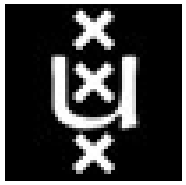


Applied Machine Learning
Introduction

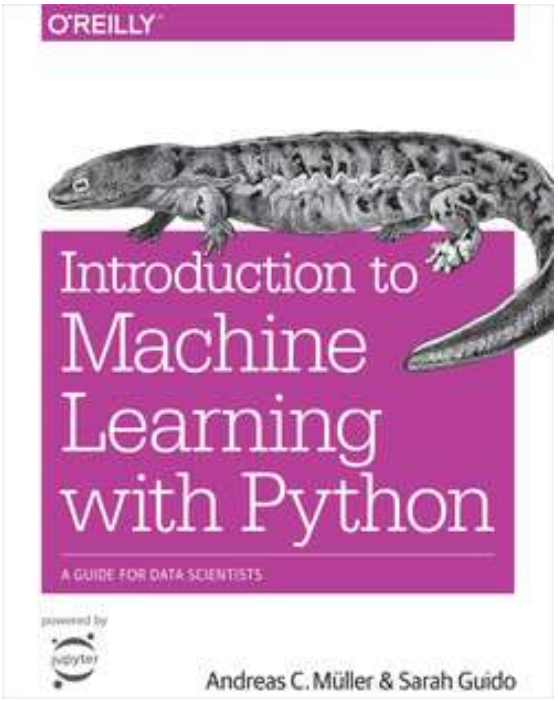
BSc course Informatiekunde 2026
https://staff.fnwi.uva.nl/a.visser/education/AML

Arnoud Visser
Intelligent Robotics Lab & Computer Vision Lab
Informatics Institute
Universiteit van Amsterdam
A.Visser@uva.nl

Illustrations courtesy of Maarten Marx, Sarah Guido, Benjamin Bengfort, and many others.

Introduction to Machine Learning

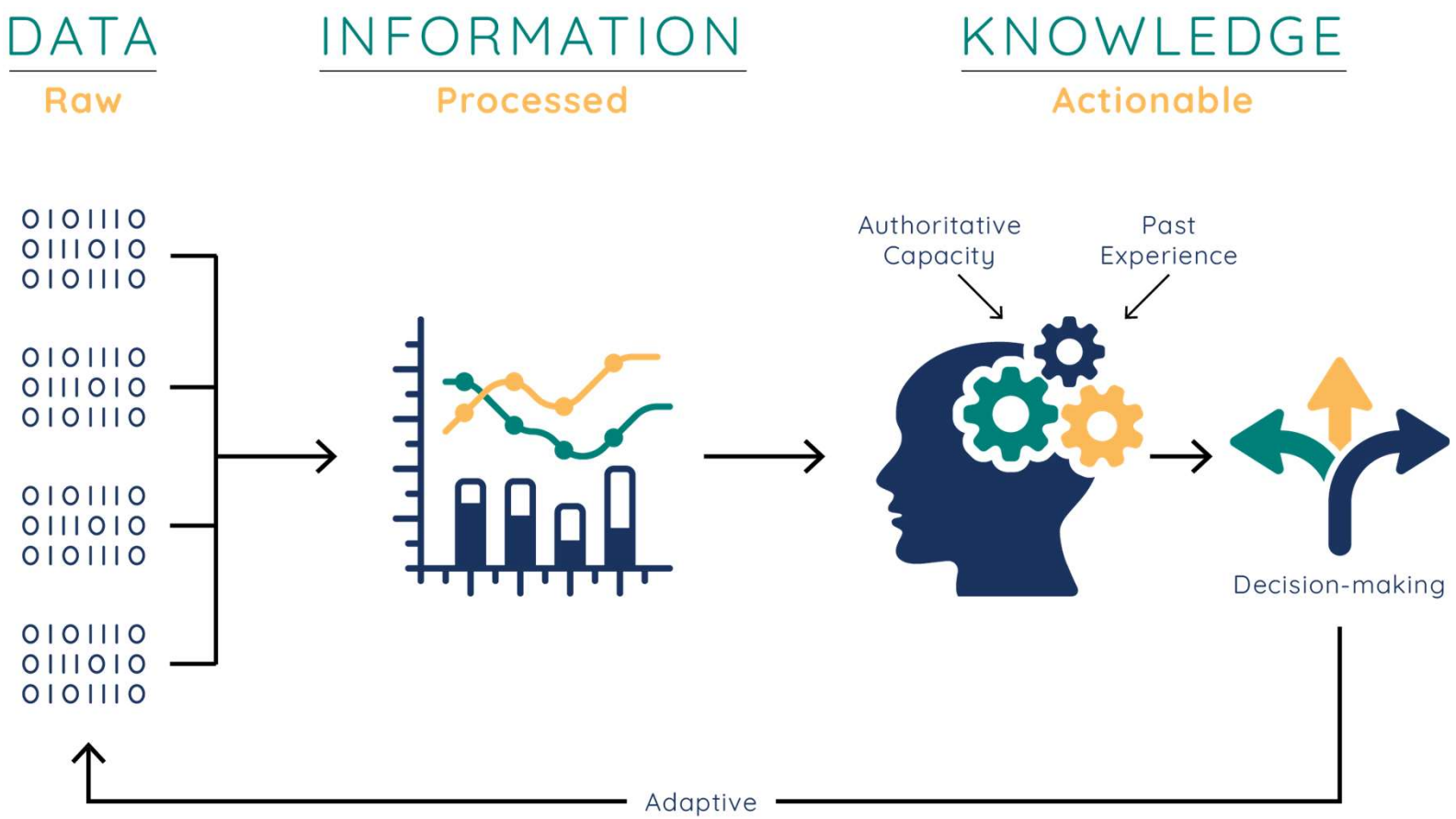
"Machine learning is about extracting knowledge from data."



Andreas C. Müller, Sarah Guido, Introduction to Machine Learning with Python, O'Reilly Media, October 2016

Introduction to Machine Learning

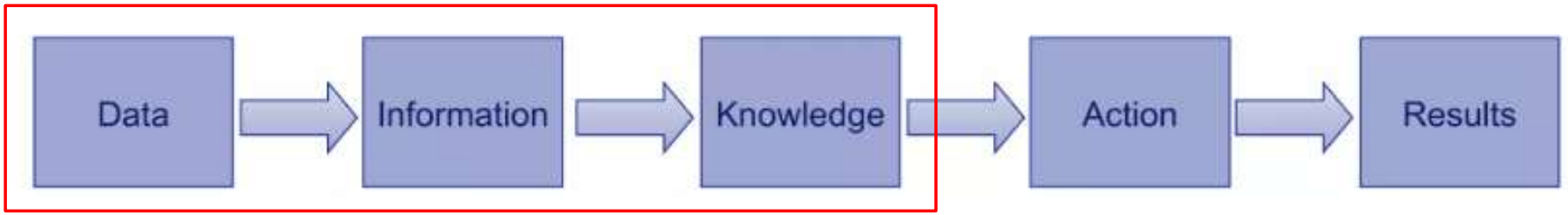
"Machine learning is about extracting knowledge from data."



Internet of Water coalition, What are data, information, and knowledge?, November 2018

Introduction to Machine Learning

"Machine learning is about extracting knowledge from data."



