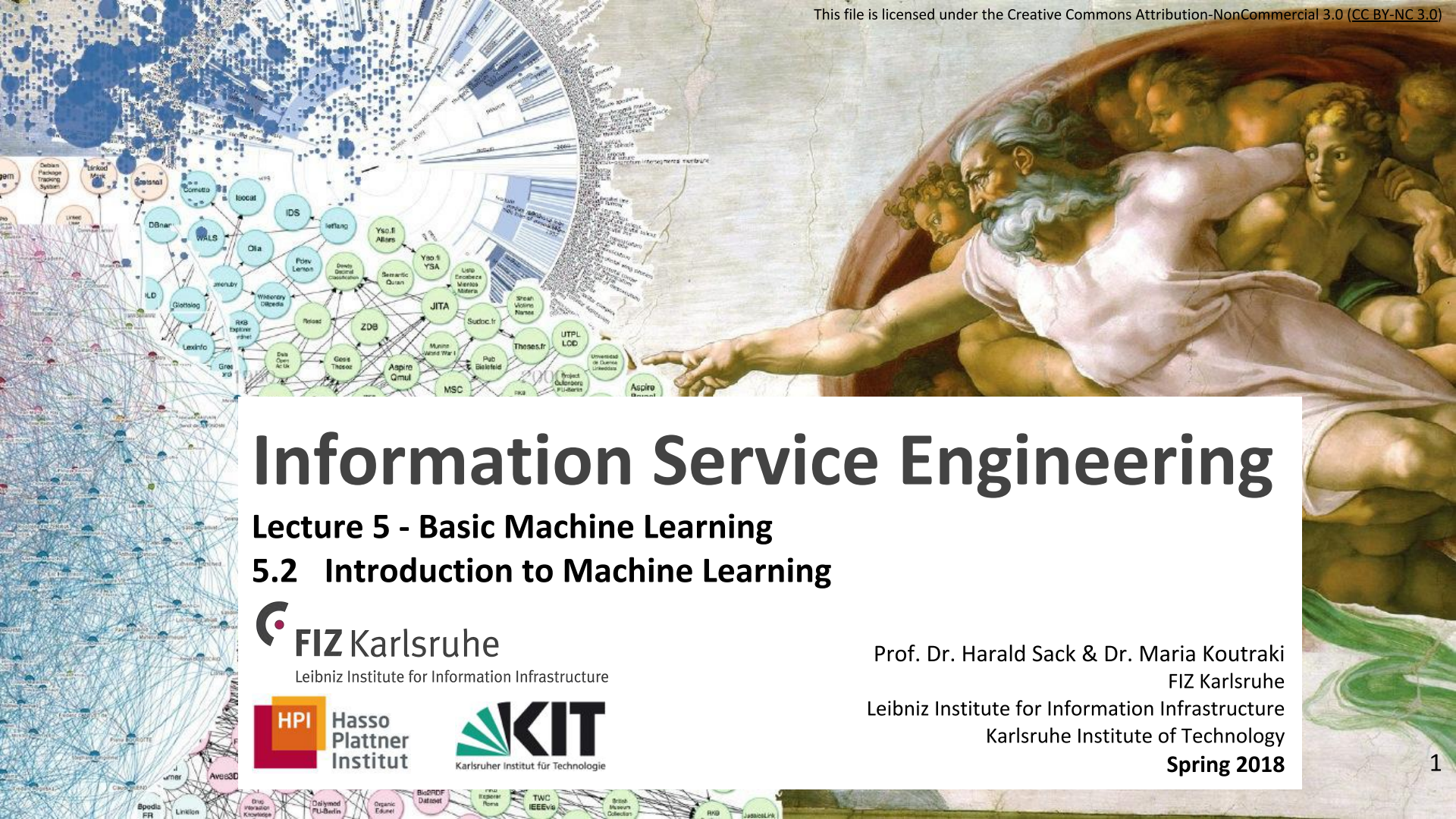




Information Service Engineering

Lecture 5 - Basic Machine Learning

5.2 Introduction to Machine Learning



FIZ Karlsruhe

Leibniz Institute for Information Infrastructure



Hasso
Plattner
Institut



Karlsruher Institut für Technologie

Prof. Dr. Harald Sack & Dr. Maria Koutraki
FIZ Karlsruhe

Leibniz Institute for Information Infrastructure
Karlsruhe Institute of Technology
Spring 2018

