

Introduction to Machine Learning

University of Bremen

Informatica Feminale 2018

Georgia Olympia Brikis

06th - 10th August 2018

About Myself

- **Computer Science** and **Philosophy** in Munich and Warsaw
- **Machine Learning Research** @ Siemens Corporate Technology
- **Current project** : Autonomous detection of coniferous seedlings in UAV images of Canadian forests
- Swimming, Building Bikes, Learning
- German, English, Polish, Greek, French and Japanese

About the Course I

- Course language : *English*
- Classroom : MZH, 1460

- Time Table

• Mon, 6 th	11:00-12:30	14:00-16:30
• Tue, 7 th	09:00-12:30	14:00-16:30
• Wed, 8 th	09:00-10:30	-
• Thu, 9 th	09:00-12:30	14:00-16:30
• Fri, 10 th	09:00-12:30	14:00-16:30

About the Course II

- Teaching method
 - 10-30 min introductory presentation
 - 15-30 min individual exercises
 - 15 min quizzes

Certificate / Credit

- *Participation Certificate (Teilnahmenachweis)*
attendance, participation in class
- *Participation Certificate + Credit (1 ECTS) (Leistungsnachweis)*
attendance, participation in class, successful completion of quizzes (>75%)
- *Participation Certificate + Credit (2 ECTS) (Leistungsnachweis)*
attendance, participation in class, successful completion of quizzes (>75%),
successful completion of homework assignment
- *General Participation Certificate*

Social Program

- Daily (afternoon) activities
- **Sign up** at the conference office

<https://www.informatica-feminale.de/eng/social-events/>

- | | | | |
|-------------------------|-------|-------------------|----------------------------------------------------|
| • Mon, 6 th | 16:30 | <i>Cafeteria</i> | Opening-Get together with Cake Buffet |
| • Tue, 7 th | 16:30 | <i>Field trip</i> | Airbus |
| • Wed, 8 th | 11:00 | <i>MZH 1470</i> | Talks, Lunch Buffet & Field trips to Research Labs |
| • Thu, 9 th | 16:30 | <i>Field trip</i> | Bremen Insitute for Production and Logistics |
| • Fri, 10 th | 16:30 | <i>Field trip</i> | Bremen Cotton Exchange |
| • Sat, 11 th | 18:00 | <i>Teerhof</i> | Networking Party |

Let's get to know each other



Key words you associate with Machine Learning
and/or what do you expect to learn about in this course



Python programming experience (e.g. project size, years of
experience)



What skills do you have that could be usefull for doing Machine
Learning? (e.g. Linear Algebra, Programming)

About the Course II

- Teaching method
 - 10-30 min introductory presentation
 - 15-30 min individual exercises
 - 15 min quizzes

Certificate / Credit

- *Participation Certificate (Teilnahmenachweis)*
attendance, participation in class
- *Participation Certificate + Credit (1 ECTS) (Leistungsnachweis)*
attendance, participation in class, successful completion of quizzes (>75%)
- *Participation Certificate + Credit (2 ECTS) (Leistungsnachweis)*
attendance, participation in class, successful completion of quizzes (>75%),
successful completion of homework assignment
- *General Participation Certificate*

