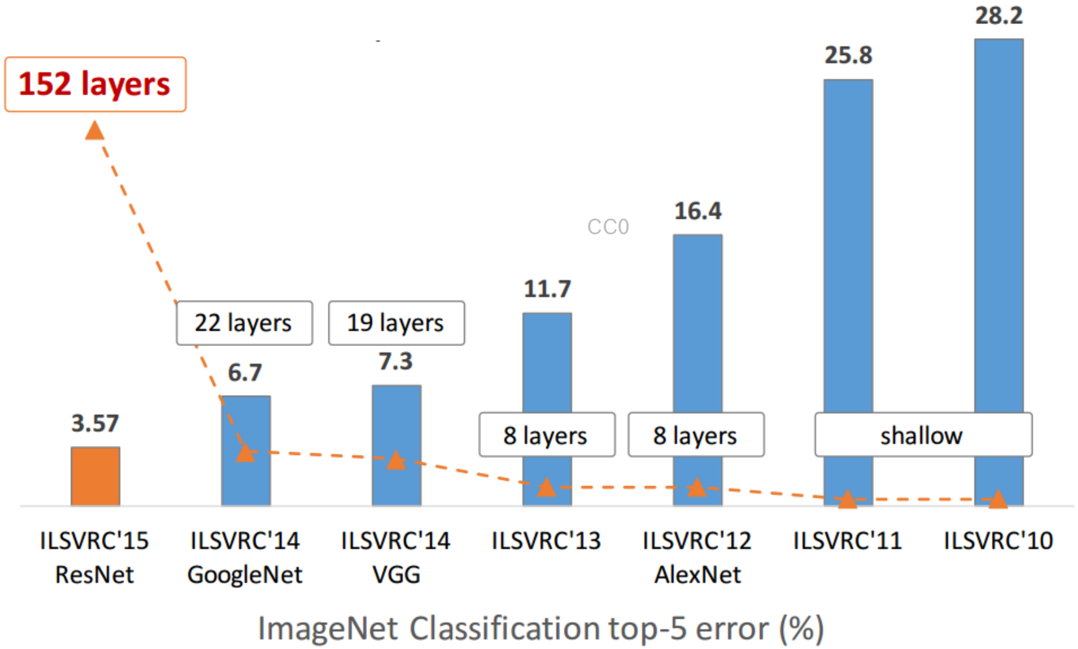
Representation learning: Determining the relevant features for a task algorithmically!

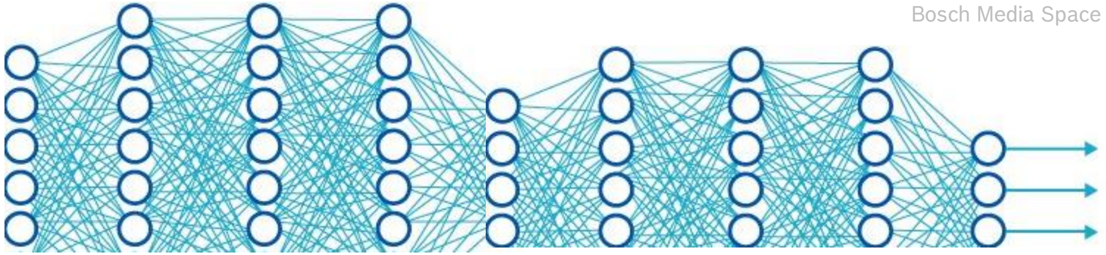
Deep Learning
Feature & Deep Learning

Artificial neural networks (ANN) can learn features!

The more hidden layers, the more complex features can be learned.

This is called deep learning.





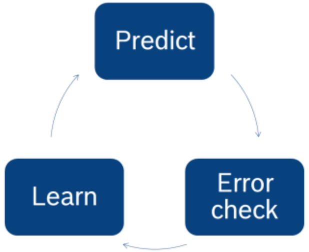
BOSCH

Deep Learning
Supervised Learning Workflow

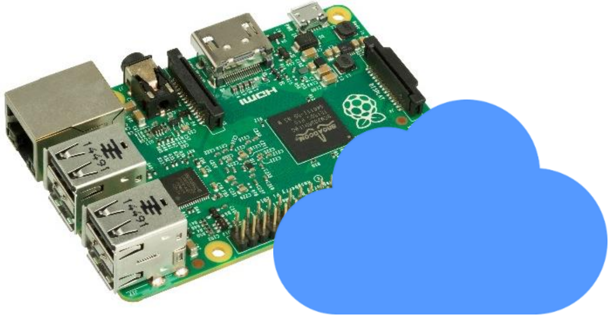
Task

Classification or Regression

Model training



Deployment



Flower height [cm]	Flower type
13	A
18	A
7	B
9	A
6	B

Labeled training data

Validation data

Flower height [cm]	Flower type
19	A
9	A
5	B
14	A
11	B

Deep Learning
Example: Intelligent Washing Machine Classifier

Bosch Media Space

You develop the next generation Bosch washing machine using image recognition to classify laundry:

30° laundry

60° laundry

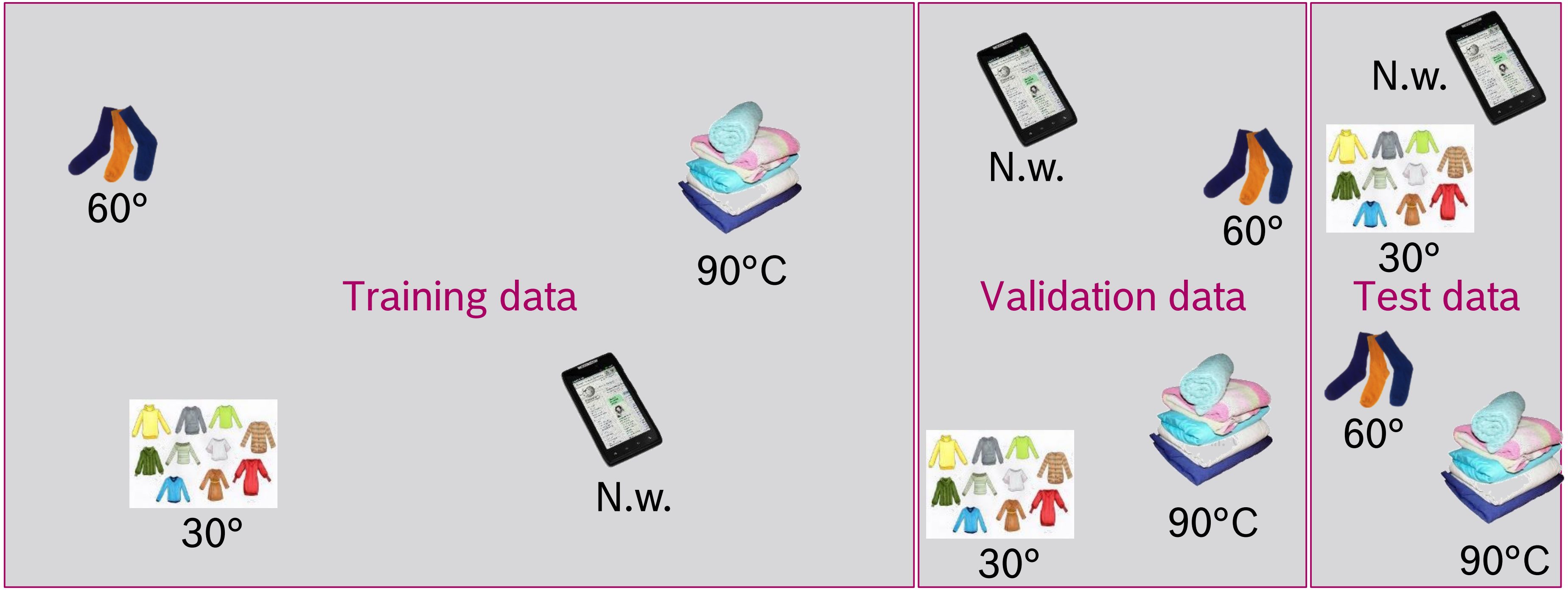
90° laundry

Non washable

All: CC0

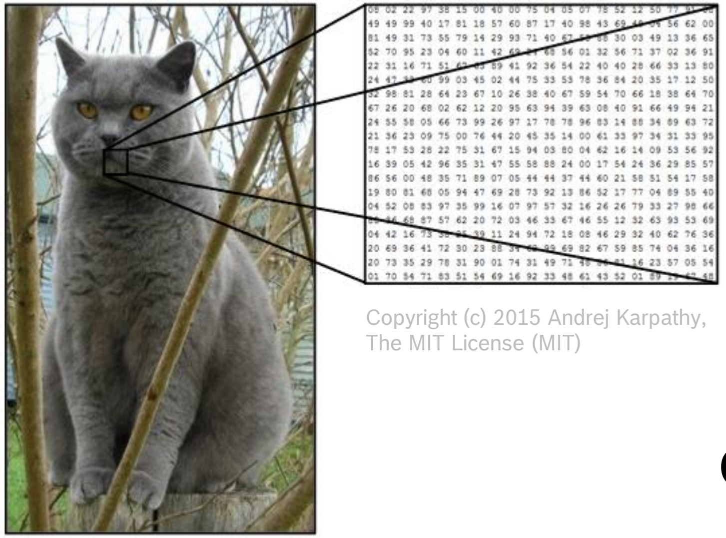
Deep Learning
Data collection & split

1. Collect pictures of laundry, label and preprocess them
2. Split your data:



Deep Learning
Network Setup

How many input and output neurons do you need?