Information Service Engineering

Lecture 5: Basic Machine Learning

5.1 A Brief History of AI

5.2 Introduction to Machine Learning

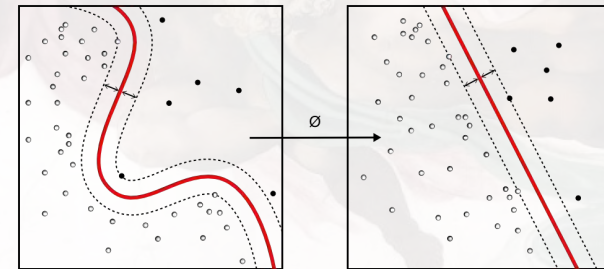
5.3 Evaluation and Generalization Problems

5.4 Basic ML Algorithms

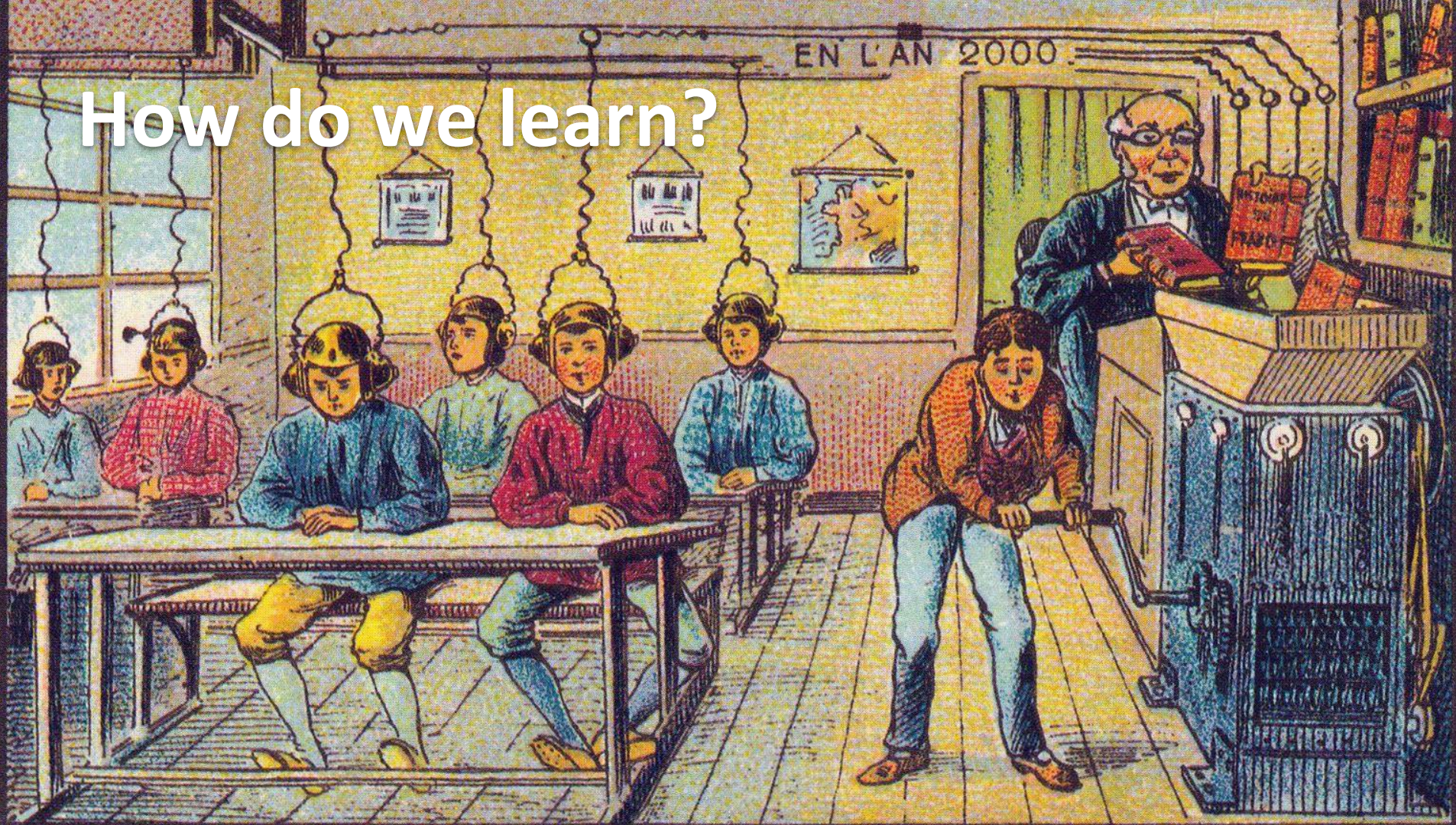
5.5 Basic ML Algorithms II

5.6 Unsupervised Learning

5.7 Deep Learning



How do we learn?