Pattern Recognition

Reinforcement Learning

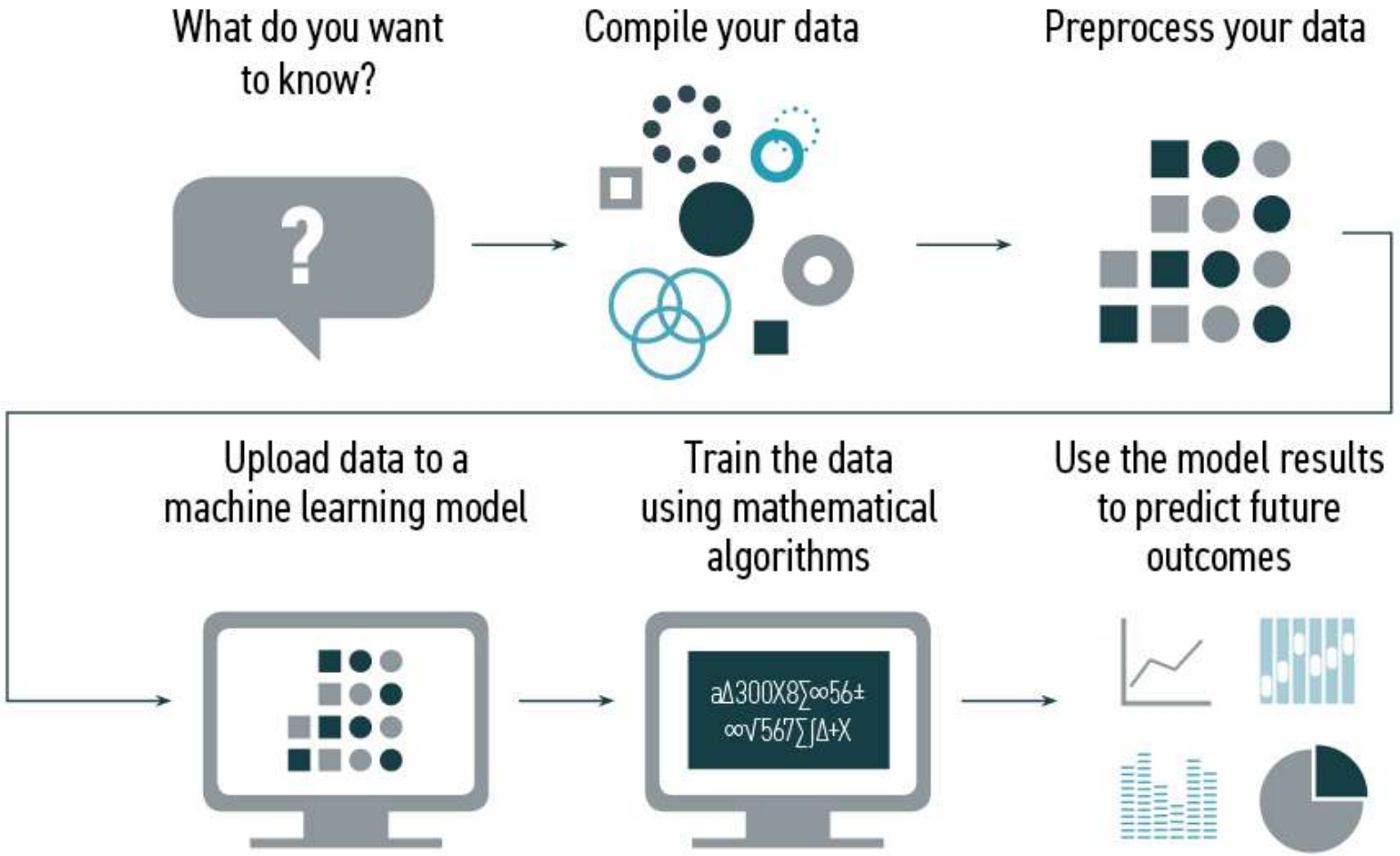
Expertise

- Depends on the organization
- Interpret the information in the right way

Kees de Vos, Knowledge Management, Feb 2002

Introduction to Machine Learning

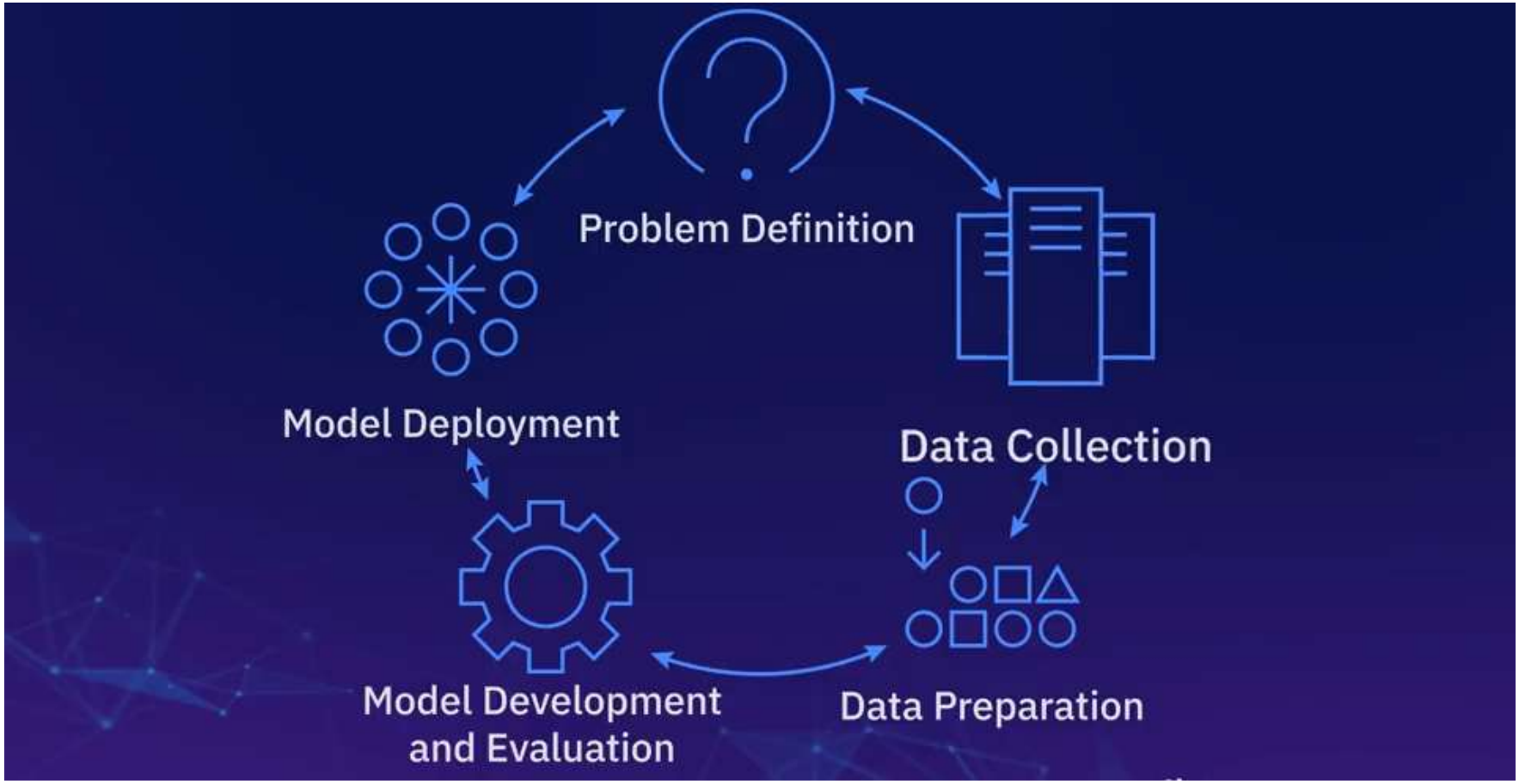
Machine learning's knowledge ≈ a predictive model.



Prediction Models in Machine Learning, 2025

Introduction to Machine Learning

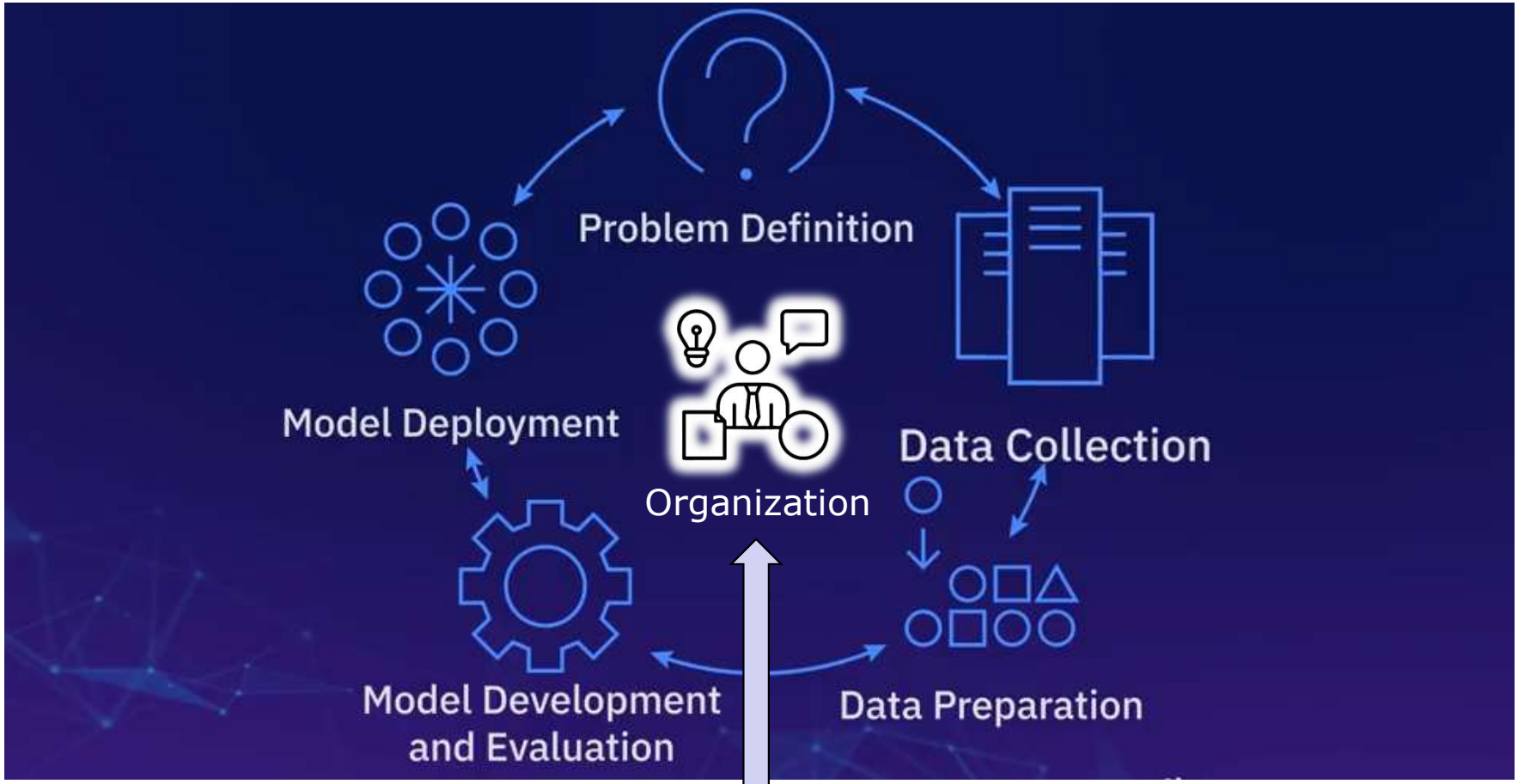
Building a predictive model is a circular process.



Joseph Santarcangelo & Joss Grossman, Machine Learning with Python, IBM Skills network, 2025

Introduction to Machine Learning

Building a predictive model is a circular process.