Input: Pixels

Choose a standard network

for your kind of problem,

e.g. convolutional NN

Output: 4 classes

30° laundry

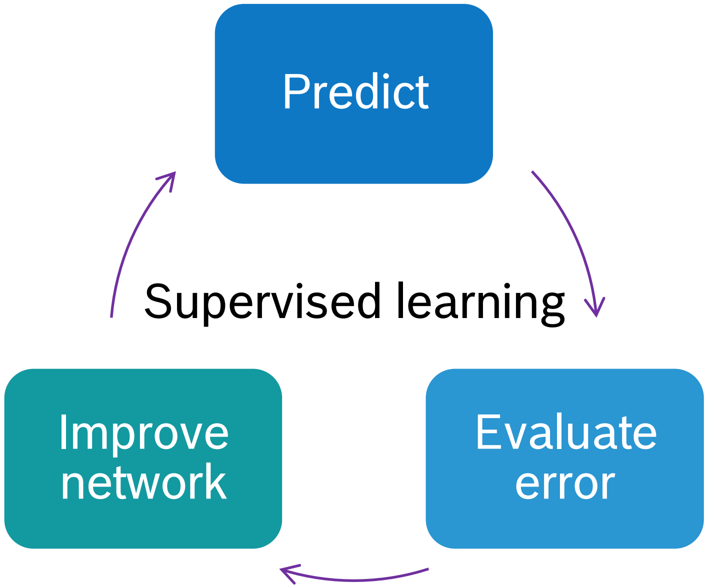
60° laundry

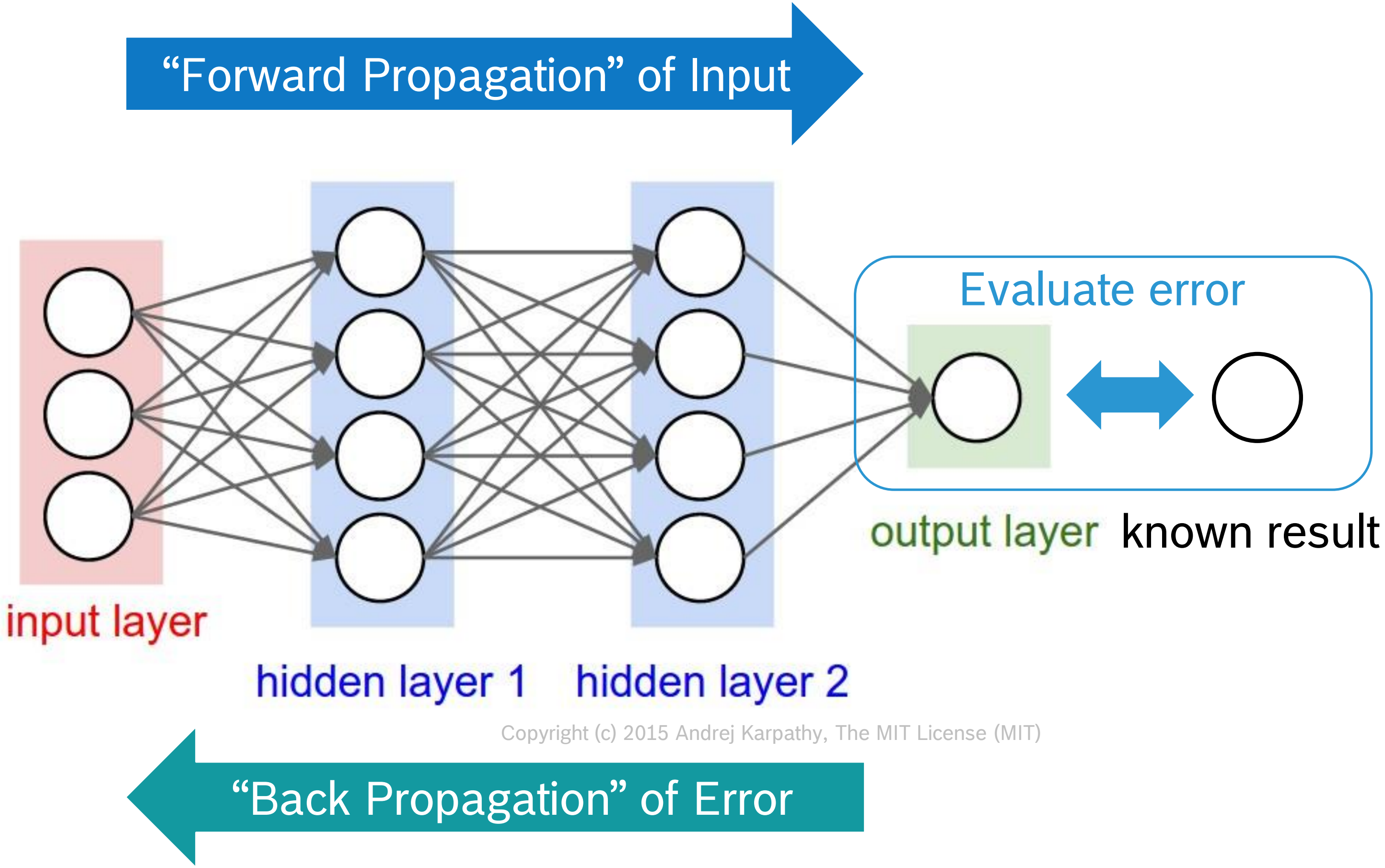
90° laundry

Non washable

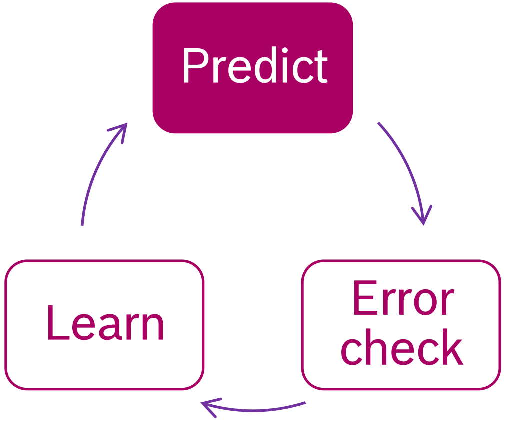
BOSCH

Deep Learning
Network Training

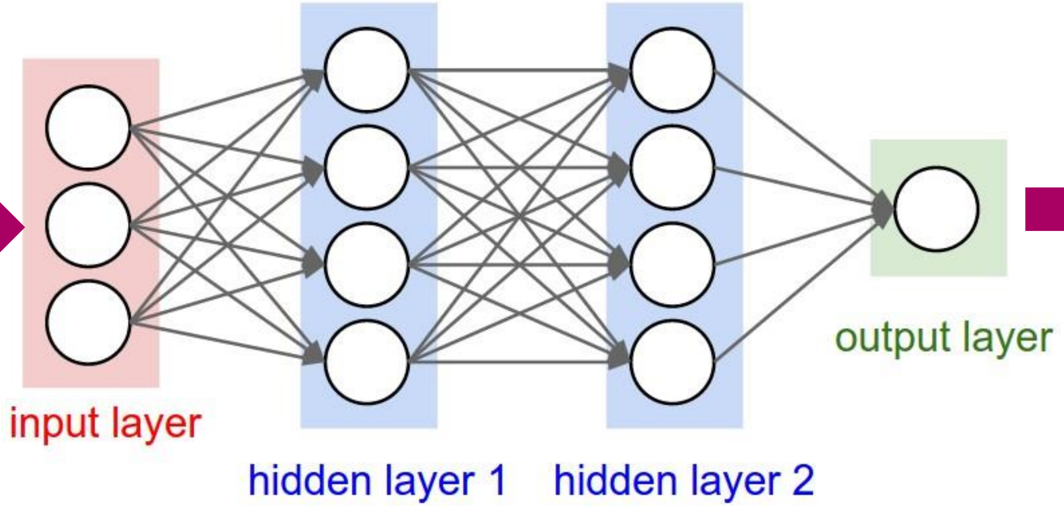




Deep Learning
Forward Propagation



Input

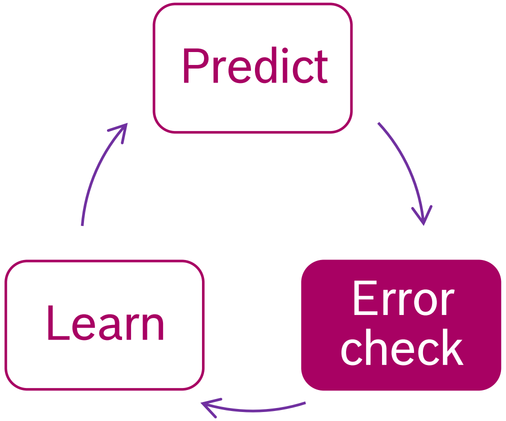


Forward Propagation of Input

Class	Prediction
30°C	0.6
60°C	0.05
90°C	0.01
N. w.	0.34

Output: Class “30°C”

Deep Learning
Error check



Prediction

Class	Prediction	Label
30°C	0.649	0
60°C	0.001	0
90°C	0.15	0
N. w.	0.2	1

Error = $\sqrt{(0.649^2 + 0.001^2 + 0.15^2 + 0.8^2)} = 1.041$

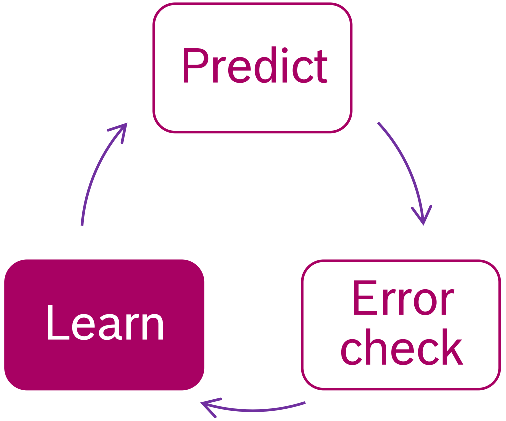
Prediction

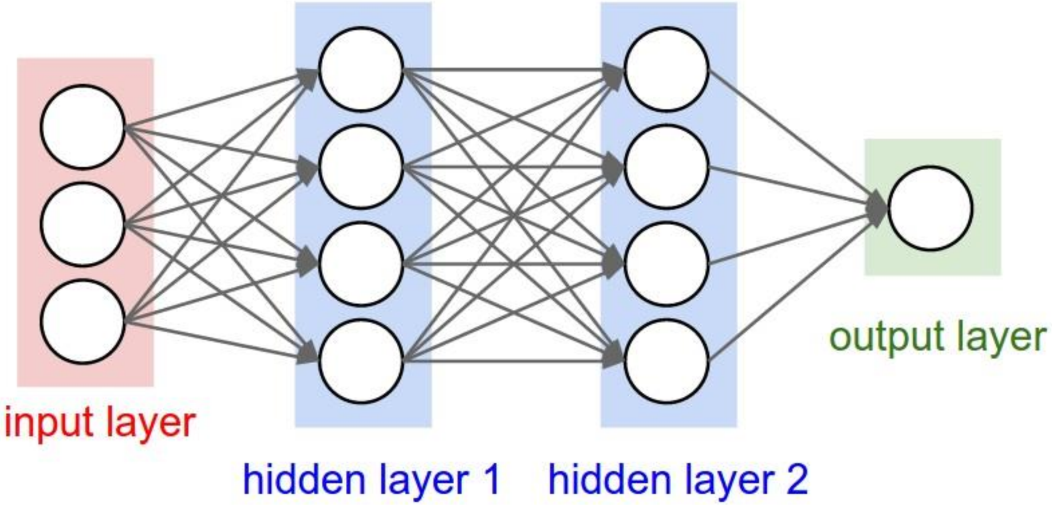
Class	Prediction	Label
30°C	0.6	1
60°C	0.05	0
90°C	0.01	0
N. w.	0.34	0

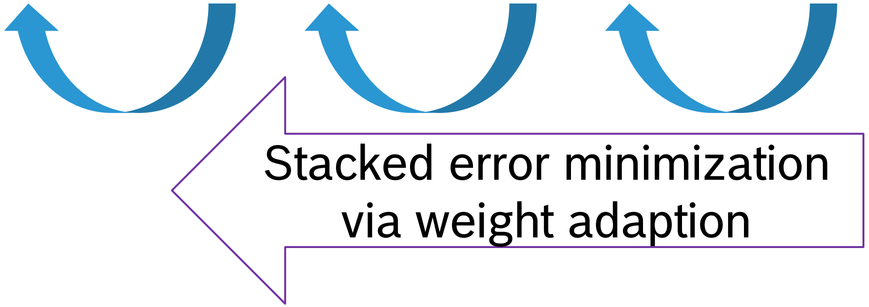
Error = $\sqrt{(0.4^2 + 0.05^2 + 0.01^2 + 0.34^2)} = 0.536$

BOSCH

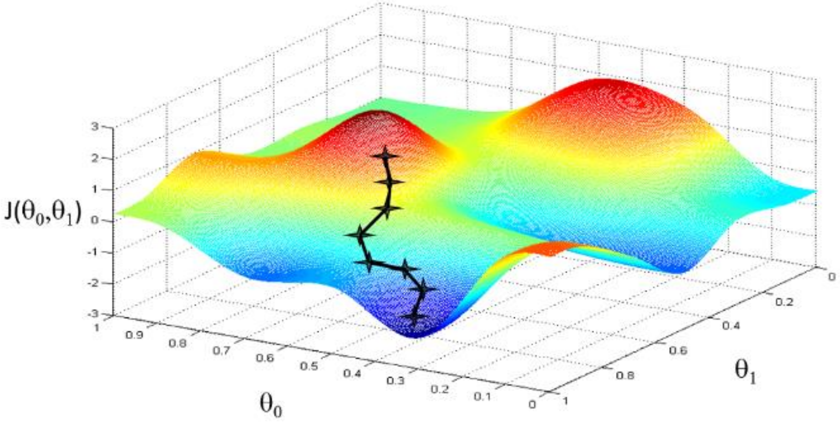
Deep Learning
Backpropagation







Error minimization using method of gradient descent



BOSCH

Deep Learning
Validation

Predict

Learn

Error check

Network training

Validation

Validation data set

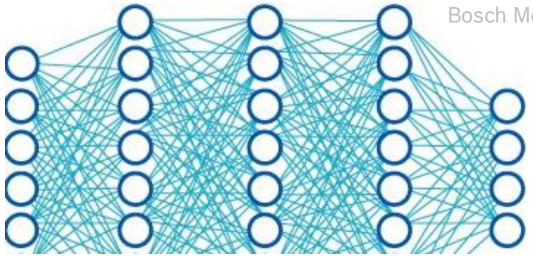
CC0

Check the error/accuracy on a different data set!