How do we learn?

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

- Recognising that last time we were in this situation (*saw this data*)
- we tried out some particular action (*gave this output*) and
- it worked (*was correct*), so we'll try it again,
- or it didn't work (*was not correct*), so we'll try something different.



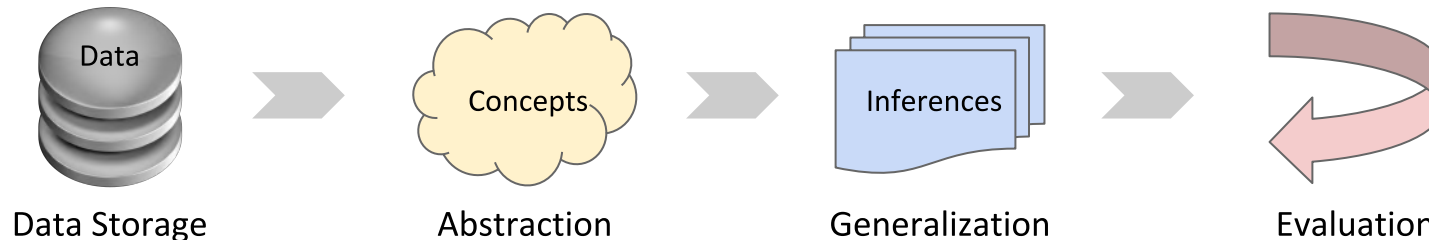
- we make an **observation**,
- we **remember**,
- we **adapt**,
- and we **generalize**

The Basic Learning Process

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

The **Basic Learning Process** is similar regardless whether the learner is human or machine:

- **Data storage** utilizes observation, memory, and recall to provide a factual basis for further reasoning.
- **Abstraction** involves the translation of stored data into broader representations and concepts.
- **Generalization** uses abstracted data to create knowledge and inferences that drive action in new contexts.
- **Evaluation** provides a feedback mechanism to measure the utility of learned knowledge and inform potential improvements.