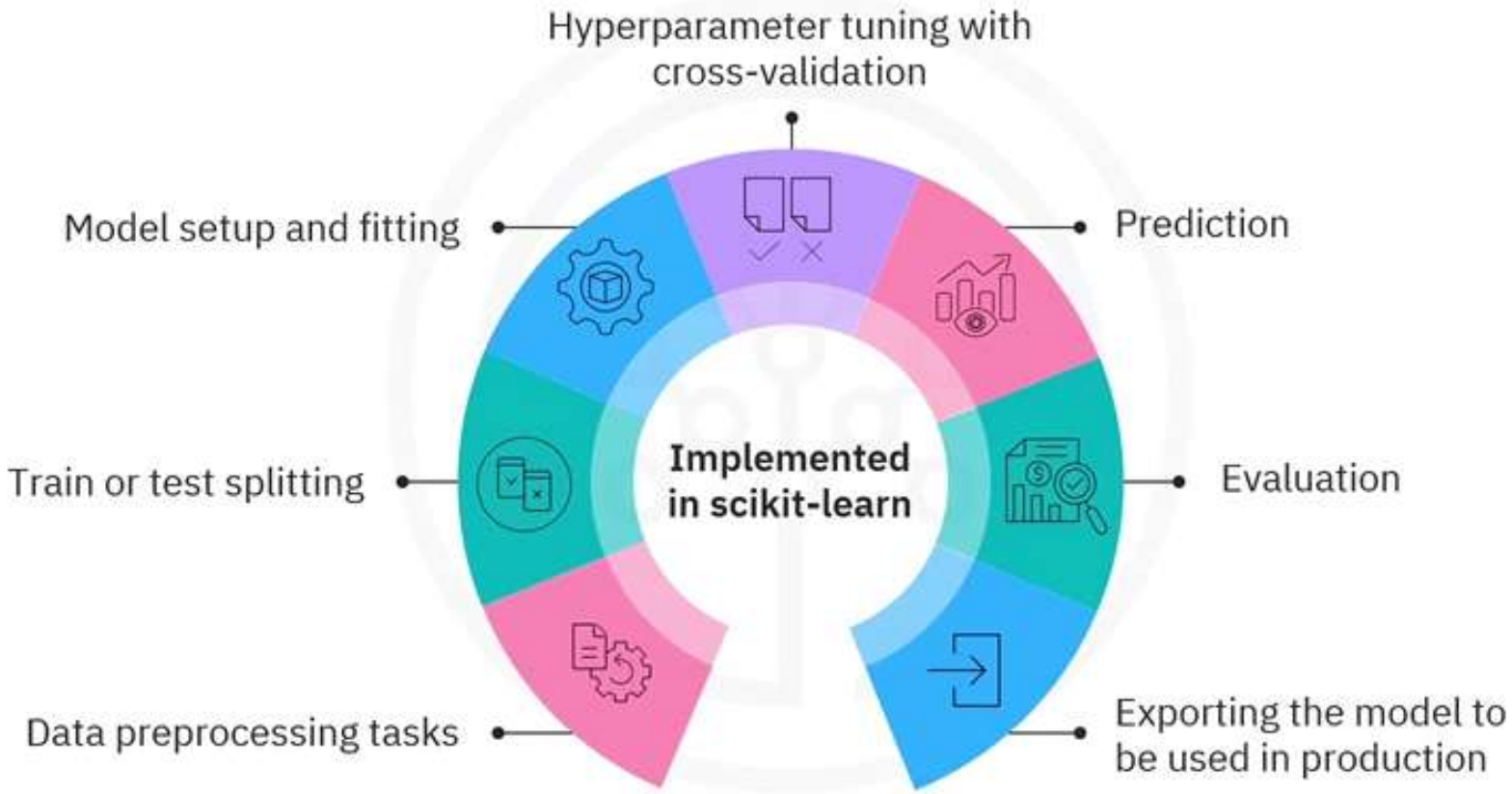


Expertise

- Depends on the organization
- Interpret the information in the right way

Introduction to Machine Learning

Building a predictive model is a circular process.



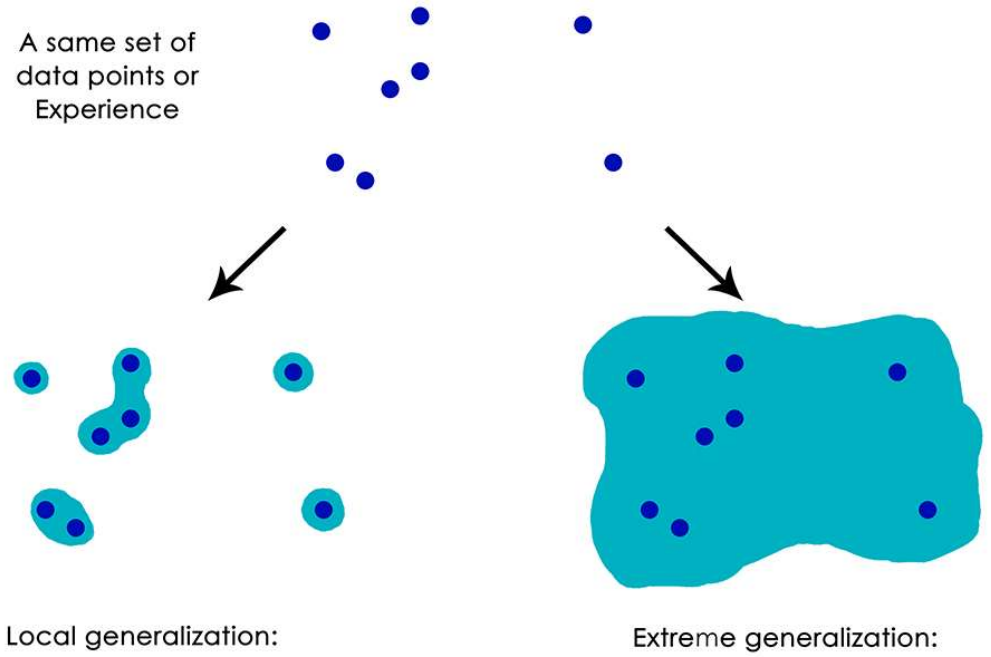
Joseph Santarcangelo & Joss Grossman, Machine Learning with Python, IBM Skills network, 2025

What is Machine Learning?

"*Machine learning algorithms build a model based on [patterns in] sample data, known as training data, in order to make predictions or decisions on new unseen data, without being explicitly programmed to do so.*"

What is Machine Learning?

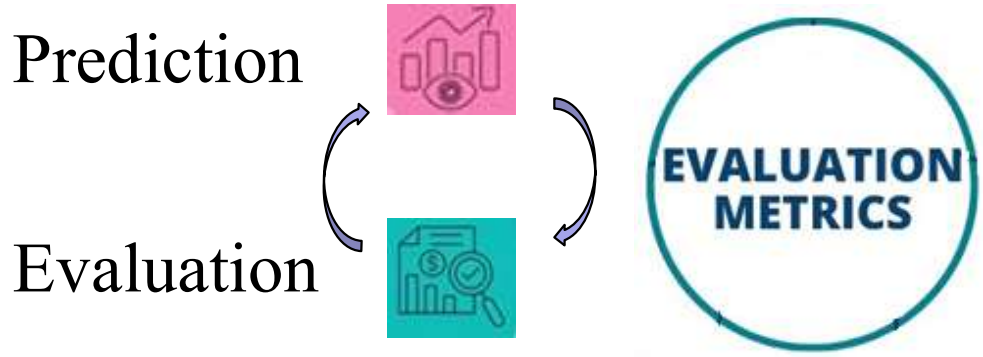
"*Machine learning algorithms build a model based on [patterns in] sample data, known as training data, in order to make predictions or decisions on new unseen data, without being explicitly programmed to do so.*"



Francois Chollet, The limitations of deep learning, 2017

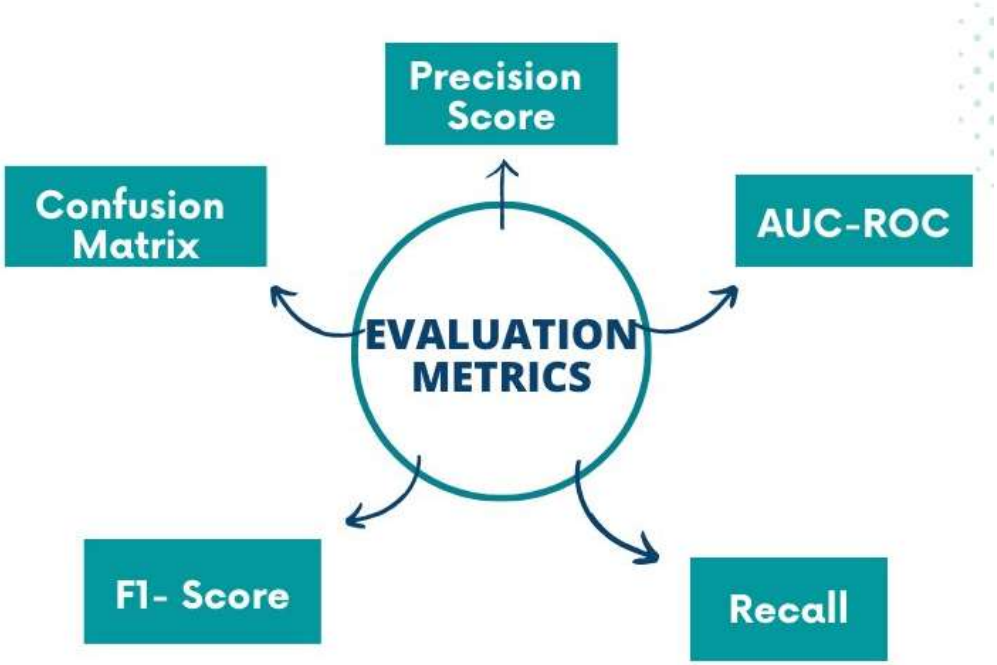
What is Machine Learning?

"*Machine learning algorithms build a model based on [patterns in] sample data, known as training data, in order to make predictions or decisions on new unseen data, without being explicitly programmed to do so.*"



What is Machine Learning?