nok

Network adaption

Adaption of network type or other hyperparameters

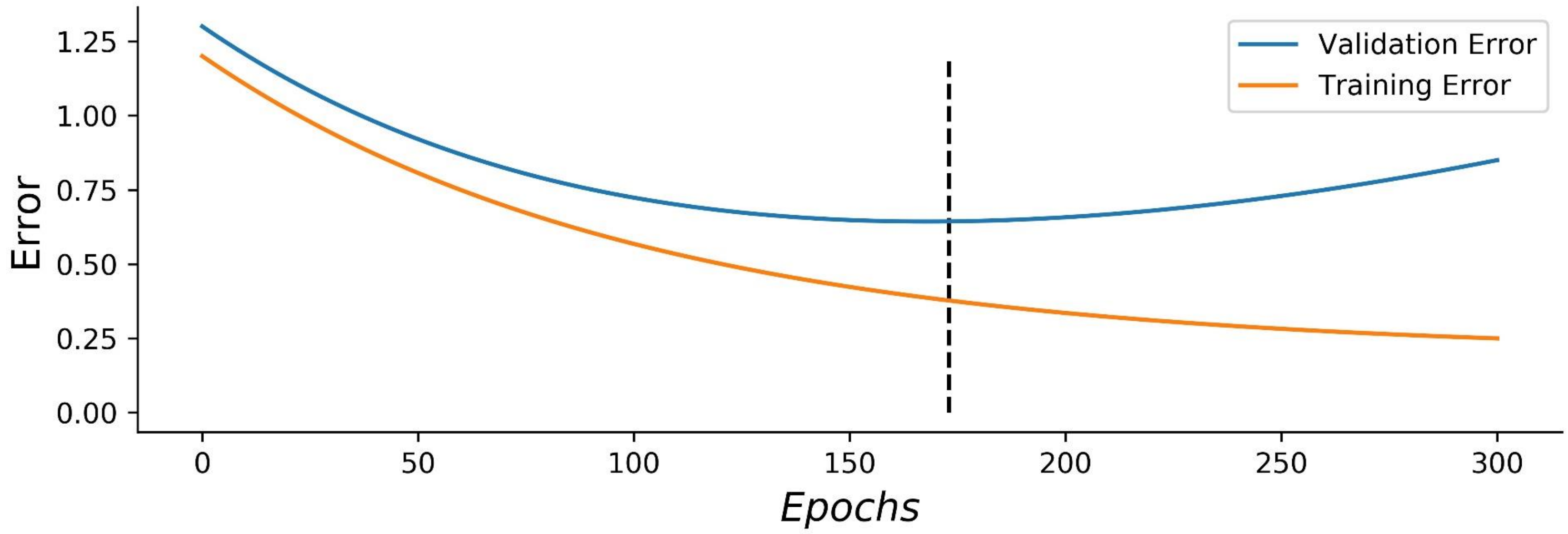


Bosch Media Space

BOSCH

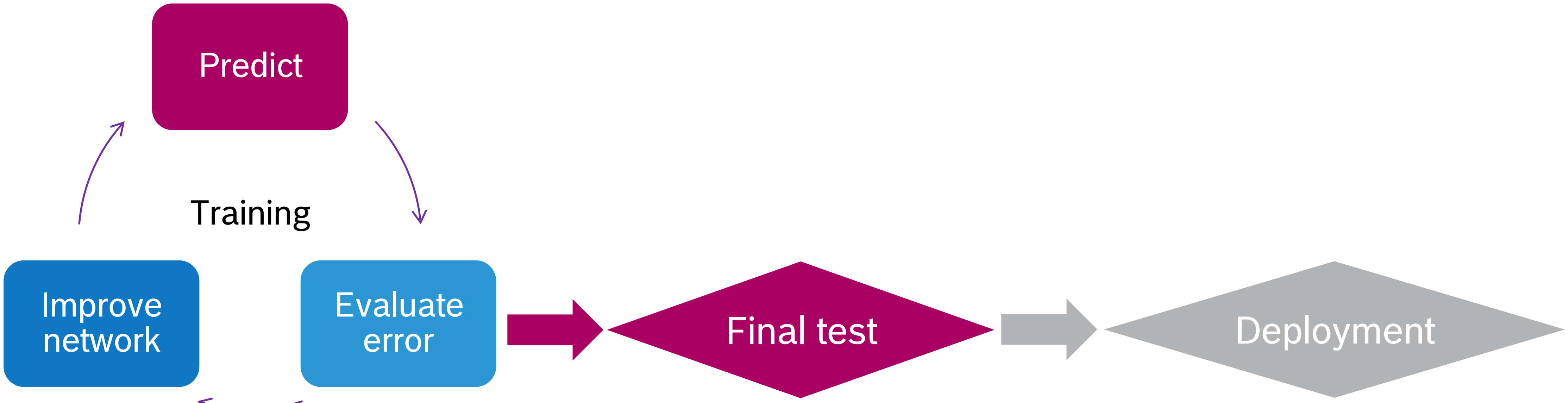
Deep Learning
Quantification of Improvement

"Learning curves"



BOSCH

Deep Learning
Test & Deployment

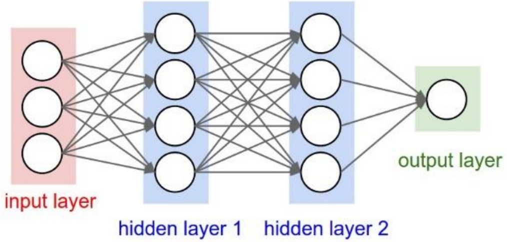


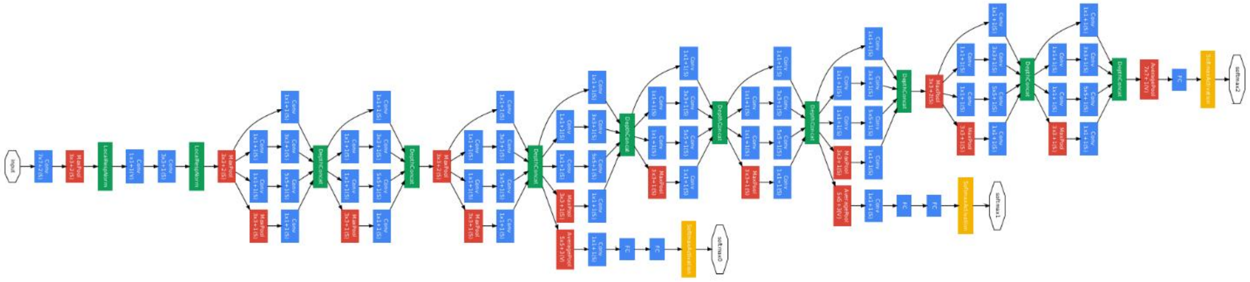
Check the error on a different data set!

Feed forward only

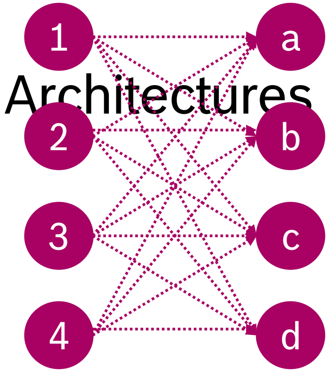
Deep Learning
Drivers of the Revolution

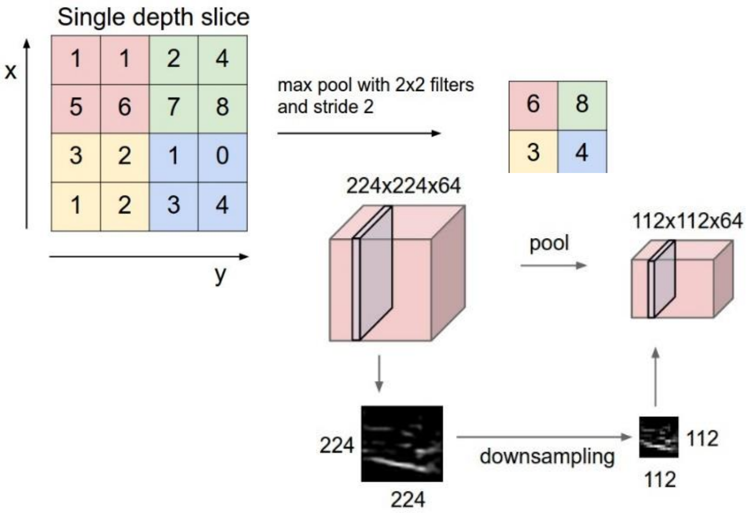
Deeper Networks

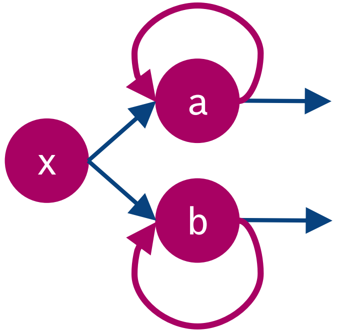




Novel Architectures



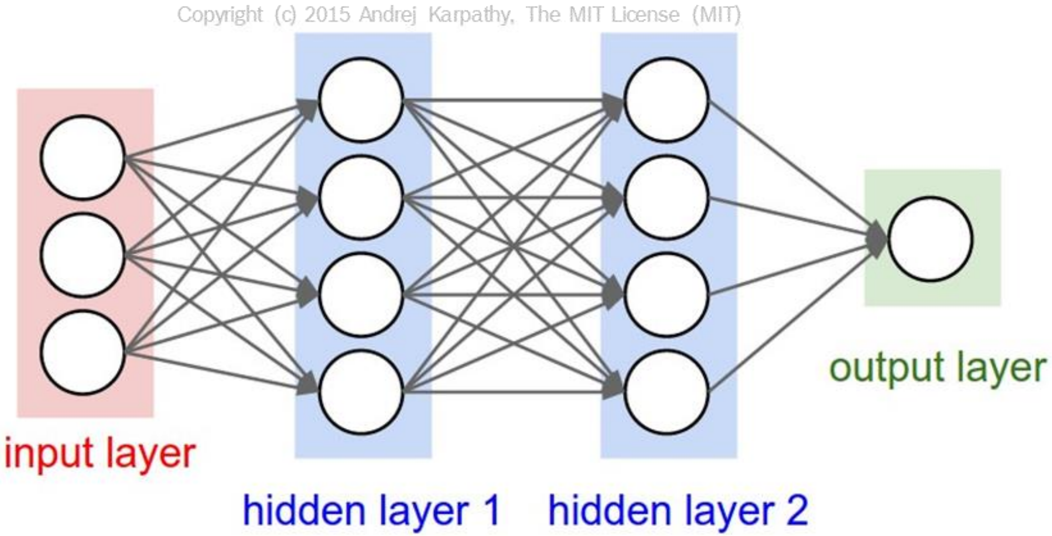




Deep Learning
Neural Network Architectures

You already know the

Multilayer Perceptron:



Every neuron in the previous layer is connected to every neuron in the next

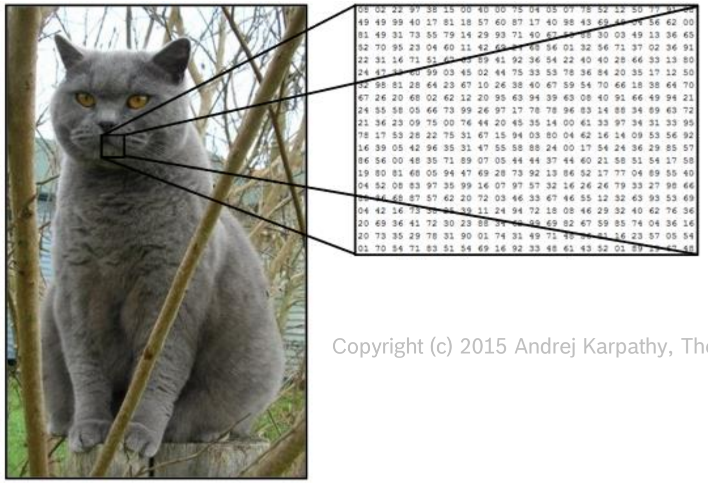
→ Doesn't scale with large inputs - too many weights!