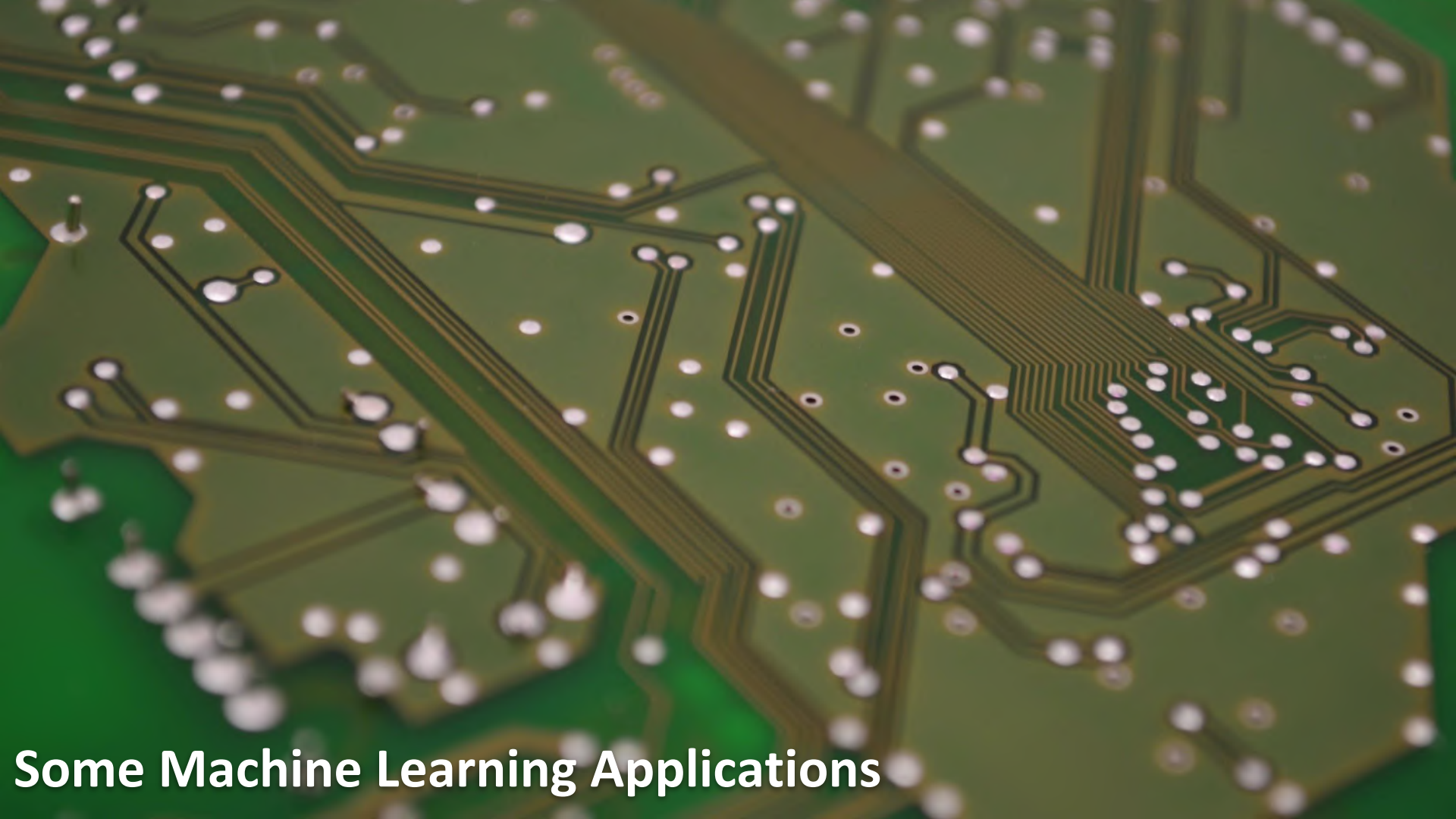


What is Machine Learning?

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

Definition:

A computer program is said to learn from **experience** E with respect to some class of **tasks** T and **performance measure** P , if its performance at tasks in T , as measured by P , improves with experience E .



Some Machine Learning Applications

Collaborative Filtering

Recommend item(s) to users that are of interest for them

- A. Always recommend the most popular items. This strategy works for a significant number of cases.

OR

- B. **Predict interesting item(s) based on previous actions and decisions of “similar” users**

Machine Learning Applications

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

Automated Translation

Translation of a text from a source language into a target language while preserving its semantics

- A. Fully understand a text before translating it using a curated set of rules crafted by a computational linguist well versed in the two languages we would like to translate

or

- B. **Use examples of translated texts in the two languages we would like to translate to “learn” how to translate**

Machine Learning Applications

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

Named Entity Recognition

Identifying named entity types in natural language texts, such as persons, places, organisations, times, etc.

A. Use hand-crafted rules to filter out entity types

or

B. Use examples of marked-up documents to learn dependencies automatically to identify entity types

Visual Concept Detection

Assign descriptive categories to images based on their visual content

- A. Use hand-crafted rules to be applied on features extracted from the images to classify the image

or

- B. **Learn to classify the images from examples of marked-up images. Even learn the best suited features to be extracted from the image examples**

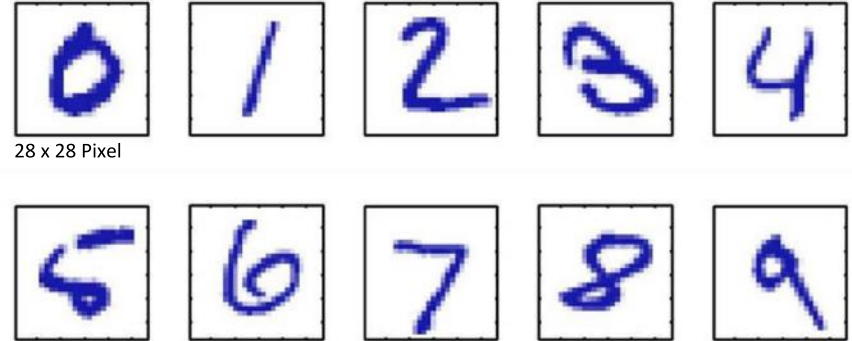
Machine Learning Applications

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

Hand-written Digit Recognition

Assign the correct value to a hand-written digit.