"*Machine learning algorithms build a model based on [patterns in] sample data, known as training data, in order to make predictions or decisions on new unseen data, without being explicitly programmed to do so.*"

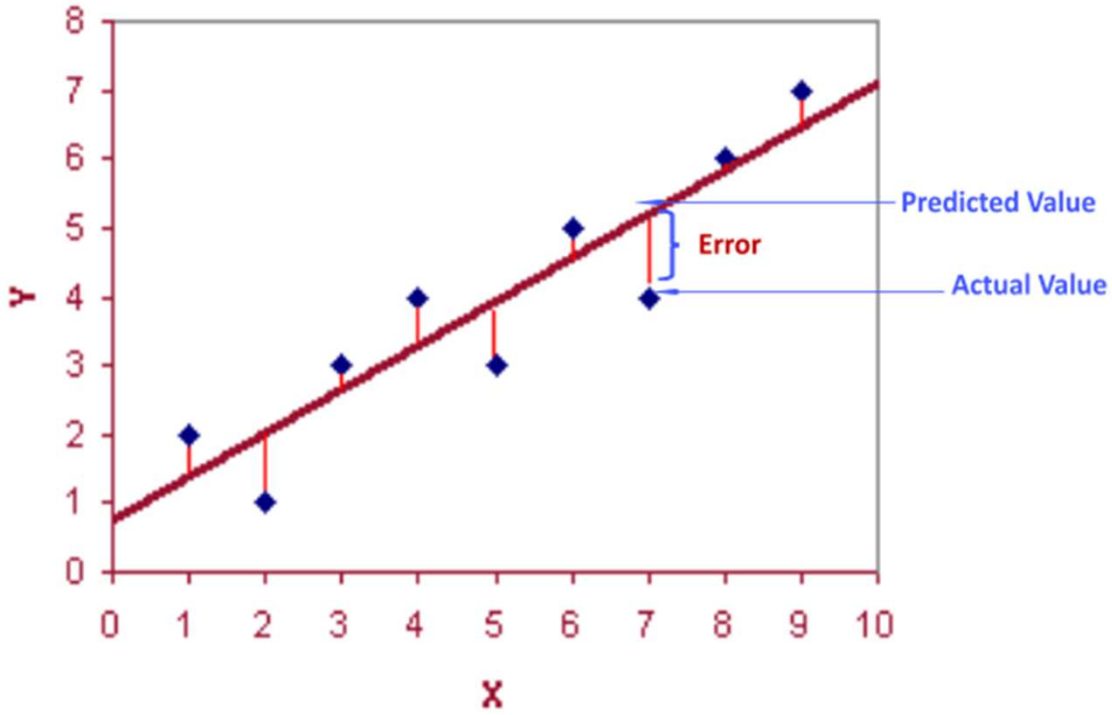


Vikram Singh, Evaluation Metrics in Machine Learning, 2023

ML and evaluation

In machine learning we always **evaluate** models

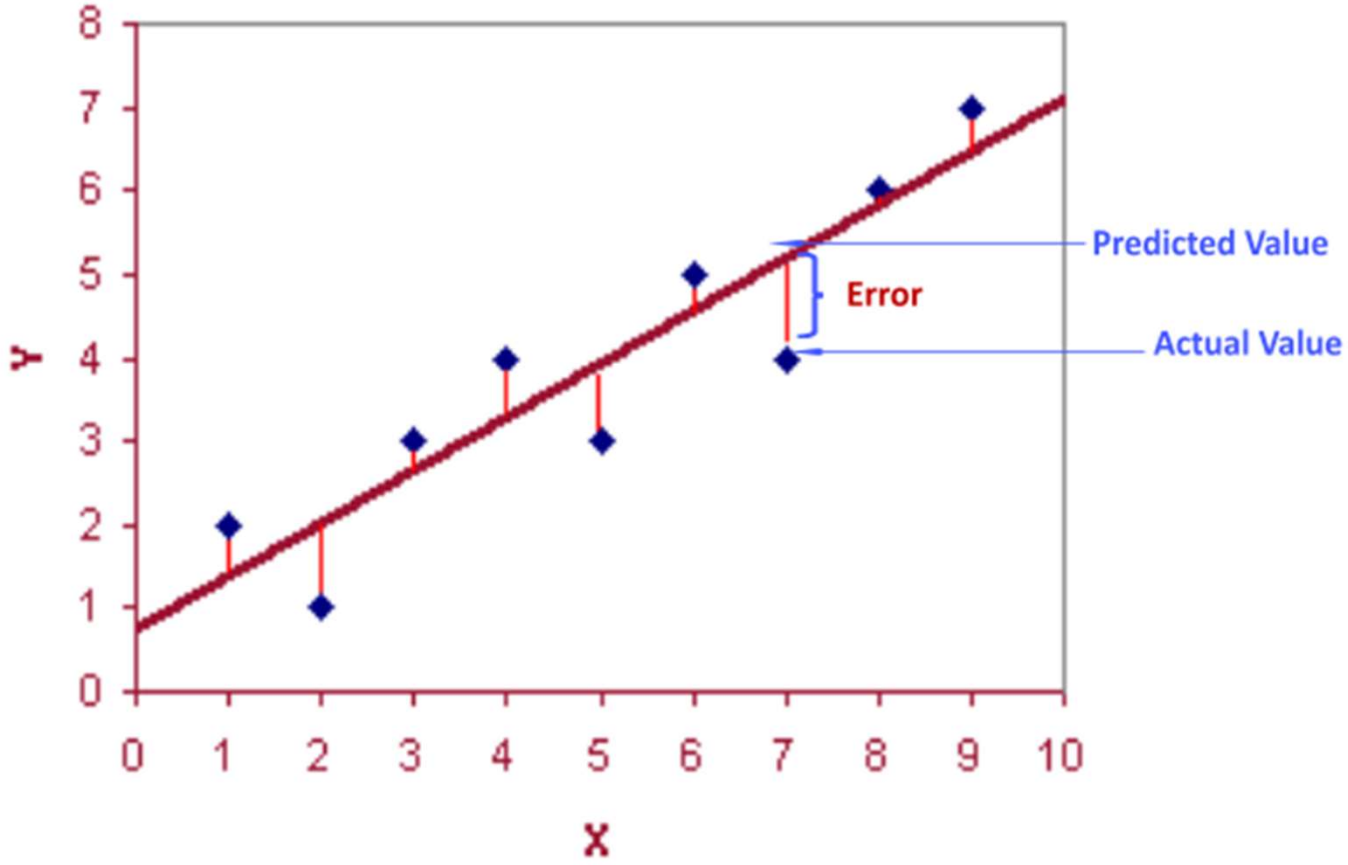
- Classification
  - Accuracy, Precision, Recall, F1
- Regression
  - RMSE