Deep Learning
Convolutional Neural Network (CNN)

Challenge:

- Vision involves a large number of input neurons (e.g. 1000x1000 pixel values)

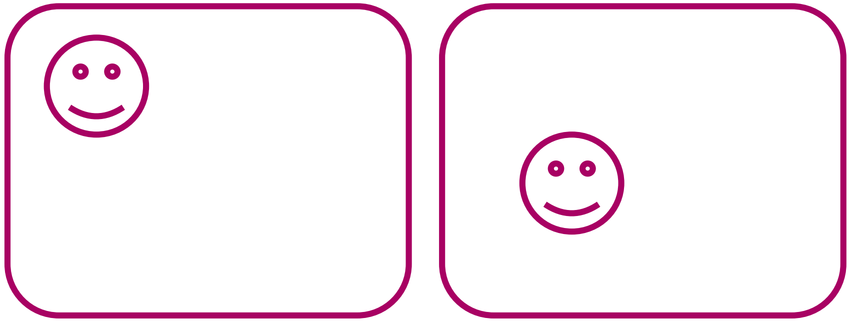
➡ Perceptron unfeasible



Copyright (c) 2015 Andrej Karpathy, The MIT License (MIT)

Solution:

- Take advantage of invariance
- Use “connectedness” of objects in nature

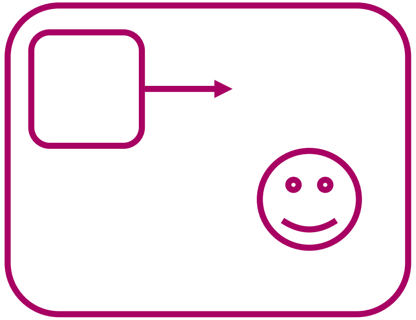


 Internal | CR/PJ-AI-E & CI/OST | 2019

BOSCH

Deep Learning
Filters

Filters can be seen as image cutouts to detect certain features:



- Filters applied over the whole image
- Weights have to be learned only once!

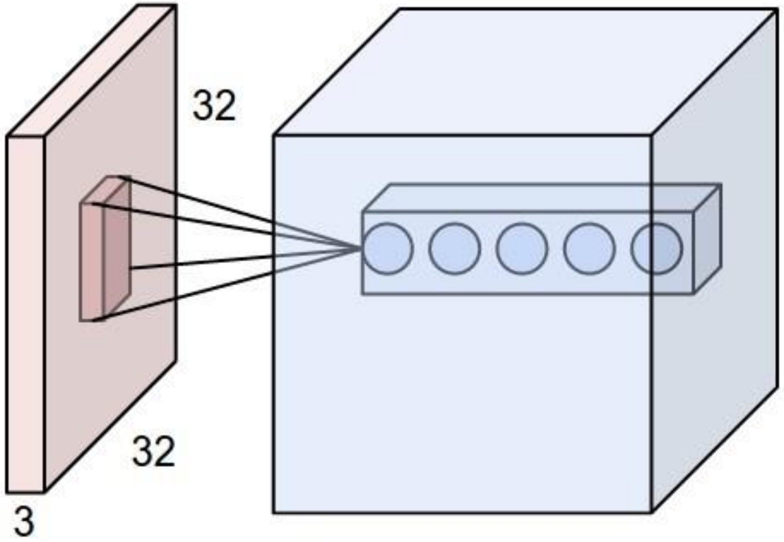
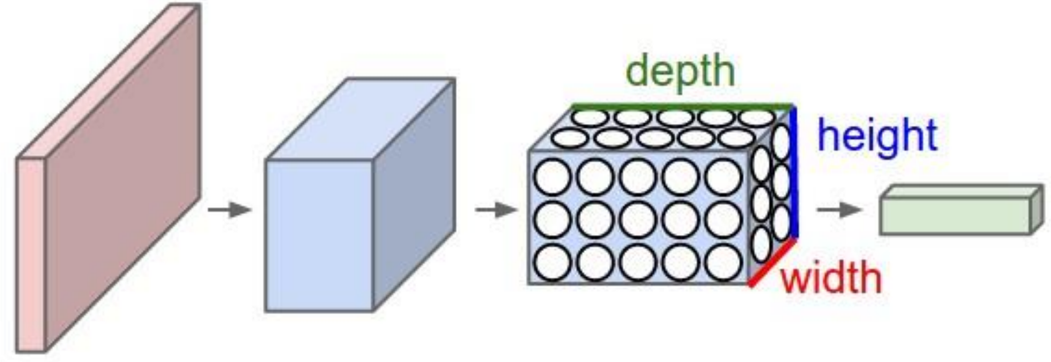


Image Source: http://cs231n.github.io/

Deep Learning
CNN Architecture



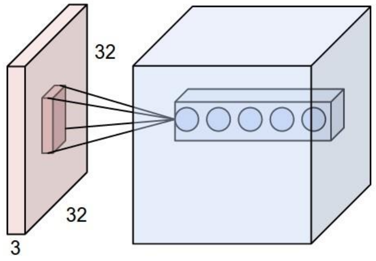
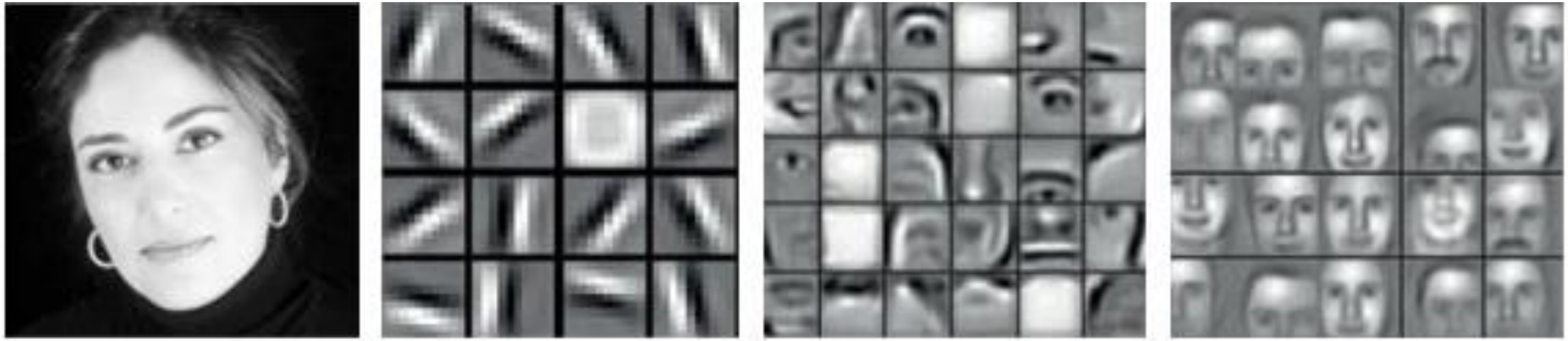


Image Source: http://cs231n.github.io/

- Height and width: Connection only to nearby neighbors
- Depth: Different filters for detection of different features!

Deep Learning
“Stacked” Feature Detection



edges | combinations of edges | object models

Deep Learning
Sweet Spot of Deep Learning

Use deep learning if you have ...

- High dimensional data like pictures
- Temporal/sequential data like text or speech
- Sufficiently large amounts of labeled data which are relevant for your target
- Sufficient computational resources

Depositphotos, Bosch License

Deep Learning
How much data do I actually need for training?

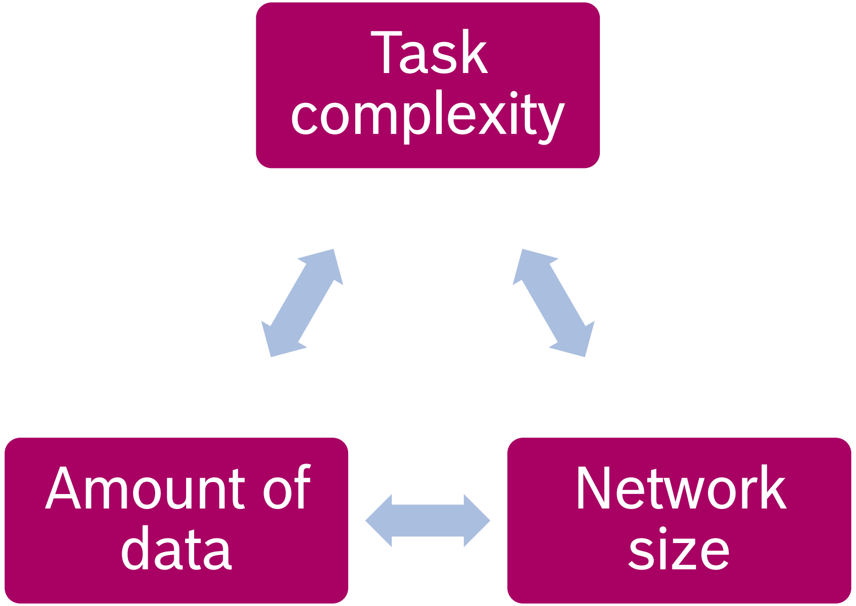
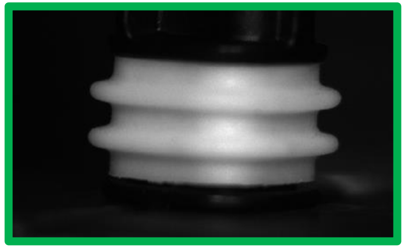
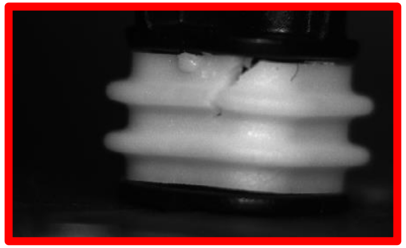


	Image Net Image Recognition	MNIST Digit recognition
# classes	1000 – 10000	10
# data samples	14 000 000	42 000
# neurons	50 000 000	< 100 000
Training time on 1 GPU	1-2 weeks	5 minutes

Deep Learning
Challenge “Blindness”

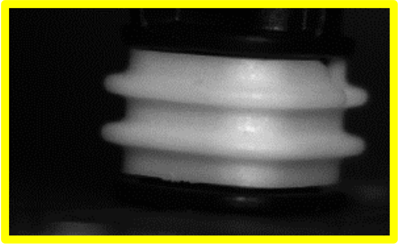


Model training on good parts



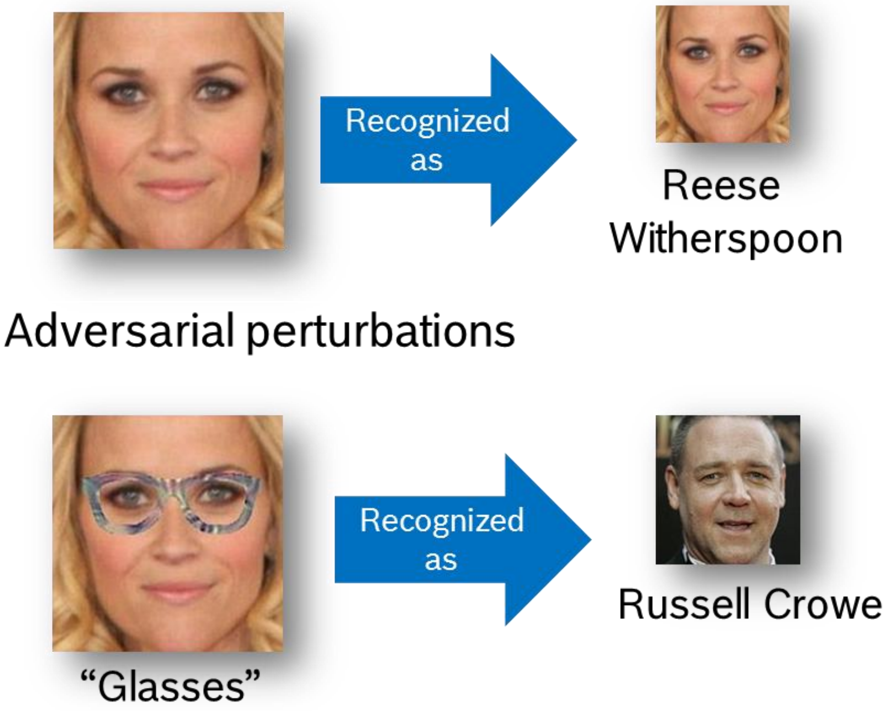
No training on certain defects

Unseen defects



- Can lead to severe consequence (defect parts go into the field)
- Must be excluded by a broad range of training data

Deep Learning
Challenge Robustness



Sharfi et al. Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition, CCS 2016
Eykholt et al.: Robust Physical-World Attacks on Deep Learning Visual Classification, CVPR 2018

BCAI activities:

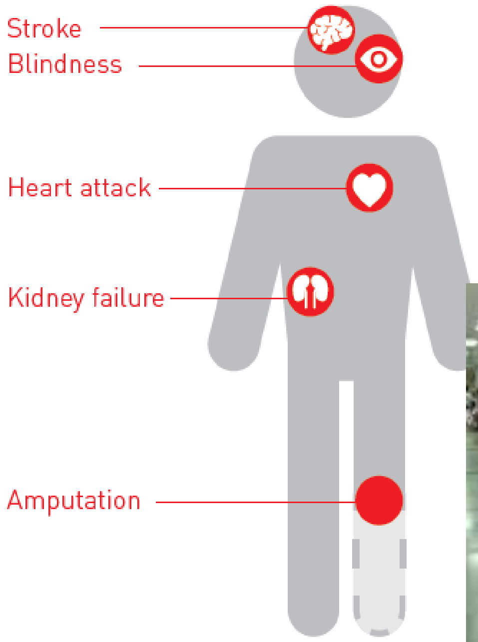
- Detection of adversarial attacks
- Development of provable defenses
- Implementation of robustness toolbox
- Invariance learning
- …

Deep Learning
Bosch Eye Care Solution: Challenge

422 Million suffer from diabetes worldwide

Consequences

Diabetes can lead to complications in many parts of the body and increase the risk of dying prematurely.