- Represent the input image as a vector $x \in \mathbb{R}^{784}$
- Learn a classifier $f(x)$ such that,
 $f: x \rightarrow \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$





Machine Learning Categories

Again - what is Machine Learning?

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine LearningA

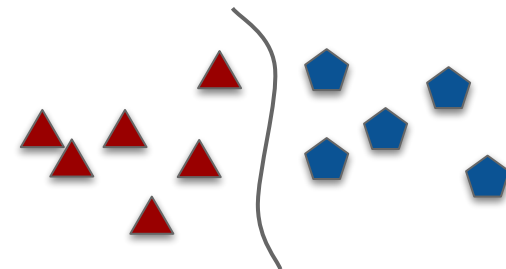
- Algorithms that can improve their performance using training data
- Typically the algorithm has a
 - **(large) number of parameters**
 - **whose values are learnt from the data**
- It can be applied in situations, where it is very **challenging (= impossible) to define rules by hand**

Classification Problems

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

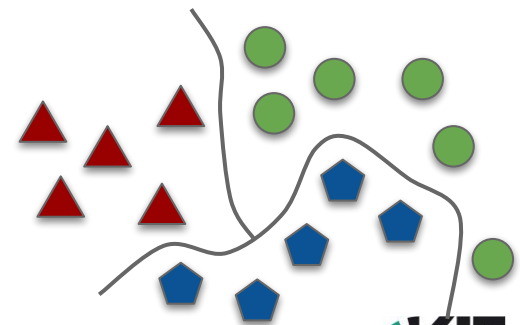
- **Binary Classification**

- given a pattern x drawn from a domain X , estimate which value an associated binary random variable $y \in \{\pm 1\}$ will assume.



- **Multi Label Classification**

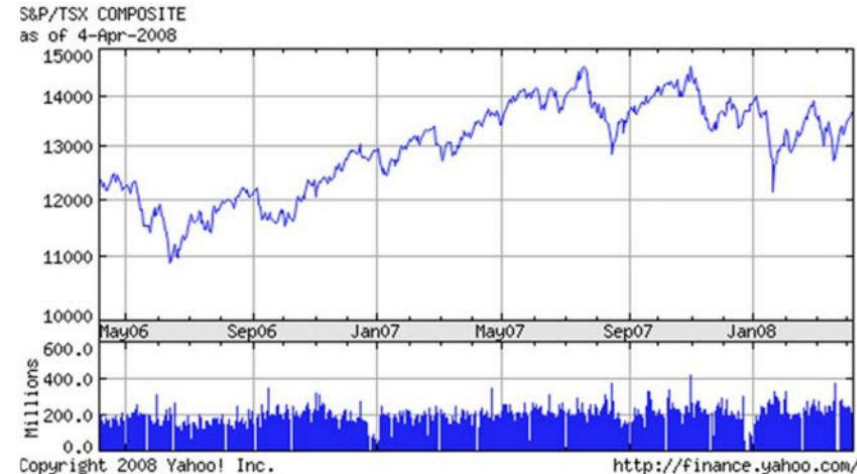
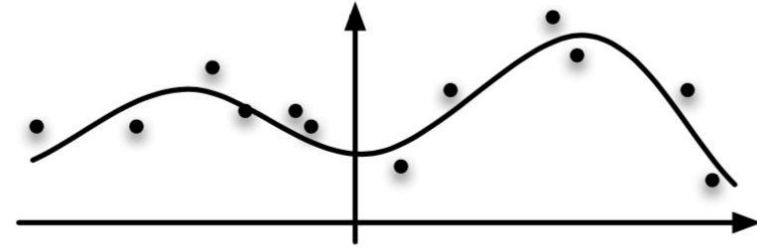
- given a pattern x drawn from a domain X , estimate which value an associated binary random variable $y \in \{1, \dots, n\}$ will assume.