INTRODUCTION

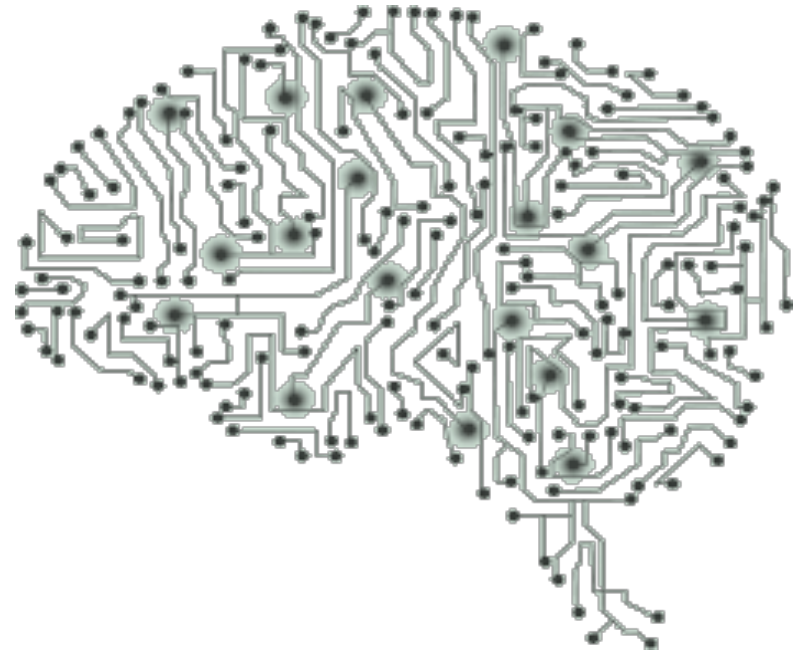
Machine Learning

Supervised learning

Unsupervised learning

Syllables

- I. Introduction
 - What is machine learning ?
 - Learning Paradigms
 - Elements of a learning task
- II. Linear Regression
- III. Logistic Regression
- IV. Decision Trees
- V. Clustering
- VI. Neural Networks
- VII. Model selection



What is machine learning ?

Machine Learning is a field of study that gives computers the **ability to learn without being explicitly programmed.**

Arthur Samuel (1959)

What is learning ?

- Progressively improve performance on a specific task
- Learning in other Sciences

What is machine learning ?

Machine Learning is a field of study that gives computers the **ability to learn without being explicitly programmed.**

Arthur Samuel (1959)

Well-posed Learning Problem: A computer program is said to learn from **experience E** with respect to some **task T** and some **performance measure P**, if its performance on T, as measured by P, improves with experience E.

Tom Mitchell (1998)