4.5 Million suffer from diabetes induced vision loss

BOSCH

Deep Learning
Bosch Eye Care Solution: Challenge

Early detection

of diabetic retinopathy can avoid vision loss completely!

4.5 Million suffer from diabetes induced vision loss

Missing infrastructure and costs

are the most prevalent obstacles for early detection in many countries in the world.

In India, there are rates of

1 eye doctor / 83.000 people

1 doctor / 2000 people

Deep Learning
Bosch Eye Care Solution: Overview

Revolutionary handheld fundus camera brought to you by Bosch

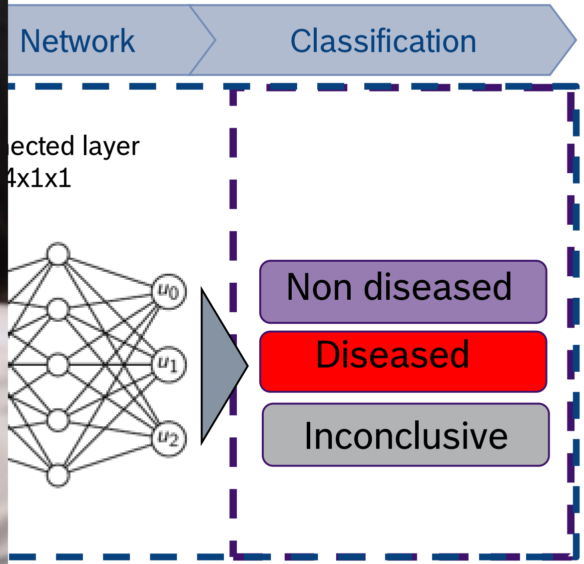
www.boscheyecare.com

A world-class ecosystem of hardware, software, and algorithm for the detection of Diabetic Retinopathy

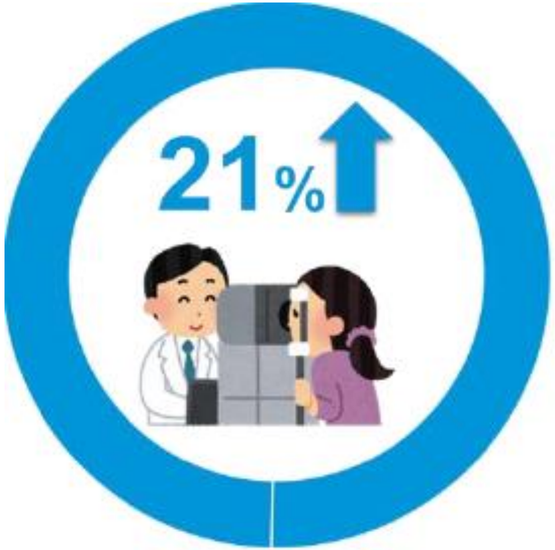
- Portable
- Affordable (~3300€)
- Easy, quick examination
- Understandable results

BOSCH

Deep Learning
Bosch Eye Care Solution: Method



Deep Learning
Bosch Eye Care Solution: Results for representative Eye Centers



21% more check-ups
due to quicker and painless examination

17% increase in detection
of diabetic retinography

8% increase in detection
of glaucoma

Deep Learning
Bosch Eye Care Solution: Application Possibilities

Public Health clinics

Eye camps

Mobile Health clinics

Home Health

Out patient dept

Multi speciality Hospital

BOSCH

IMPLEMENT– ROLES, TRAINING AND CONSULTING

Implement
Overview

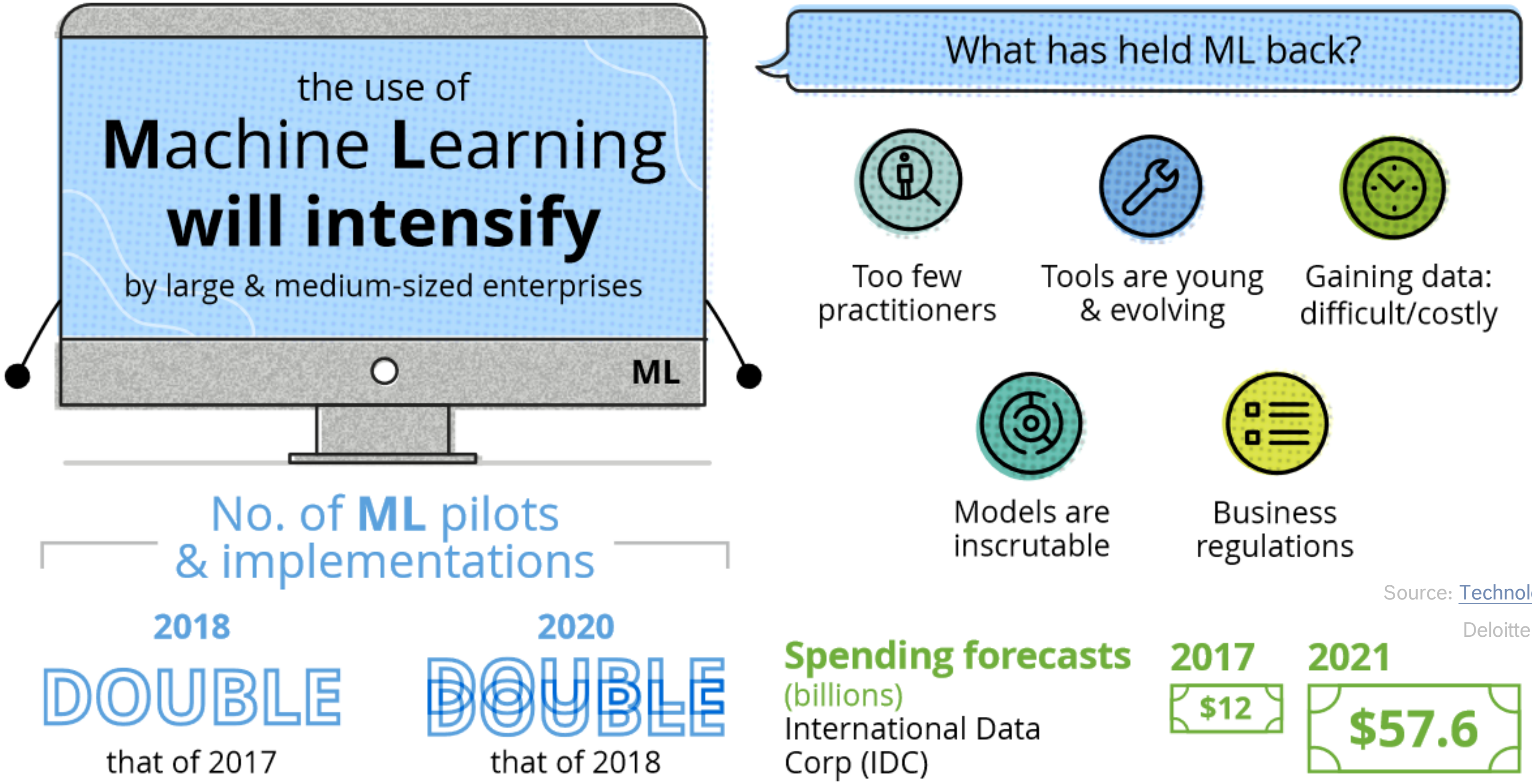
Data-driven mindset

Roles, Training and Consulting

Platform providers

BOSCH

Roles, Training and Consulting
Challenges in Machine Learning



Source: [Technology, Media and Telecommunications Predictions 2018](), Deloitte 2017

BOSCH

Roles, Training and Consulting
Roles and Competencies

Data Analyst

Has a data-driven mind-set, uses analytical tools applied to his/hers **domain.**

Data Scientist

Expert in statistics, machine learning, and big data technologies and methodologies

Data Engineer

Builds big data systems, efficient applications, and develops **deployment** strategies and IT architecture

Roles, Training and Consulting
Data Scientist – nice-to-have competencies

Data Scientist

- **Background** in Mathematics, Physics, Computer Science, Data Science, Statistics
- Has worked in **academic setting**, i.e. university research group, Max Planck Institute, Fraunhofer Institute
- Participated in **kaggle** competitions, has code on **github**
- Brushed up Machine Learning skills in **online university** such as udacity, iversity, coursera and iTunesU
- **Programming** skills in Python, tensorflow, pytorch, R, keras

BOSCH

Roles, Training and Consulting
Data Engineer – nice-to-have competencies

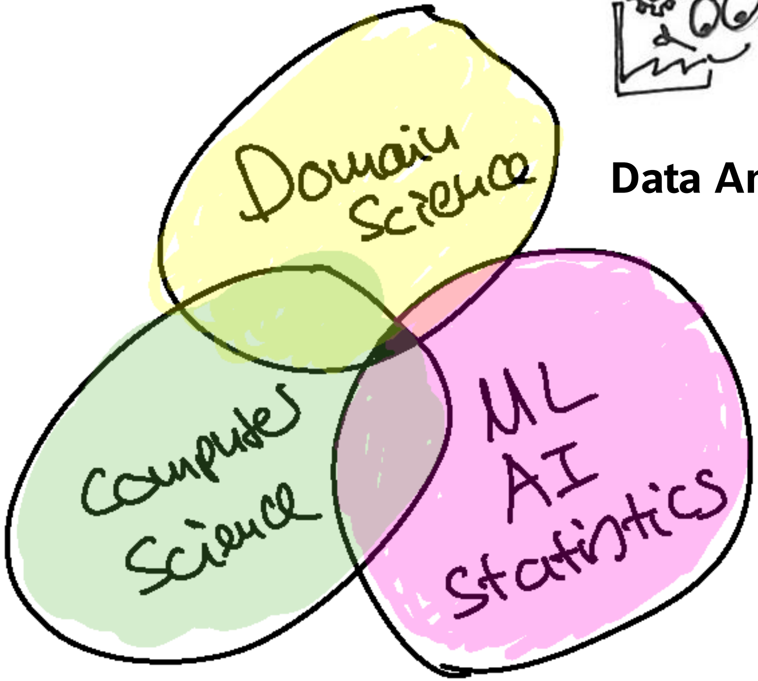
Data Engineer

- **Background** in Computer Science, Physics, Engineering, Data Science
- Has experience in **high-performance computing**
- **Knowledge** about Unix, system administration and computer networks
- Experience in **parallel computing** and distributed systems
- Past work in **academic setting**, i.e. university research group, Max Planck Institute, Fraunhofer Institute as an additional asset
- **Programming** skills in Python, Java, bash, R and preferably Spark as a framework

BOSCH

Roles, Training and Consulting
Data Analyst, Data Scientist and Data Engineer

Data Analyst



Data Engineer

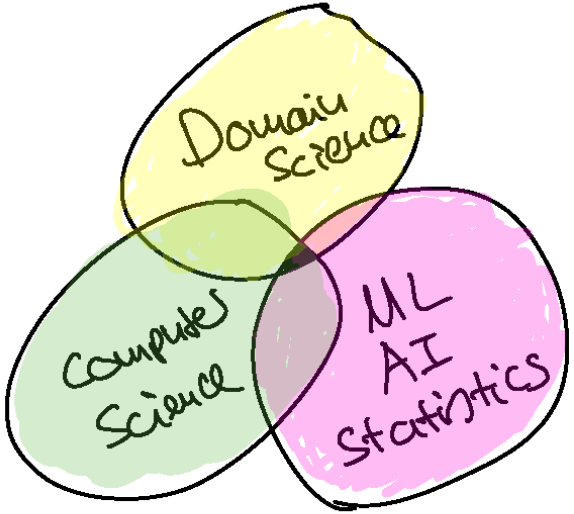
Data Scientist

Data-driven projects are a team effort where **domain experience**, **analytical skills** and **technological expertise** go hand in hand.