Regression Problems

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

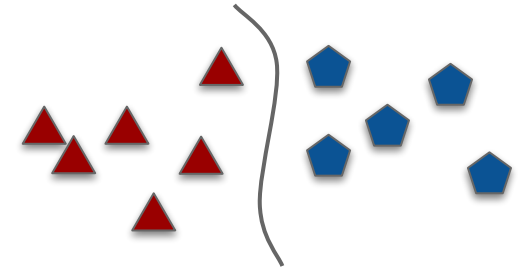
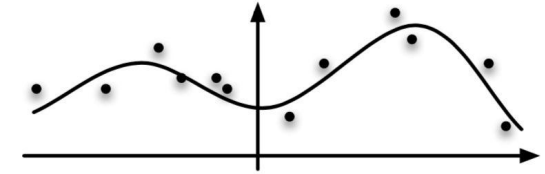
- The Goal of Regression is to estimate a real-valued variable $y \in \mathbb{R}$ given a pattern x
- Example:
Prediction of stock prices for a future date



Classification vs. Regression Problems

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

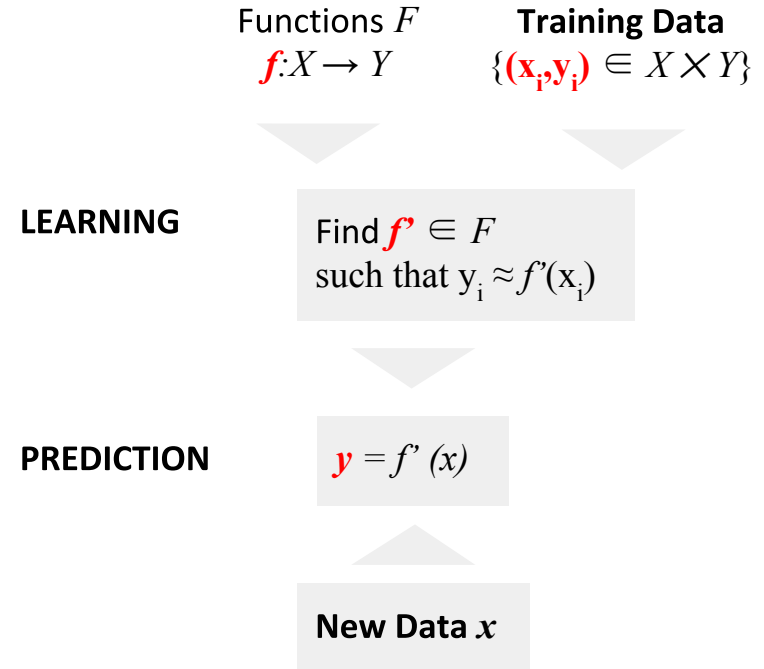
- Problems with a **quantitative response** most times are considered as **regression problems**.
- Problems involving a **qualitative response** are often referred to as **classification problems**