Maarten Marx, Applied Machine Learning, 2024

RMSE

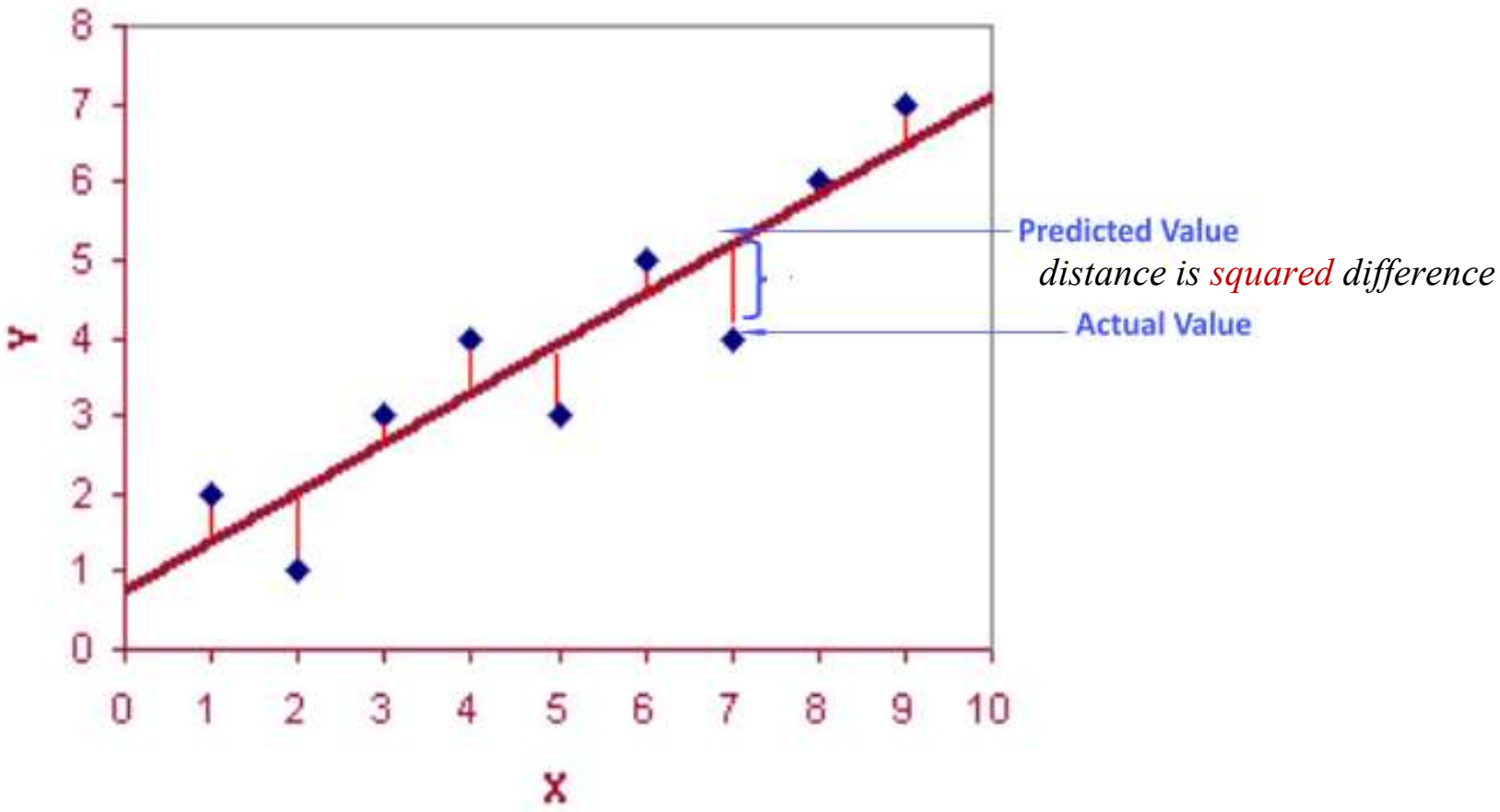
- Root – Mean – Squared – Error



Maarten Marx, Applied Machine Learning, 2024

RMSE

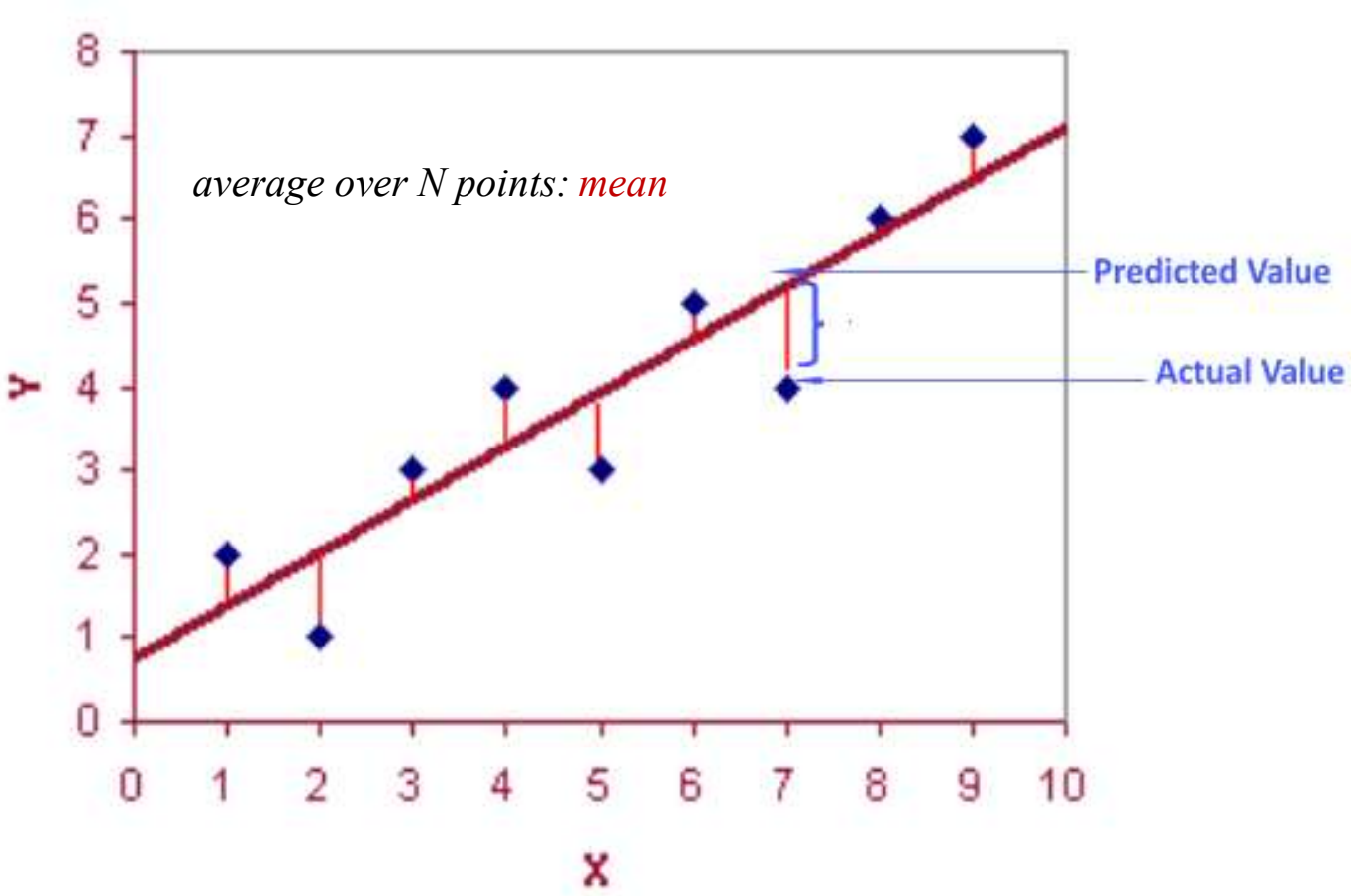
- Root – Mean – Squared – Error



Arnoud Visser, Applied Machine Learning, 2025

RMSE

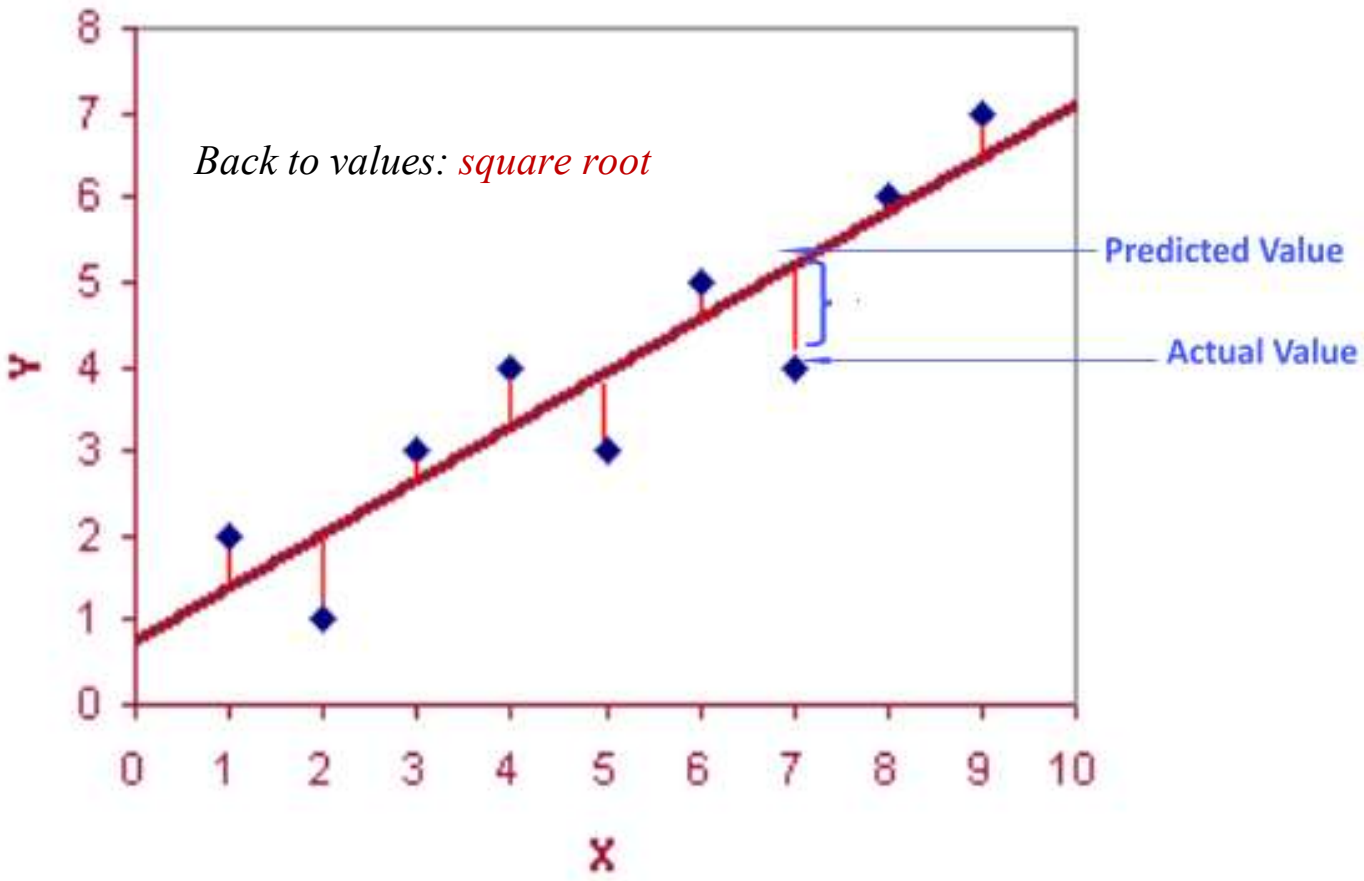
- Root – Mean – Squared – Error



Arnoud Visser, Applied Machine Learning, 2025

RMSE

- Root – Mean – Squared – Error



RMSE

- Root – Mean – Squared – Error

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i \right)^2 \; / \; N}$$

RMSE

- Root – Mean – Squared – Error

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / \ N}$$

distance is squared difference

RMSE

- Root – Mean – Squared – Error

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

average over N points: mean

RMSE

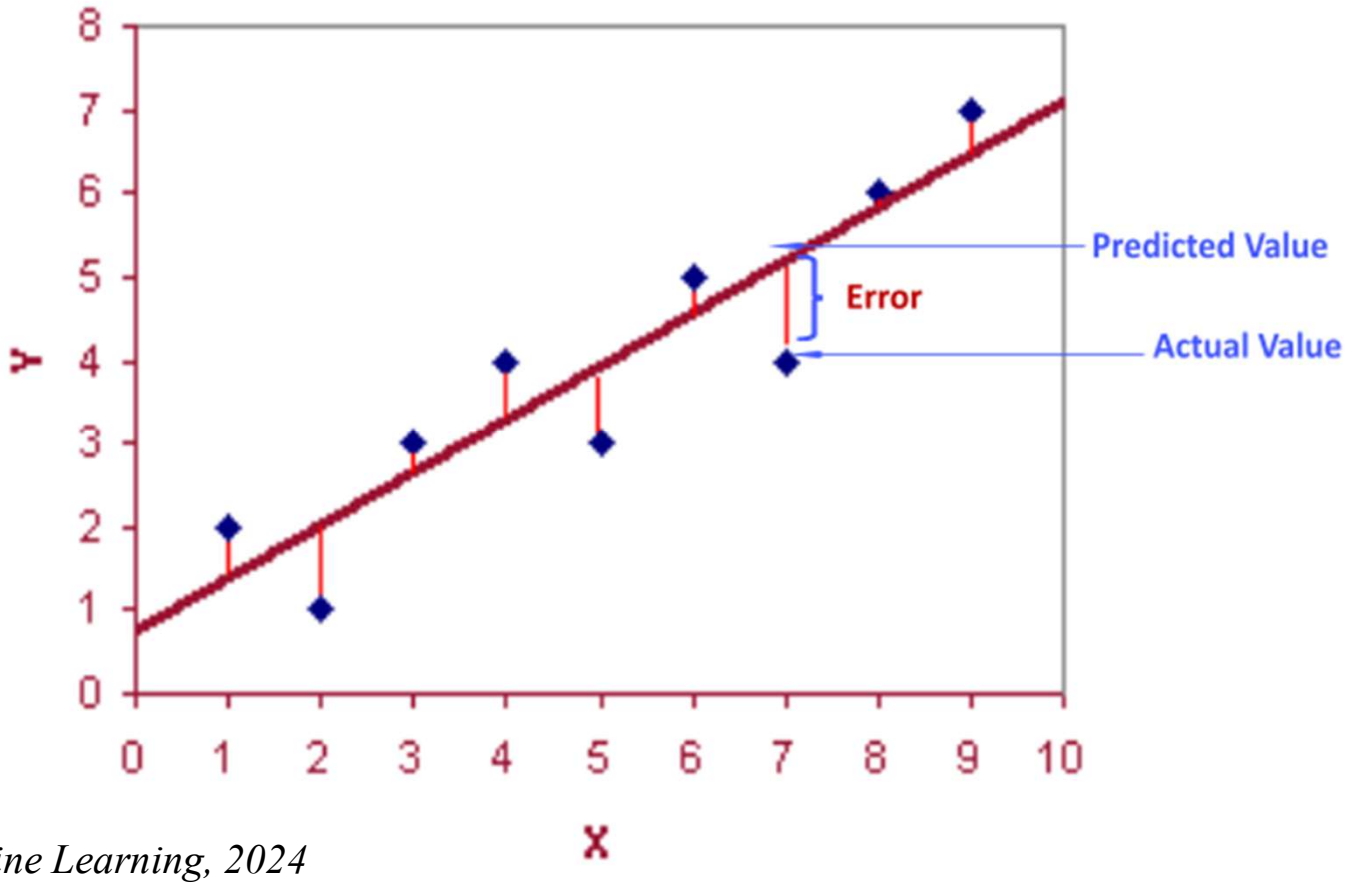
- Root – Mean – Squared – Error

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

Back to values: square root

RMSE

- Program it in Python
- You are given two arrays: `actual_y` and `predicted_y`



RMSE

- Program it in Python
- You are given two arrays: `actual_y` and `predicted_y`

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

RMSE

- Program it in Python
- You are given two arrays: `actual_y` and `predicted_y`

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

Squared

RMSE

- Program it in Python
- You are given two arrays: `actual_y` and `predicted_y`

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

Mean

RMSE

- Program it in Python
- You are given two arrays: `actual_y` and `predicted_y`

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

Root

RMSE

- Program it in Python

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

distance is squared difference

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

RMSE

- Program it in Python

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

average over N points: mean

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

RMSE

- Program it in Python

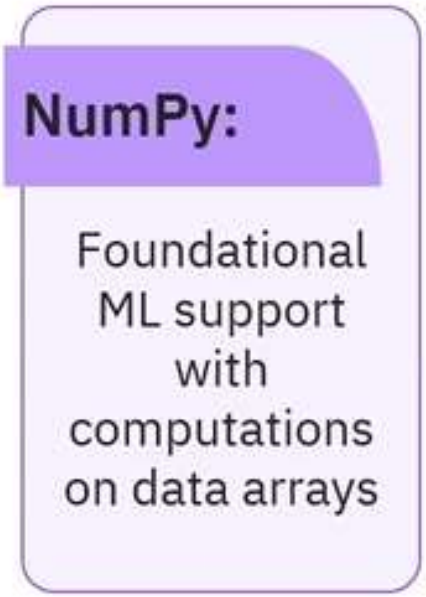
$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

```
from math import sqrt

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) )
                 /
                 N
                )

rmse(actual_y,predicted_y)
```

Back to values: square root

Computation in ML

- Get rid of for-loops
- Vectorized computations
  - Very fast