An example

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

Task

Classifying emails as spam or not spam

Experience

Watching you label emails as spam or not spam

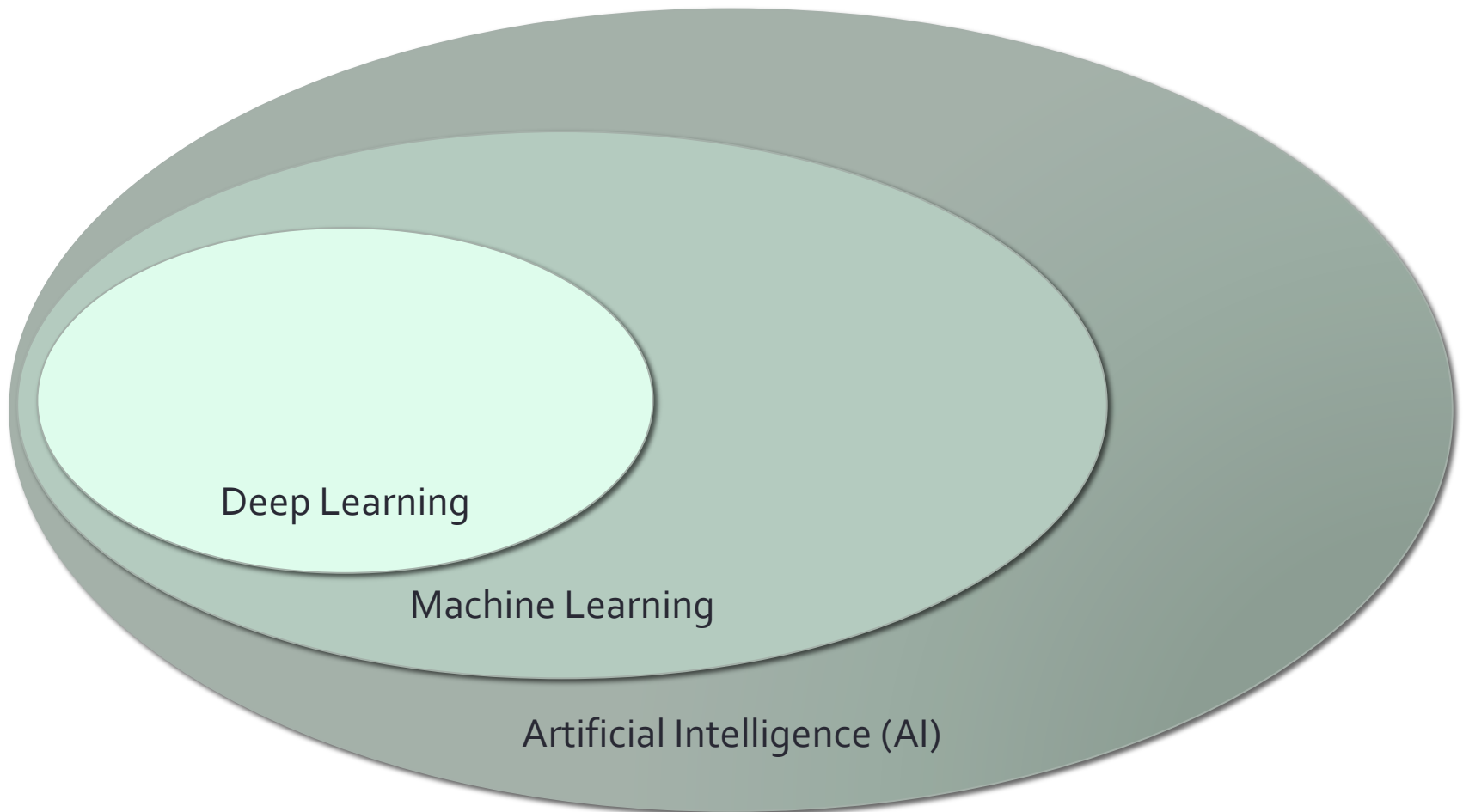
Performance Measure

The number (or fraction) of emails correctly classified as spam/not spam

Elements of a learning task

- Collection of **training data**
- A class of **learning models**. Often defined by the free model parameters in a learning model with a fixed structure
- Selection of a **cost function** which is a function of the data and the free parameters; a good model has a low cost
- Optimizing the cost function via a **learning rule** to find the best model in the class of learning models under consideration.

Terminology



Learning paradigms

Example: Digit Classification (I)

- *Goal*: Classification of printed or handwritten digits
- *Application*: automatic reading of postal codes
- A task in the field of optical character recognition

Example: Digit Classification (II)

- Goal:
- ...