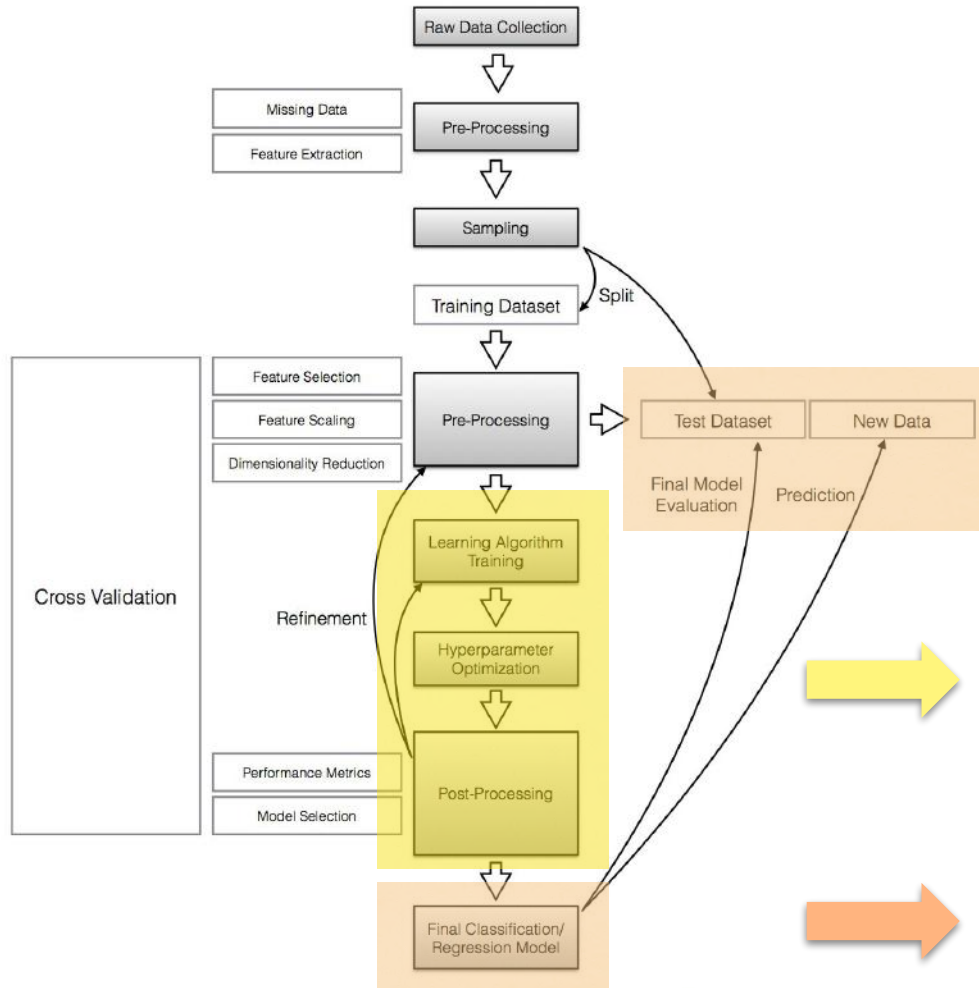
Supervised Learning

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

- A **training set** of examples with the **correct responses** (targets) is provided and, based on this training set,
- the **algorithm generalises to respond correctly to all possible inputs**.
- Typically, each **example** is a pair consisting of
 - an **input object** (typically a vector) and
 - a desired **output value** (also called the supervisory signal)