  - Super elegant



RMSE

- Program it in Python with numpy

```
import numpy as np

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    A,B = np.array(A), np.array(B)
    return sqrt( ((A-B)**2).mean()  )

rmse(actual_y,predicted_y)
```

RMSE

- Program it in Python with numpy

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

distance is squared difference

```
import numpy as np

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    A,B = np.array(A), np.array(B)
    return sqrt( ((A-B)**2).mean()  )

rmse(actual_y,predicted_y)
```

RMSE

- Program it in Python with numpy

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

average over N points: mean

```
import numpy as np

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    A,B = np.array(A), np.array(B)
    return sqrt( ((A-B)**2).mean()  )

rmse(actual_y,predicted_y)
```

RMSE

- Program it in Python with numpy

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 / N}$$

```
import numpy as np

actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]

def rmse(A,B):
    A,B = np.array(A), np.array(B)
    return sqrt( ((A-B)**2).mean()  )

rmse(actual_y,predicted_y)
```

Back to values: square root

0.7745966692414834

RMSE

- Program it in Python

```
actual_y= [2,3,5,2,4]
predicted_y = [2,2,4,2,5]
```

```
from math import sqrt

def rmse(A,B):
    N= len(A)
    return sqrt( sum( (A[i]-B[i])**2  for i in range(N) ) / N  )

rmse(actual_y,predicted_y) )
```

→ 0.7745966692414834

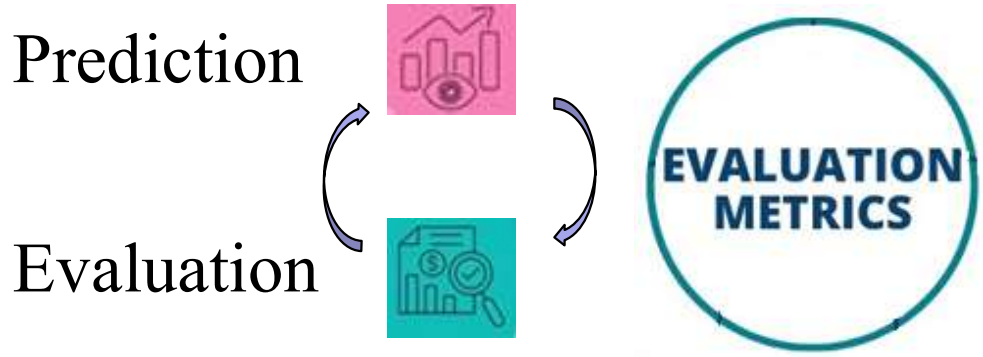
```
import numpy as np

def rmse(A,B):
    A,B = np.array(A), np.array(B)
    return sqrt( ((A-B)**2).mean()  )

rmse(actual_y,predicted_y)
```

→ 0.7745966692414834

Maarten Marx, Applied Machine Learning, 2024

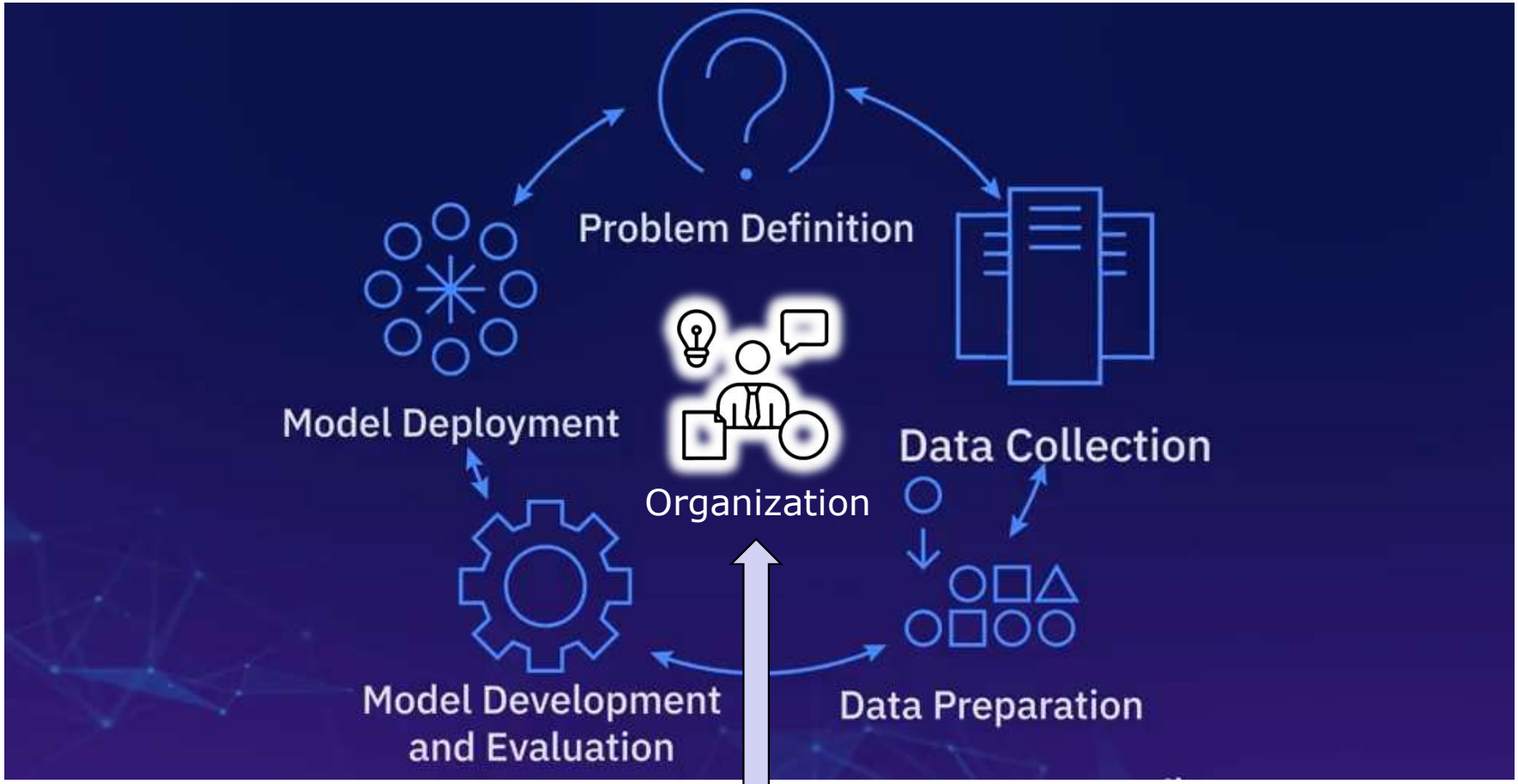
Machine Learning: minimize the error

"*Machine learning algorithms build a model based on [patterns in] sample data, known as training data, in order to make predictions or decisions on new unseen data, with an evaluation metric as implicit measure what is expected.*"



The right way

Building a predictive model is a circular process.



Expertise

- Depends on the organization
- Interpret the information in the right way

The right way

What model is optimized?

What is the problem to be solved?

Expertise

- Depends on the organization
- Interpret the information in the right way

Start with a baseline

What is the problem to be solved?