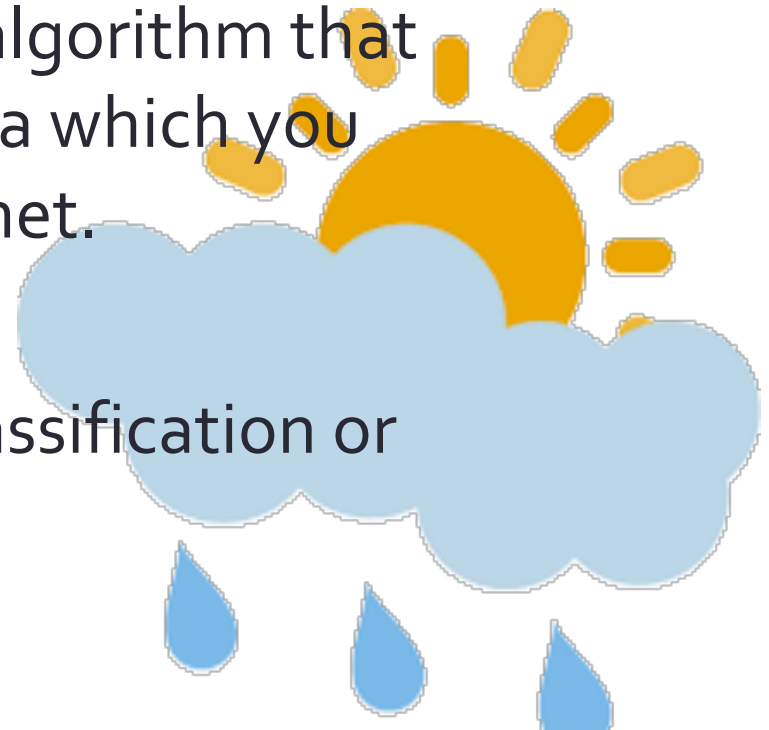
Unsupervised learning

• ... ■

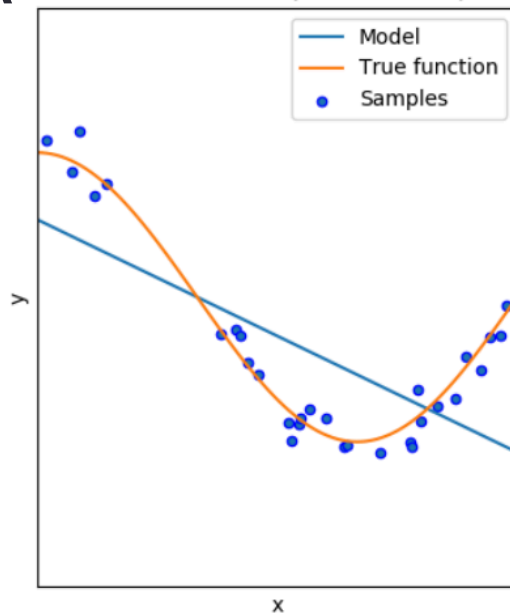
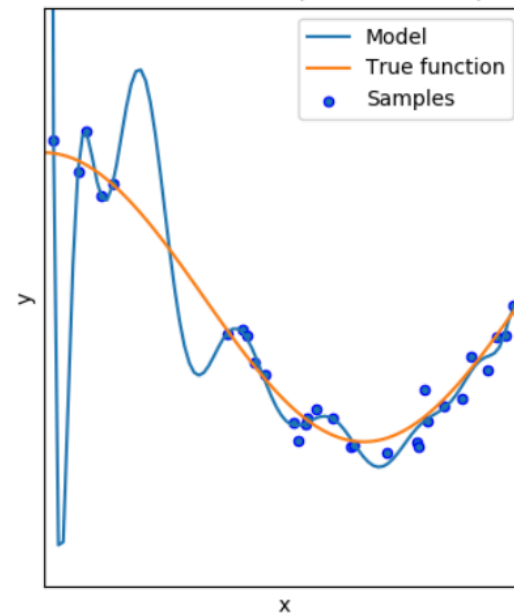
Imagine you want to predict whether or not it is going to be raining on your birthday.

You want to use a learning algorithm that utilizes historic weather data which you downloaded from the internet.