BOSCH

Roles, Training and Consulting
Interplay at Controlling Use Case

Michael Binder (G2/PJ-FCE-IT)
Background: Controlling
Project topics: Product lead



Phil Gaudreau (CR/PJ-AI-S1)
Background: Mathematics
Project topics: Design of prediction algorithm

Dominik Mayer
Background: Hadoop and Linux administration

Cinar Goktug (CR/PJ-AI-S1)
Background: PhD in Machine Learning
Project topics: Design and implementation of prediction algorithm

BOSCH

Roles, Training and Consulting

Bosch Center for Artificial Intelligence: Key facts

Expanding

Locations

Germany (Renningen), U.S.A. (Sunnyvale and Pittsburgh), India (Bangalore), Israel (Haifa)

>170 AI Experts

already filed 81 patents and 24 Top Tier publications since foundation of BCAI in 2017

Our Target

AI-driven efficiency improvements account for **1% reduction** in Bosch Group operating costs in year 5 after formation

AI-driven Bosch products account for **10%** of Bosch Group **profit** in year 5 after formation

Strong

Academic Partnerships

e.g. Delta Lab in Amsterdam, CMU Pittsburgh, Cyber Valley initiative ….

Business Impact

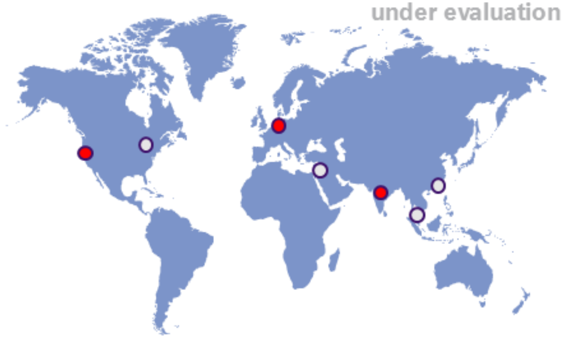
created across all functions and GBs with >125 projects e.g. in Manufacturing, Engineering, SCM & Controlling, Intelligent Services

Central contact: bosch.center.for.ai@bosch.com

BOSCH

Roles, Training and Consulting

Bosch Center for Artificial Intelligence: Divisions & Offers



BCAI
CR/PJ-AI

AI Consulting
support project strategies

AI Marketing
raise awareness of BCAI

AI Enabling
accelerate digital business

AI Research
gain best in class AI technology

AI Services
drive commercialization

- Use-case definition
- AI infrastructure strategy
- Business models

- Competency ramp up
- Trainings, workshops
- Data strategy support
- Community build up
- AI exchange platform

- AI platform
- AI turnkey projects
- Model maintenance

BOSCH

Roles, Training and Consulting
BCAI Enabling Value Proposition

- Competency ramp-up
- Community build-up
- Leading edge work environment
- Data strategy support

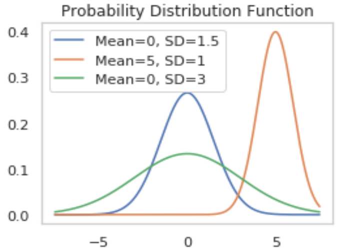
Roles, Training and Consulting
BCAI Enabling Training Offers

Management & Project Leads

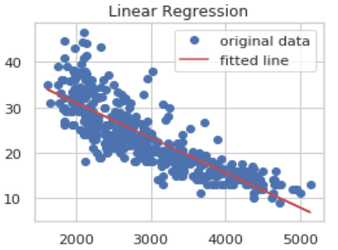
AI Awareness

Deep Learning Deep Dive

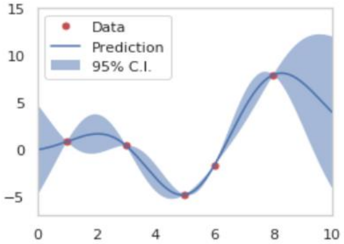
Technical Trainings



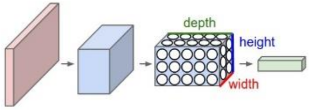
Statistics



Data Science



Machine Learning



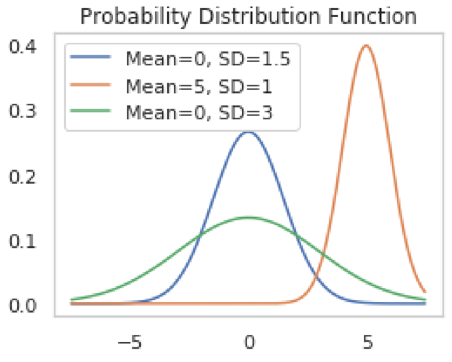
Deep Learning

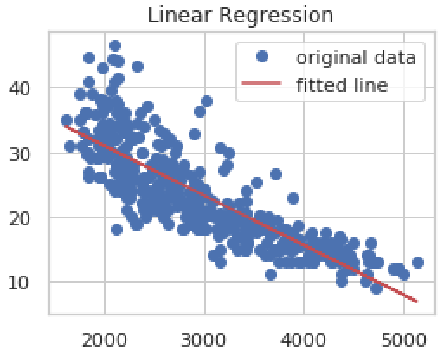
BOSCH

Roles, Training and Consulting
Basics of Statistics

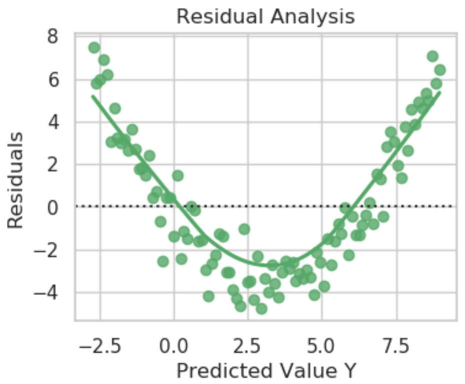
Training duration: 1 day

Max. 15 participants





Decision	The Null Hypothesis: True	The Null Hypothesis: False
Accept H_0	(1-α)	β
Reject H_0	α	(1-β)



LO1: Understand Basic Statistical Terminology

- (Conditional) probabilities
- Location and distribution parameters
- Common distributions

LO2: Correlation and Linear Regression

- Correlation
- Linear regression
- Parameter interpretation

LO3: Parameter Estimation and Statistical Tests

- Parameter estimation
- Hypothesis tests
- Confidence intervals

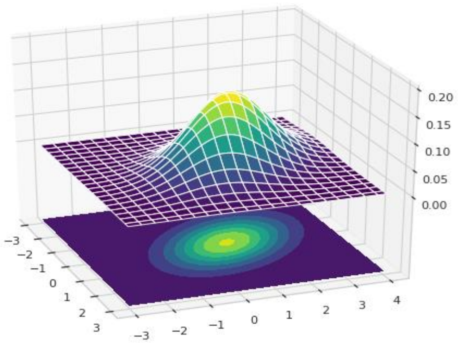
LO4: Linear Regression Assumptions and Diagnostics

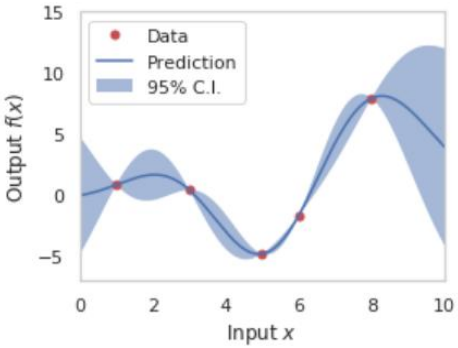
- Basic assumptions, tests and remedies against assumption violation
- Residual analysis

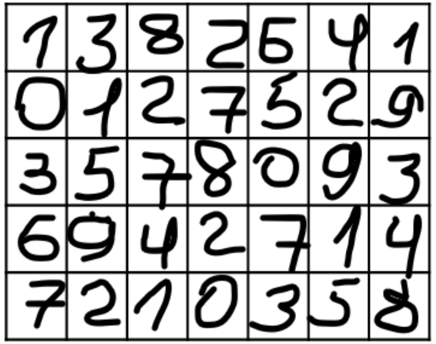
Roles, Training and Consulting
Advanced Statistics & Gaussian Processes

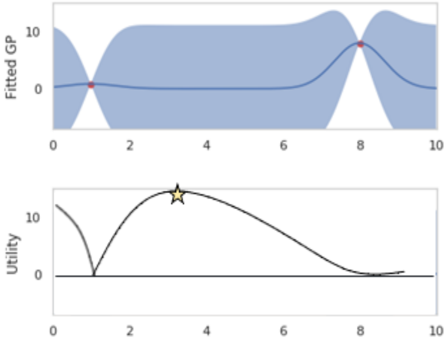
Training duration: 1 day

Max. 15 participants









LO1: Gaussian Mixture Models

- Multivariate Gaussian distribution
- Gaussian mixture models
- Estimation by expectation maximization

LO2: Understanding Gaussian Processes (GP)

- Covariance function
- Noise-free vs. noise case
- Hyperparameters (signal and noise variance, length-scale)

LO3: Advanced Topics and Applications of GP

- Types of covariance functions
- Regression and classification with GP
- Model estimation and selection

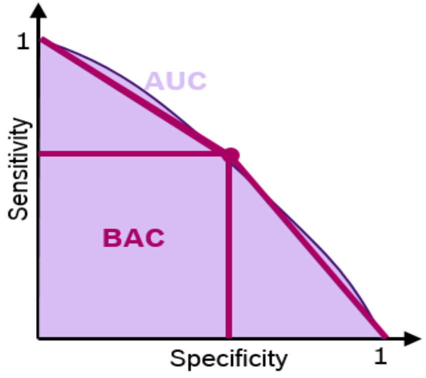
LO4: Bayesian Optimization (for Hyperparameters)

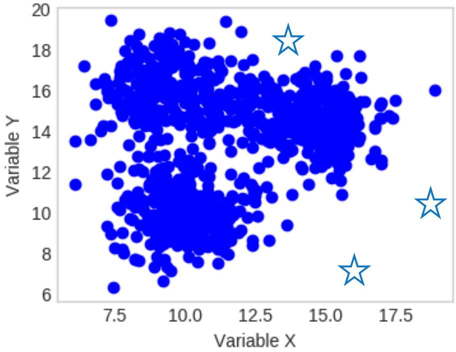
- General purpose
- Acquisition functions
- Exploration vs. Exploitation
- Application for GP

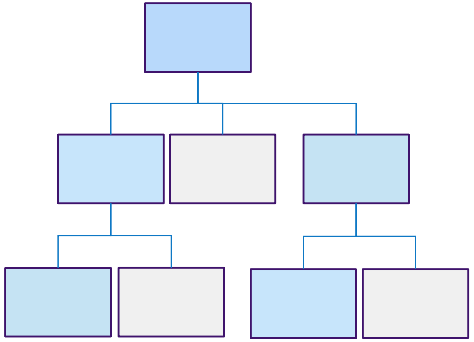
BOSCH

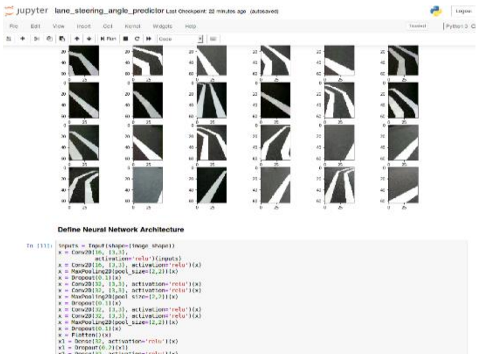
Roles, Training and Consulting
Basics of Machine Learning