What model is optimized?

Expertise

- Depends on the organization
- Interpret the information in the right way

Start with a baseline

What is the weight of a child?

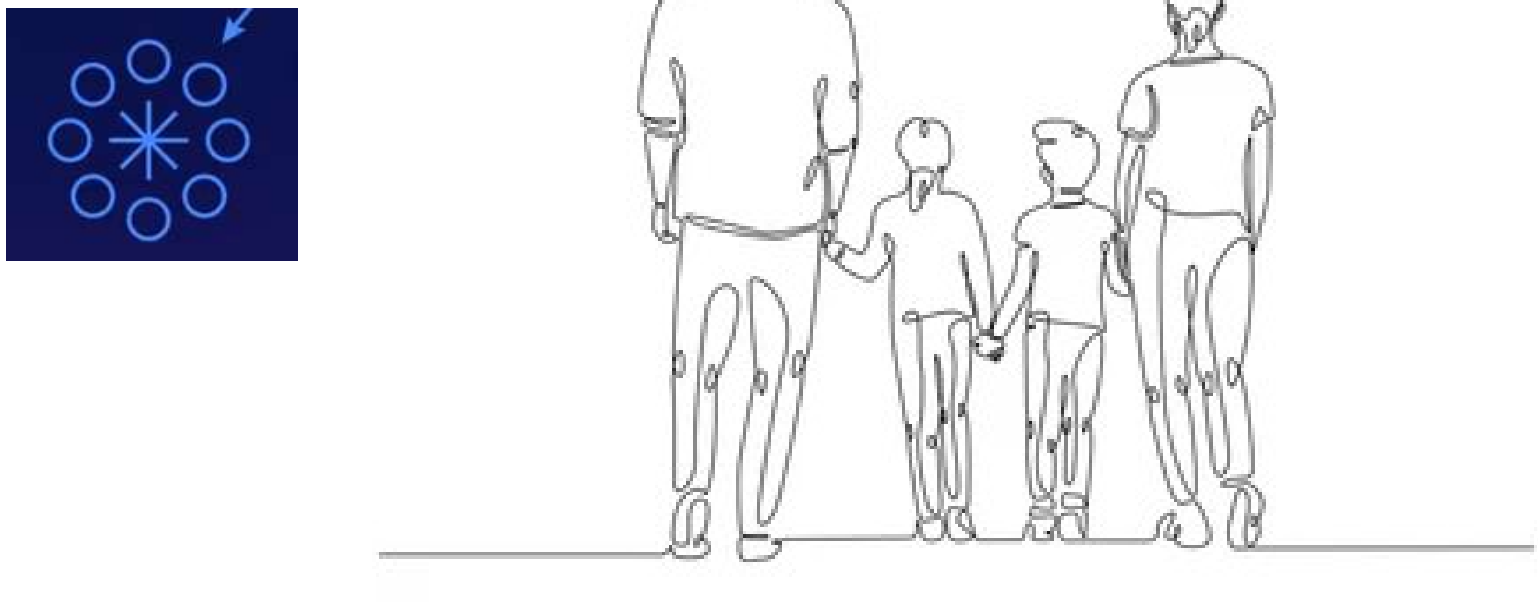
What could be the simplest model?

Expertise

- Depends on the organization
- Interpret the information in the right way

Baseline model

What the weight of a kid?

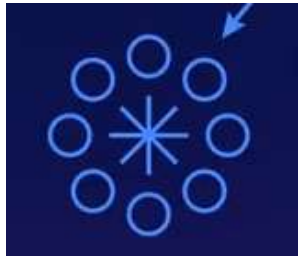


Expertise

- Depends on the organization
- Interpret the information in the right way

Baseline model

What the weight of a child?



Half size of adult → Half weight of adult

Expertise

- Depends on the organization
- Interpret the information in the right way

Baseline model – version 0.0

A child is half the weight of an adult

Average adult
~80 kg

Average child
~40 kg

Expertise

- Depends on the organization
- Interpret the information in the right way

Baseline model – version 0.0

A child is half the weight of an adult

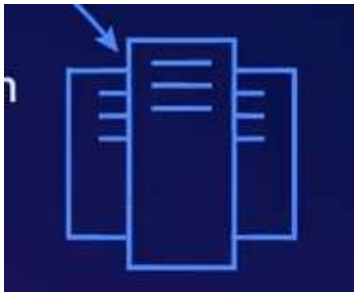
- Your model is based on “prejudice”
- It does not use any feature of the child

Expertise

- Depends on the organization
- Interpret the information in the right way

Evaluate baseline model – version 0.0

A child is half the weight of an adult



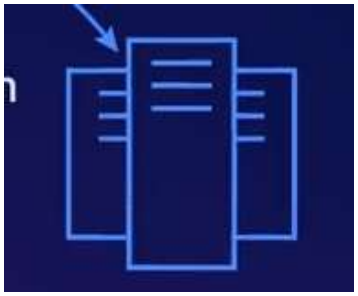
Dataset

Expertise

- Depends on the organization
- Interpret the information in the right way

Evaluate baseline model – version 0.0

A child is half the weight of an adult



regensburg_pediatric_appendicitis

Dataset

Description

This dataset was acquired in a retrospective study from a cohort of pediatric patients admitted with abdominal pain to Children's Hospital St. Hedwig in Regensburg, Germany. Multiple abdominal B-mode ultrasound images were acquired for most patients, with the number of views varying from 1 to 15. The images depict various regions of interest, such as the abdomen's right lower quadrant, appendix, intestines, lymph nodes and reproductive organs. Alongside multiple US images for each subject, the dataset includes information encompassing laboratory tests, physical examination results, clinical scores, such as Alvarado and pediatric appendicitis scores, and expert-produced ultrasonographic findings. Lastly, the subjects were labeled w.r.t. three target variables: diagnosis (appendicitis vs. no appendicitis), management (surgical vs. conservative) and severity (complicated vs. uncomplicated or no appendicitis). The study was approved by the Ethics Committee of the University of Regensburg (no. 18-1063-101, 18-1063_1-101 and 18-1063_2-101) and was performed following applicable guidelines and regulations.

Evaluate baseline model – version 0.0

A child is half the weight of an adult

56 Features

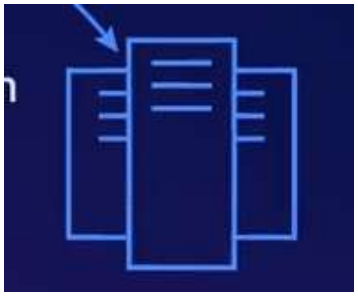
	Feature Name	Type	Distinct/Missing Values
?	Management (target)	string	4 distinct values 1 missing attributes
▬	Age	numeric	577 distinct values 1 missing attributes
▬	BMI	numeric	510 distinct values 27 missing attributes
?	Sex	string	2 distinct values 2 missing attributes
▬	Height	numeric	187 distinct values 26 missing attributes
▬	Weight	numeric	268 distinct values 3 missing attributes

regensburg_pediatric_appendicitis

Dataset

Evaluate baseline model – version 0.0

A child is half the weight of an adult



regensburg_pediatric_appendicitis

Dataset

43.17 kg

```
import numpy as np
from sklearn.datasets import fetch_openml

patients = fetch_openml(name='regensburg_pediatric_appendicitis', version=1)
w = np.array(patients.data.Weight))
w = w[~np.isnan(h)]

print("Average weight of a child at Regensburg hospital is %.2f kg" % w.mean())
```

Evaluate Baseline model – version 0.0

A child is half the weight of an adult

regensburg_pediatric_appendicitis

Model

40 kg

Dataset

43.17 kg

- Only 8% off
- Our model is based on “prejudice”
- It does not use any feature of the child

Germany	85.9 kg (189.4 lb)	69.2 kg (152.6 lb)	18+	Self-reported	2021

Evaluate Baseline model – version 0.1

A child is half the weight of a German adult

regensburg_pediatric_appendicitis

Model
38.78 kg

Dataset
43.17 kg

- Now it is 11% off
- Our model is based on "prejudice"
- It does not use any feature of the child