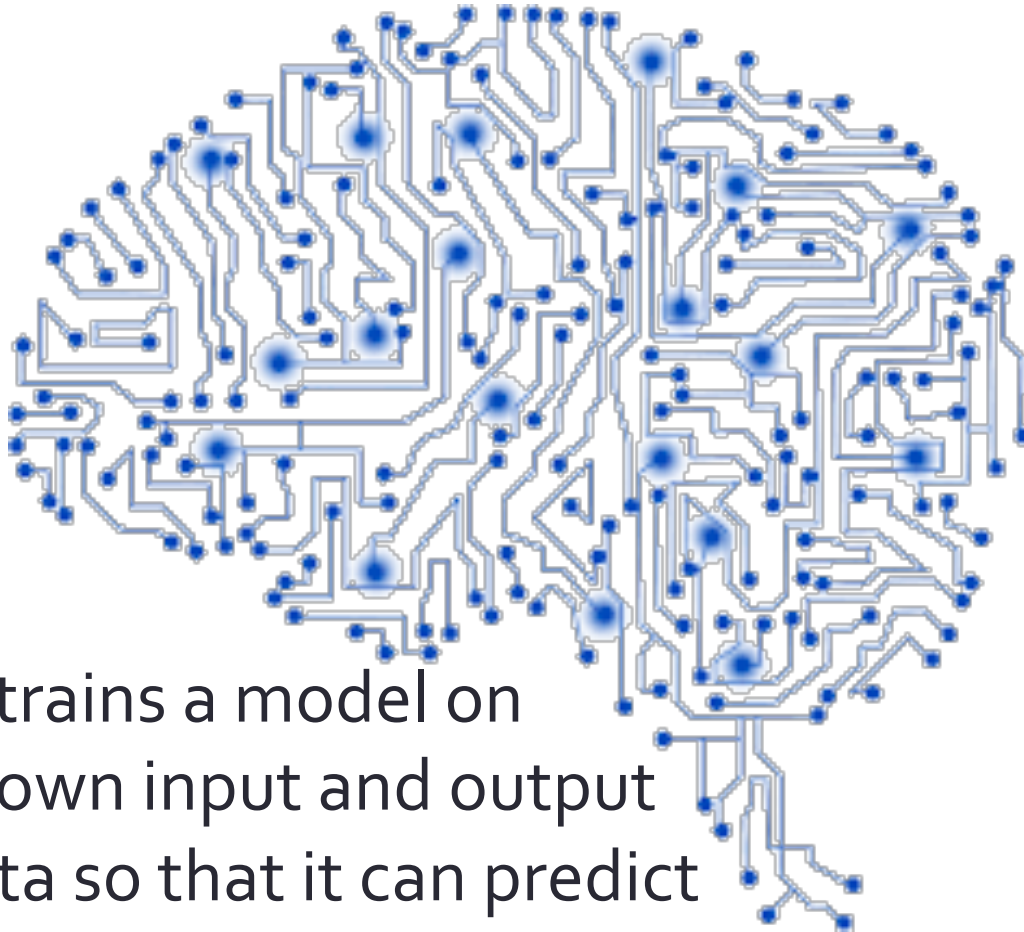
Would you treat this as a classification or regression problem ?



These are examples of ... ?

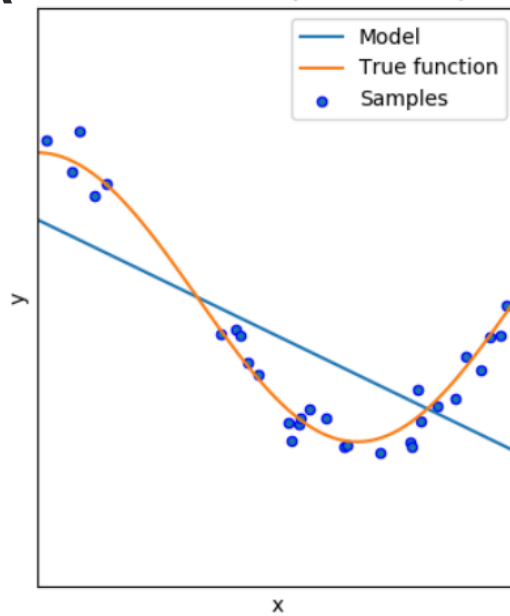
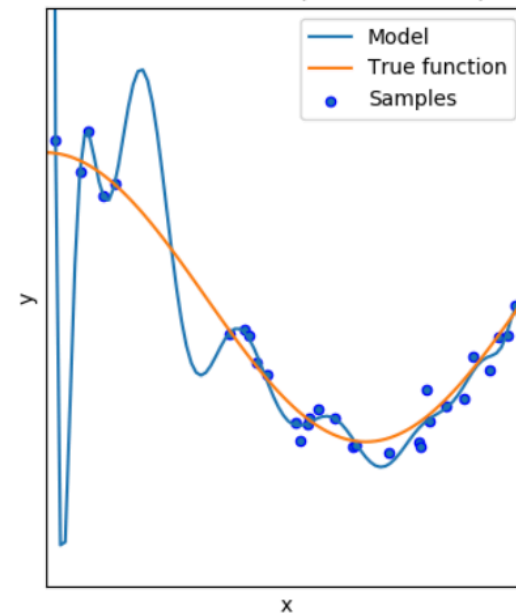
A**B**

Which of these is a reasonable definition of machine learning ?



... trains a model on known input and output data so that it can predict future outputs ?

These are examples of ... ?

A**B**

RESOURCES

Python, numpy, pandas

Jupyter Notebooks

Kaggle, Github