Training duration: 2 days
Max. 15 participants









LO1: Machine Learning (ML) Concepts

- ML workflow
- Model complexity
- Regularization
- Ensembles

LO2: Data Science Workflow

- Data exploration and dimensionality reduction
- Outlier detection
- Model validation and selction

LO3: Algorithms and Methods

- Clustering
- Decision trees
- Support vector machines
- Neural networks

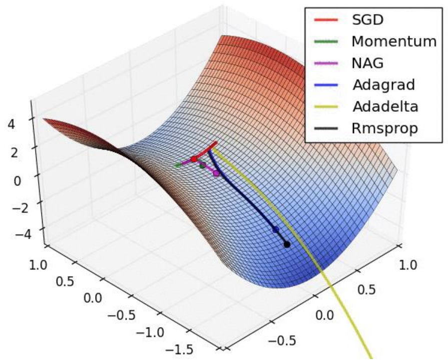
LO4: Application of Learned Concepts using Python

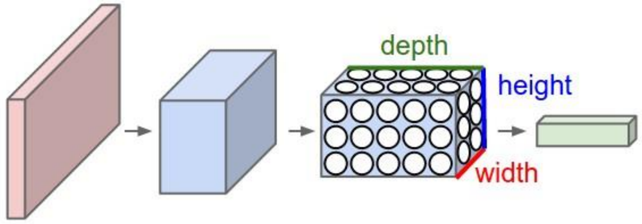
- Implementing the models
- Performance comparison
- Use cases at Bosch and potential in BUs

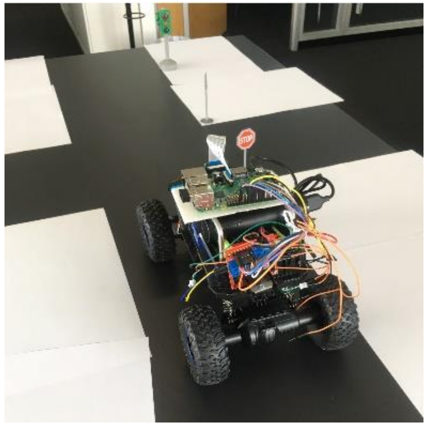
BOSCH

Roles, Training and Consulting
Deep Learning Training

Training duration: 1 day
Max. 15 participants







LO1: Key Concepts of Deep Learning

- Training and inference
- Optimization
- Activation and loss functions
- Weight initialization
- Transfer learning

LO2: Architectures and their Applications

- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Generative Adversarial Networks (GAN)

LO3: Work on a real Deep Learning Problem

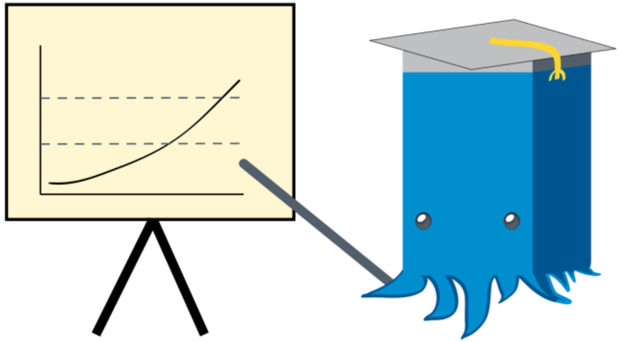
- Develop neural networks in the context of autonomous driving
- Deploy your results on a model car's edge device

Images: http://cs231n.github.io (MIT), A. Radford https://imgur.com/a/Hqolp

BOSCH

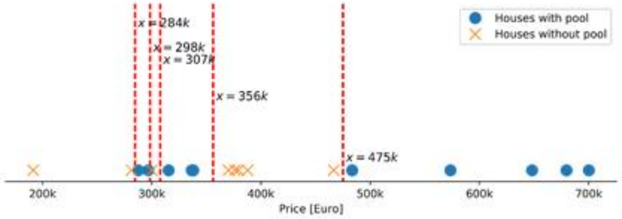
▶ Roles, Training and Consulting
BCAI Enabling Learning Platform

Target Group



Motivated AI Enthusiasts

Training Offers



- Self-paced learning on Jupyter notebooks
- Prerequisite: basic Python skills
- Linked with our community

Access Links

[Visit Our Learning Platform]

[Visit Our Community]

BOSCH

Roles, Training and Consulting
BCAI Enabling "Brokering Map"

Target Groups

Management
Project Leads

Technical Experts
Basic/Advanced

Training Offers

- Content and prerequisites
- Locations and dates
- Access and costs
- …

Self-Service Education Area

- Books and scientific articles
- University courses
- Online tutorials and lectures
- Communities (internal/external)
- Blogs
- Podcasts and talks
- Videos and demonstrators

Access Links

Visit Our Brokering Map

How to Contribute

BOSCH

Roles, Training and Consulting
Self-study material

Talks

- ❖ 2016 “KI – Vision und Wirklichkeit” by Prof. Jürgen Schmidhuber (audio + slides)
- ❖ 2018 “Ethik des "autonomen" Fahrens“ by Prof. Peter Dabrock (audio + slides)
- ❖ 2018 „Artificial Intelligence and its Impacts” by Prof. Tom M. Mitchell (video + slides)
- ❖ 2017 „AI is taking over the world” by Dr. Lothar Baum (video)

More resources here

Free Online Courses

- ❖ Udacity: Introduction to Hadoop und MapReduce by Cloudera (Website)
- ❖ CalTech course CS156 on Machine Learning by Yaser Abu-Mostafa
- ❖ Coursera: Machine Learning by Andrew Ng (Website)
- ❖ Stanford CS class CS231n: Convolutional Neural Networks for Visual Recognition

Roles, Training and Consulting
Self-study material

Podcasts

- ❖ Great for traveling or exercising time!
- ❖ Machine learning guide (iTunes / Website): introduction to ML, fun & insightful, useful resource lists
- ❖ Learning machines 101 (iTunes / Website): introduction to AI and ML, a bit more advanced
- ❖ CRE208 Neuronale Netze (iTunes / Website): german podcast about ANN & applications

Libraries & Frameworks

- Scitkit-learn
- H2O
- TensorFlow
- Microsoft CMTK
- TensorFlow
- Keras
- Caffe
- PyTorch

More resources here

AI podcasts here

Roles, Training and Consulting
Communities – stay on top of the latest trends

- Bosch Center for Artificial Intelligence
- CoC Big Data News Blog - CoC Big Data
- Machine Learning Curriculum 2016
- Deep Learning
- Reinforcement Learning
- Social Coding @ Bosch
- Python

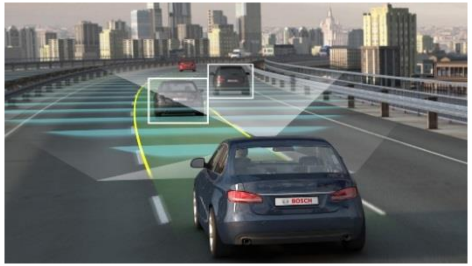
More links here

BOSCH

IMPLEMENT– PLATFORM PROVIDERS

Platform providers
CI/OST – CoC Big Data platforms

Mobility Analytics



Autonomous Driving Platform for CC

Connected Industry

RB i4.0 Analytics Platform for AE, PS, CC, DC and CM

Internet of Things

Bosch IoT Ecosystem

Big Data Platform as part of the Bosch IoT Cloud

Deep Learning

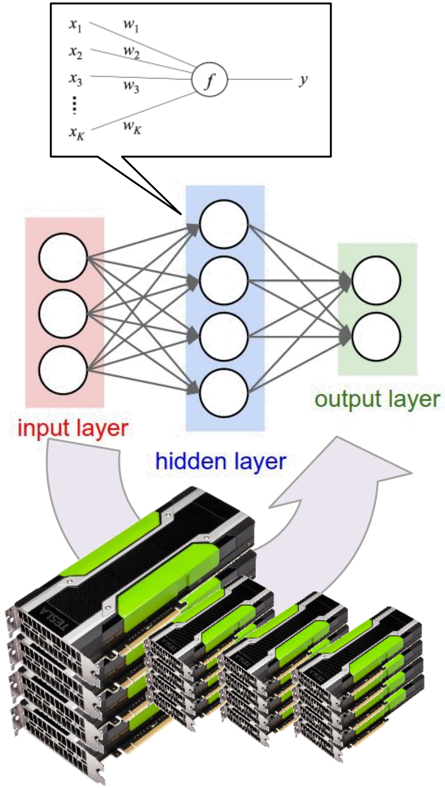
GPU Computing Platform for BCAI, CR and CC

- CoC Big Data Director: Dr. Michael Peters (CI/OST)
- CoC Big Data Senior Manager: Nicolas Rueger (CI/OST1)
- Bosch Connect: [https://connect.bosch.com/communities/community/**bigdata**](https://connect.bosch.com/communities/community/bigdata)

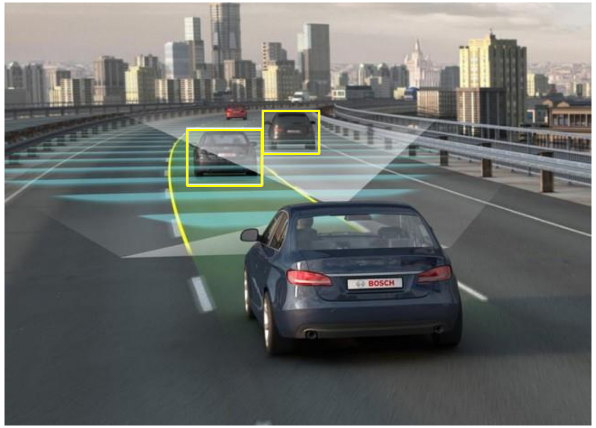
BOSCH

Platform providers
GPU Cluster: Infrastructure for Deep Learning Applications

Data sets
for training and validation



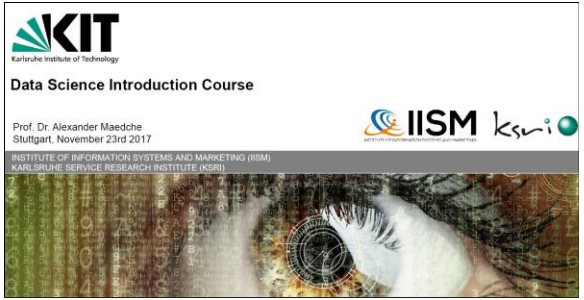
GPU Cluster
to accelerate model training



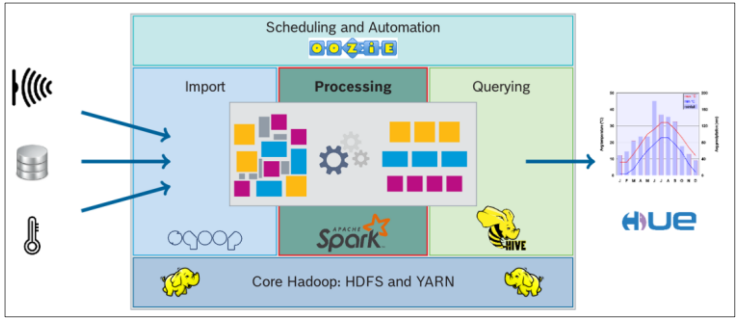
Inference
Application of models
to unknown (e.g. live) data

BOSCH

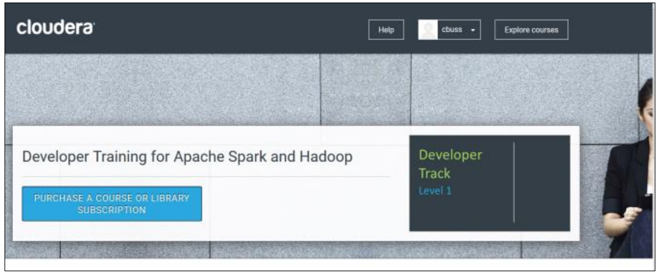
Roles, Training and Consulting
Big Data Analytics training portfolio