Supervised Learning



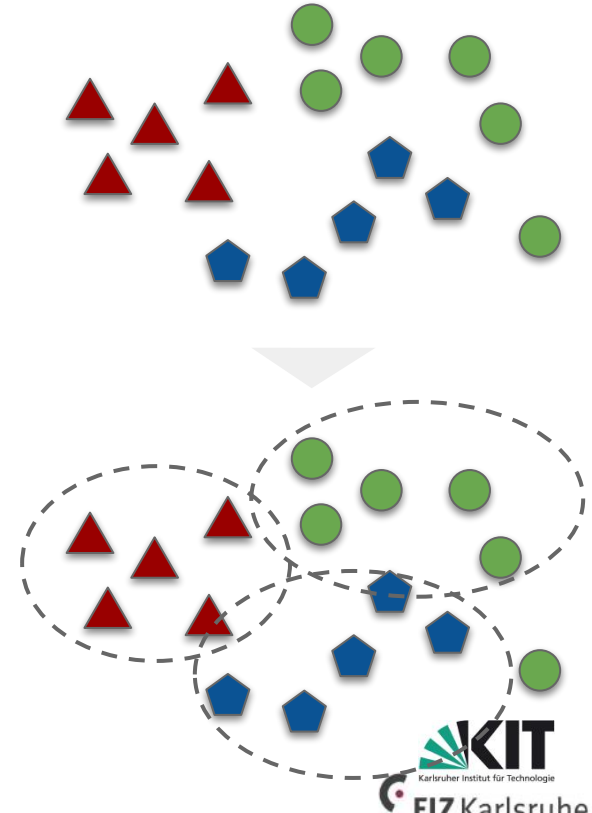
Training/Optimisation Phase

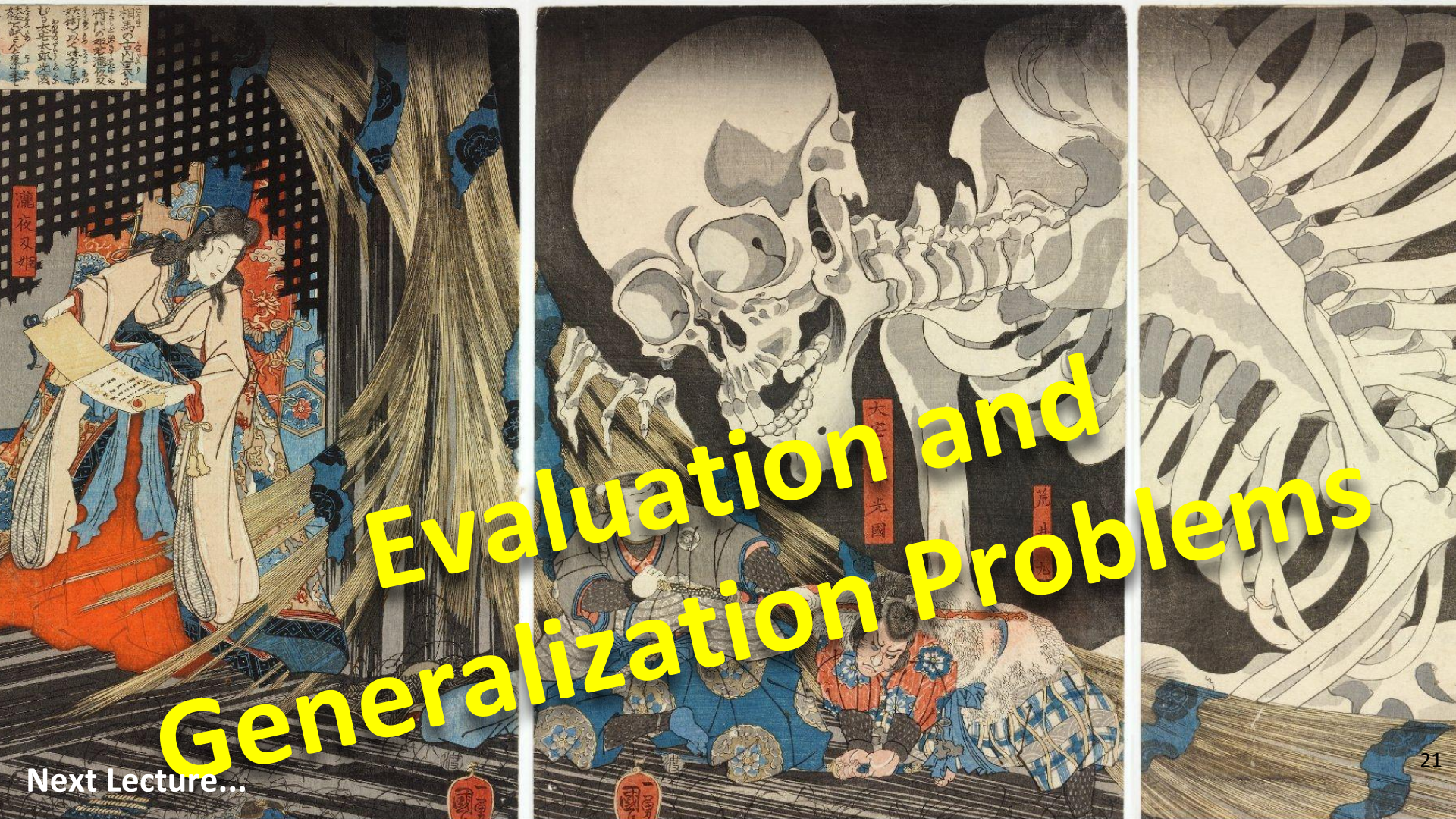
Prediction Phase

Unsupervised Learning

Information Service Engineering - Lecture 5 Basic Machine Learning - 5.2 Introduction to Machine Learning

- Inferring a function to describe hidden structure from "unlabeled" data
- **Correct responses are not provided**, but instead
- the algorithm tries to **identify similarities between the inputs** so that inputs that have something in common are categorised together.
- The statistical approach to unsupervised learning is known as **density estimation**.





Evaluation and Generalization Problems

Next Lecture...

Picture References:

- P.3: Jean Marc Cote (if 1901) or Villemard (if 1910), France in 2000 year (XXI century). Future school. France, paper card.
https://commons.wikimedia.org/wiki/File:France_in_XXI_Century_School.jpg
- P.7: Electronics Board, via pixabay, <https://pixabay.com/en/electronics-board-1955152/>
- P.13: Card catalog room at the Frick Art Reference Library taken during Open House New York 2016,
https://commons.wikimedia.org/wiki/File:Frick_Art_Reference_Library_Card_Catalog.jpg
- P.21: Utagawa Kuniyoshi, The three tiles in Takiyasha the Witch and the Skeleton Spectre
https://commons.wikimedia.org/wiki/File:Takiyasha_the_Witch_and_the_Skeleton_Spectre.jpg