77.55 kg

Germany	85.9 kg (189.4 lb)	69.2 kg (152.6 lb)	18+	Self-reported	2021

Evaluate Baseline model – version 0.1.1

Use your first ML model

regensburg_pediatric_appendicitis

Dataset

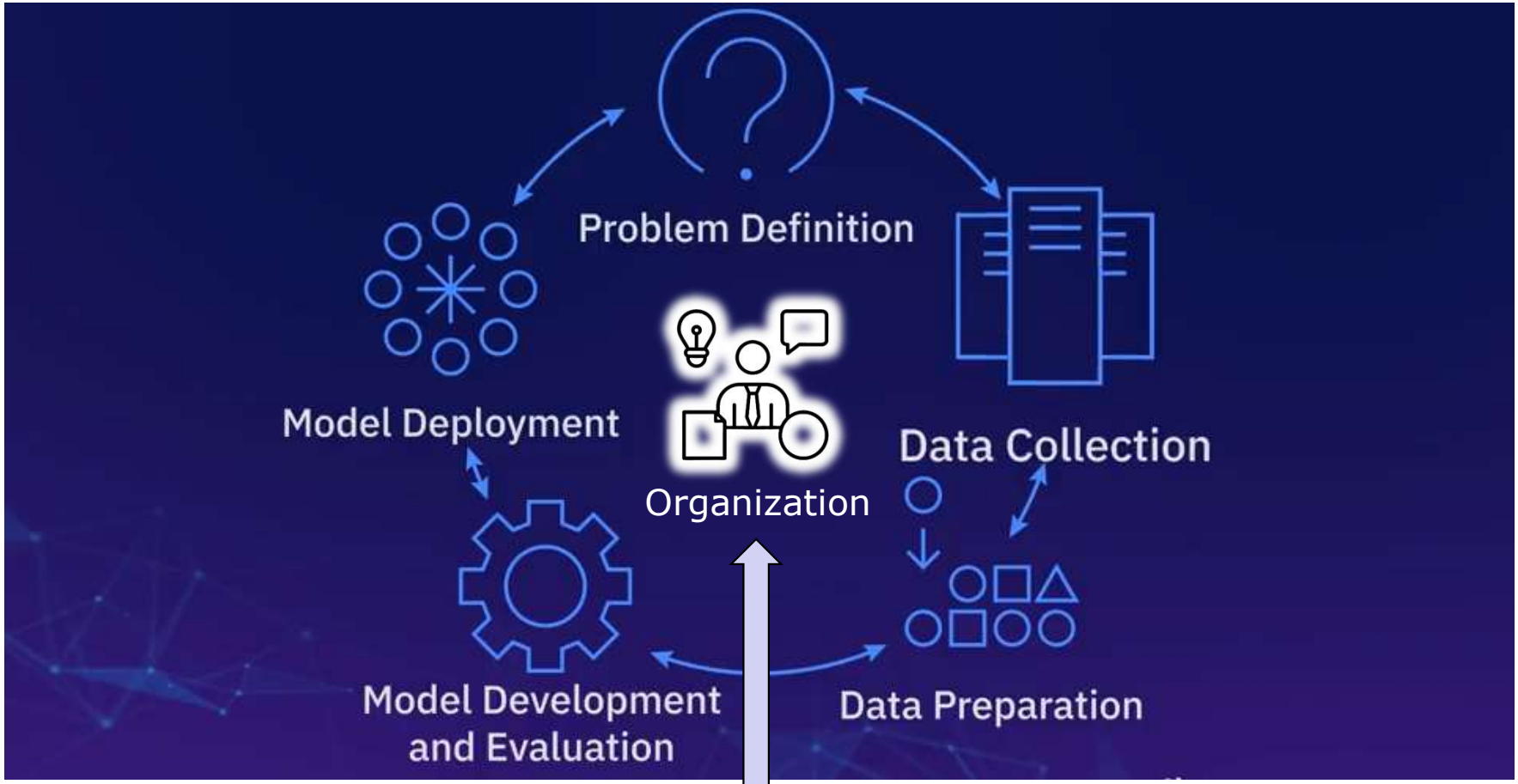
43.17 kg

```
from sklearn.dummy import DummyRegressor

model = DummyRegressor(strategy="mean")
model.fit(X_train, y_train)

print("Average weight of a child at Regensburg hospital is %.2f kg" % model.predict(X_test[0]))
```

43.65 kg

The right way

Building a predictive model is a circular process.



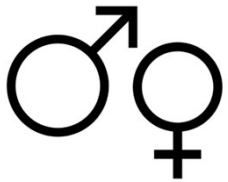
Expertise

- Depends on the organization
- Interpret the information in the right way

Evaluate informed models – version 0.2

regensburg_pediatric_appendicitis

Gender



Dataset

43.17 kg

- What happens when we add features of the child

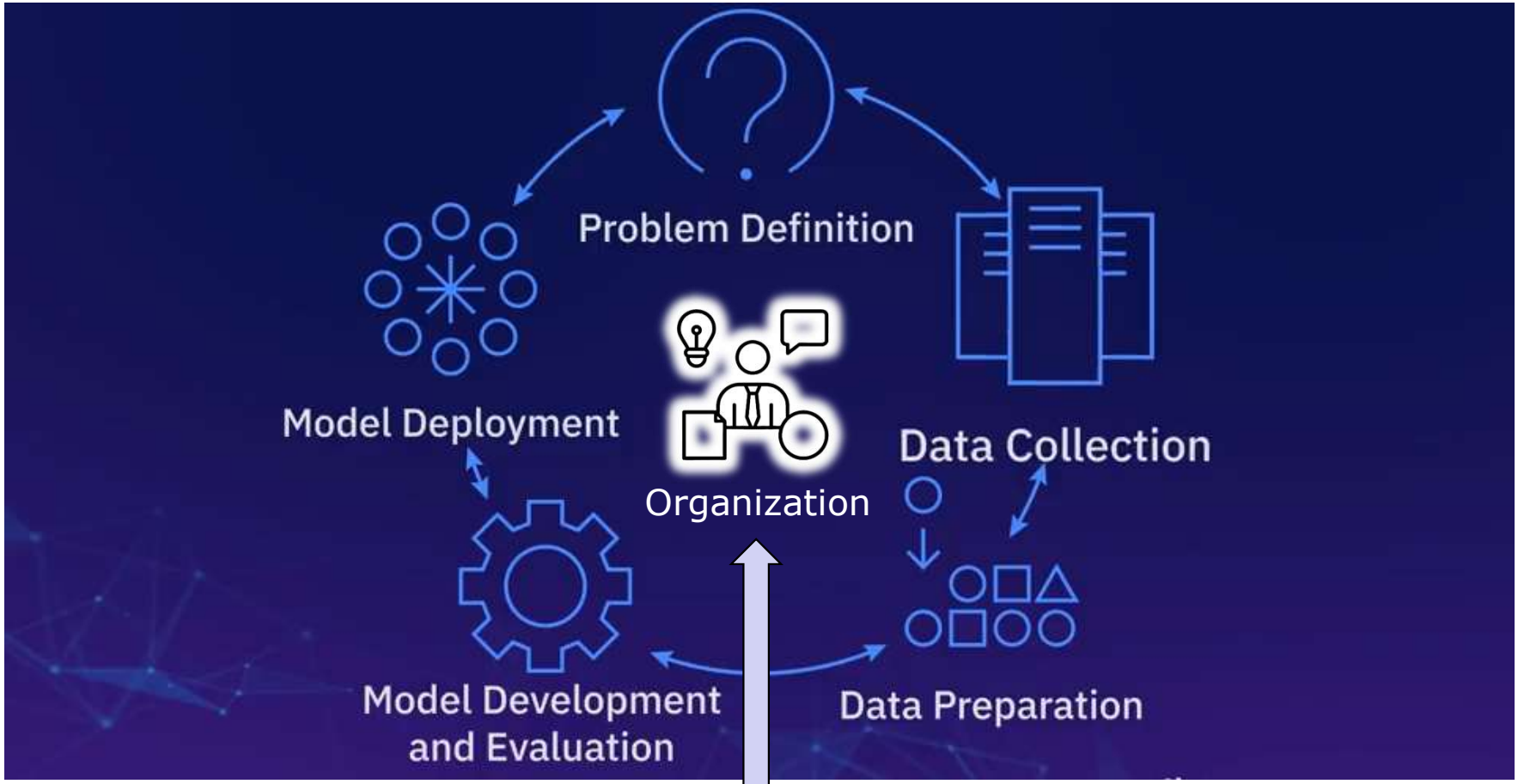
41.23 kg

45.28 kg

Germany	85.9 kg (189.4 lb)	69.2 kg (152.6 lb)	18+	Self-reported	2021

The right way

Building a predictive model is a circular process.



Expertise

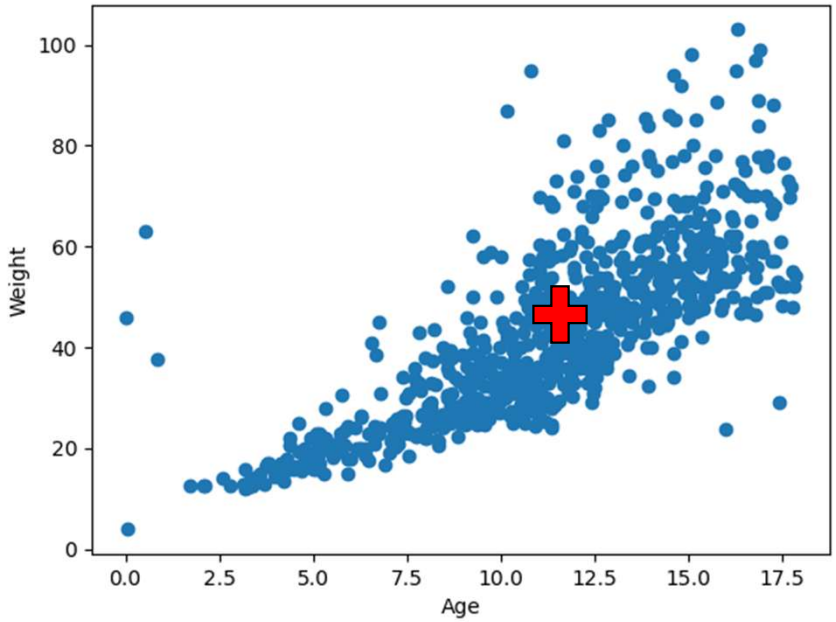
- Depends on the organization
- Interpret the information in the right way

Evaluate informed models – version 0.3

- What happens when we add features of the child

regensburg_pediatric_appendicitis

Age



Dataset

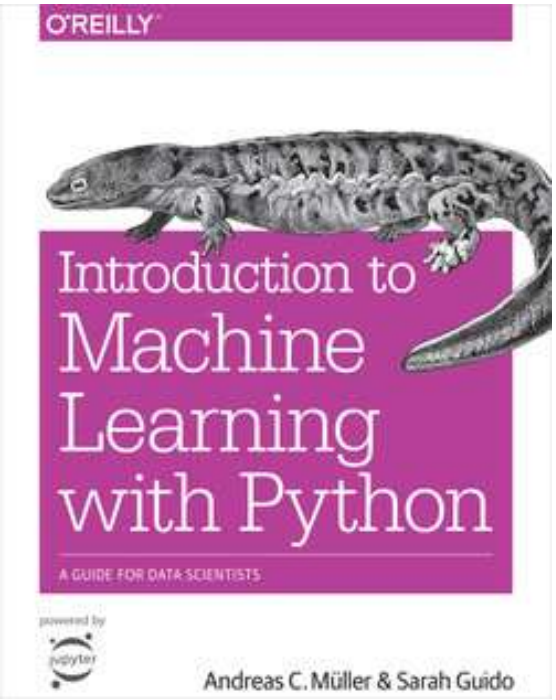
43.17 kg

- Age is continues value

Regression model

Introduction to Machine Learning

- 8.1 Approaching a Machine Learning Problem
- 8.2 From Prototype to Production
- 8.3 Testing Production Systems



Andreas C. Müller, Sarah Guido, Introduction to Machine Learning with Python, O'Reilly Media, October 2016

Conclusion

Learning outcomes of this course covered today

- What is Machine Learning?
- What can it do and what not?
- How to evaluate the quality of an ML model