[IT-BD-120] Intro to Data Science



[IT-BD-110] Hadoop Hands-On Tutorial



[IT-BD-210-A] Hadoop and Spark Developer

IMPLEMENT– DATA-DRIVEN MINDSET

Data-Driven Mindset
Example: Data-driven Mindset at Amazon

- Everything that can be **measured**, is measured: web design, product features, HR, finance, operations processes …
  - Example: impact of website load times on sales

- Prove the opportunities of a new idea with a **live test** and **result data**

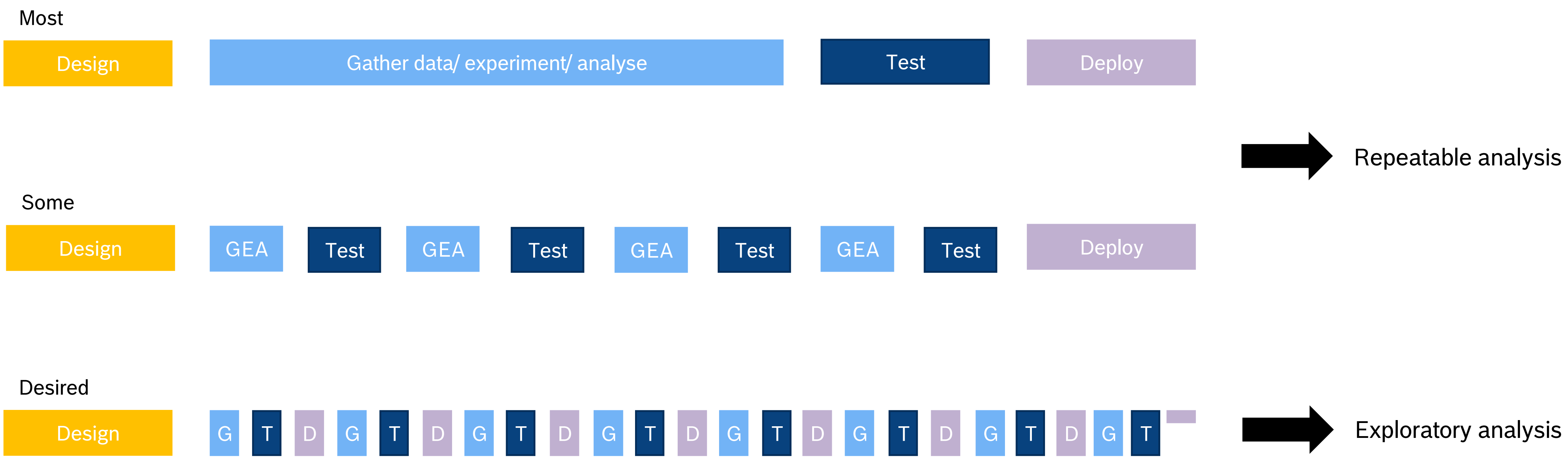
- Everyone, regardless of seniority, has **access to data** and tools to test their ideas and intuitions

Direct access to data and computing capabilites are of fundamental importance.

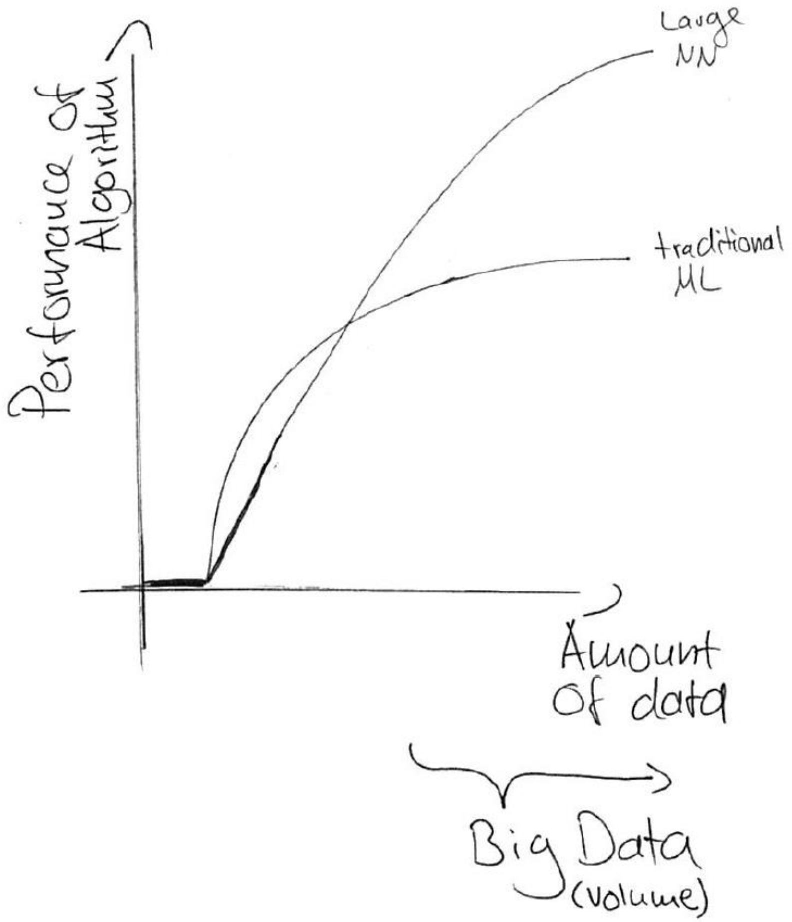
Source: https://www.entrepreneur.com/article/237326

BOSCH

Data-Driven Mindset
Data worker process

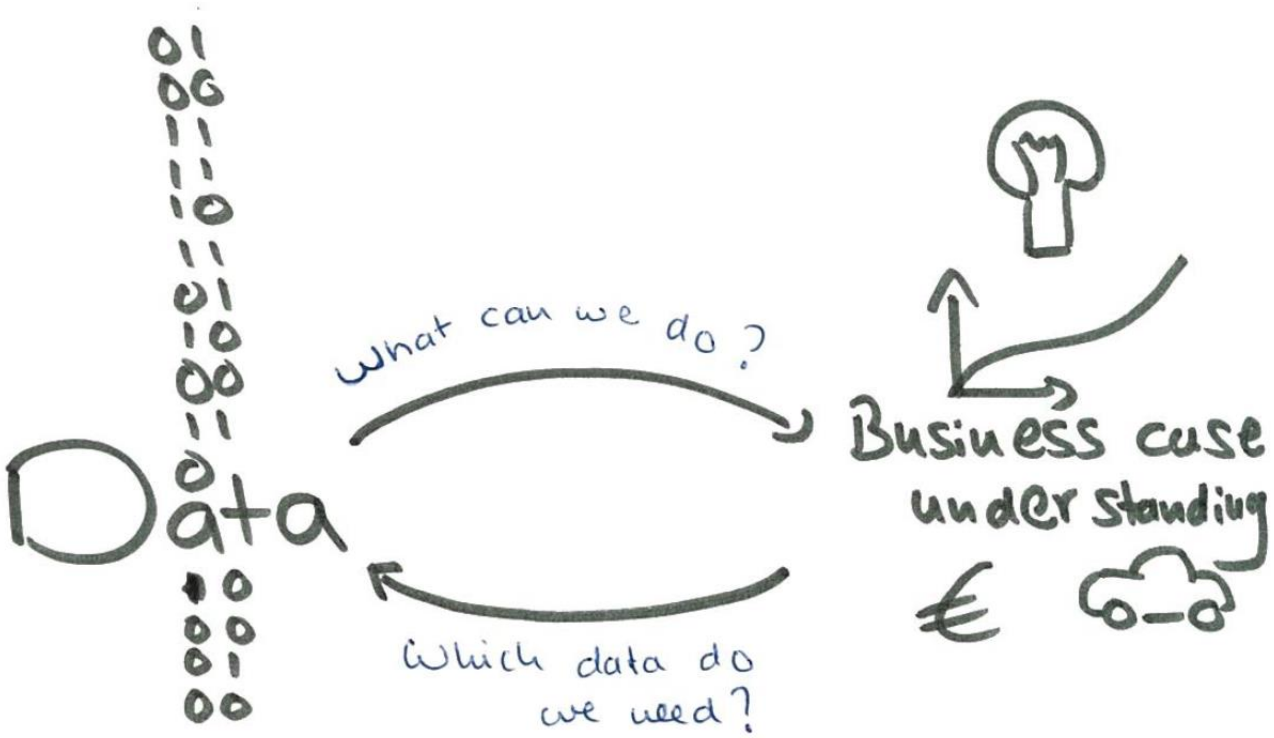


Data-Driven Mindset
The value of data



Not only algorithms **win the game** but also data.

Data-Driven Mindset
What is first: data or the business case?

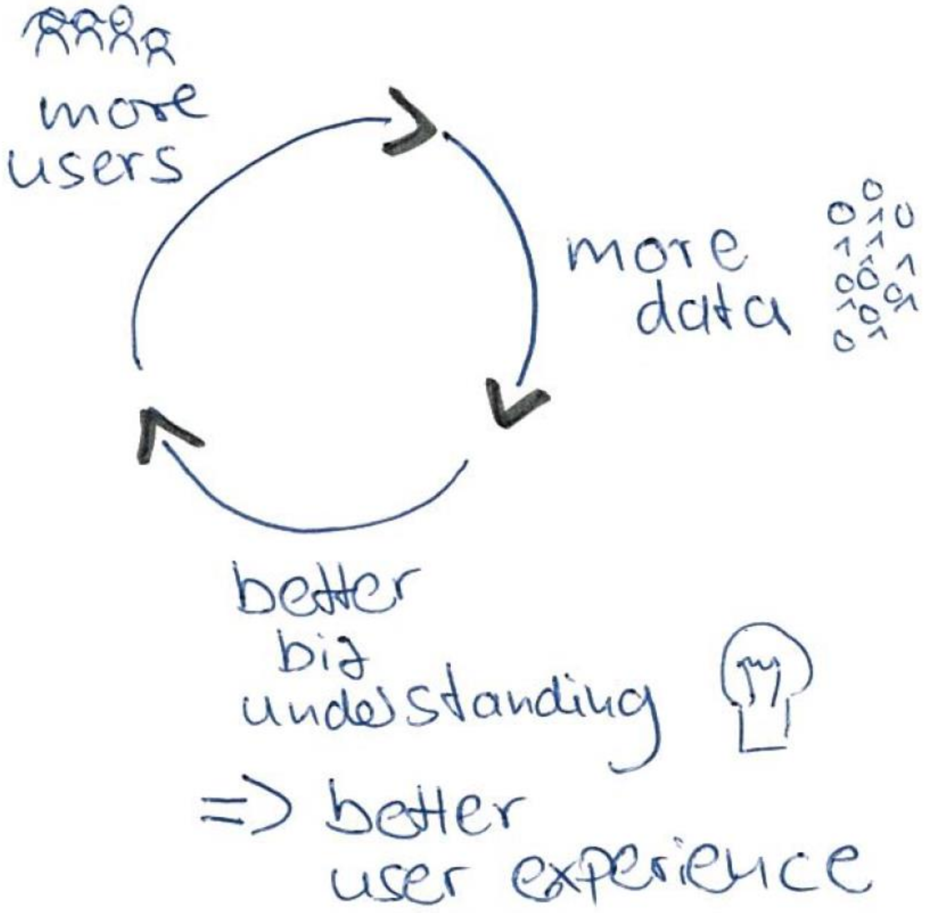


Data first: collecting data **early** is the start of **data strategy**.

Surely data can be collected before a fully defined business case.

Data-Driven Mindset

The positive data feedback loop / data advantage



Data is the defensible barrier in data-driven business.

It may lead to a continuous self-enhancing improvement cycle.

BOSCH

Retrospective
Disclaimer for pdf handout

License statement of document: not allowed to publish external to Bosch.

Usage of training material in products only after

- ❖ Written agreement,
- ❖ Disclaimer of liability,
- ❖ Reference on use of third party sources (Pictures, Logos)