TECHNISCHE UNIVERSITÄT DARMSTADT

# Statistical Machine Learning

## Lecture 01: Introduction

**Kristian Kersting**
**TU Darmstadt**

Summer Semester 2020

# Today's Objectives

- Organizational issues
- Advertisement
- Introduction

# Outline

1. Organizational Issues

2. Introduction

3. Wrap-Up

TECHNISCHE UNIVERSITÄT DARMSTADT

# Outline

TECHNISCHE UNIVERSITÄT DARMSTADT

# Instructors

- Kristian Kersting heads the AI and ML Lab at the Department of Computer Science at the TU Darmstadt. He has studied computer science and your can find him in the Alte Hauptgebäude, Room 074, Hochschulstrasse 1. You can also contact Kristian through kersting@cs.tu-darmstadt.de

- **Karl Stelzner** joined the AIML Lab as a Phd student in 2017. He is working on probabilistic (deep) learning, in particular for unsupervised image understanding. You can contact Karl via email stelzner@cs.tu-darmstadt.de.

PLEASE FEEL FREE TO EMAIL US WITH QUESTIONS!

TECHNISCHE UNIVERSITÄT DARMSTADT

# Instructors

- **Svenja Stark** joined the IAS Lab as a Phd student in December 2016. She is working on adaptive skill libraries and skill comparison. You can contact her via email svenja@robot-learning.de

- **Hany Abdulsamad** joined the Intelligent Autonomous System lab in April 2016 as a PhD student. His research interests include optimal control, trajectory optimization, reinforcement learning and robotics. During his Phd, Hany is working on the SKILLS4ROBOTS project with the aim of enabling humanoid robots to acquire and improve a rich set of motor skills. You can contact him by email at hany@robot-learning.de

PLEASE FEEL FREE TO EMAIL US WITH QUESTIONS!

# Website & Mailing list

- Moodle: https://moodle.informatik.tu-darmstadt.de/course/view.php?id=928

# Course Language

...will be in **English**

# Course Language

...will be in **English**

**Why?**

- Essentially *all* machine learning literature is in English.
- Knowing the proper *terminology* is essential!
- Good to improve your English skills!

TECHNISCHE UNIVERSITÄT DARMSTADT

# Course Language

...will be in **English**

**Why?**

- Essentially *all* machine learning literature is in English.
- Knowing the proper *terminology* is essential!
- Good to improve your English skills!

Questions and answers in emails/homework/exams may be answered in German (However, this is not encouraged...).

# Interaction: Answers & Questions

When you answer or ask a complex question, your classmates may snigger at you!

However, your Professor rewards good answers & questions! 5 with a snickers

**snigger** = a smothered or half-suppressed laugh

# Feedback: Essential for both sides...


We appreciate
**FEEDBACK!**

Jeder Prof hat 'ne Meise. Meine dürfen Sie füttern!

# Exam & Bonus Points from Homework

**There will be a written exam.**

- Approximate date: The weeks after the end of classes...

# Exam & Bonus Points from Homework

**There will be a written exam.**

- Approximate date: The weeks after the end of classes...

**Homework Exercises:**

- Homework is crucial for the exam!
- The bonus questions will count as bonus points to the lecture!
- Will max out on bonus points!
- Please register in Moodle with groups of 2 students.

# Exam & Bonus Points from Homework

**There will be a written exam.**

- Approximate date: The weeks after the end of classes...

**Homework Exercises:**

- Homework is crucial for the exam!
- The bonus questions will count as bonus points to the lecture!
- Will max out on bonus points!
- Please register in Moodle with groups of 2 students.

**Question**: Favorite Homework-Frequency? 4 homeworks

TECHNISCHE UNIVERSITÄT DARMSTADT

# Homework Assignments

- There will be **4** homework assignments!
- Each assignment will contain:
  - A few multiple choice questions
  - A few essay questions
  - Some programming exercises.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Background Reading

- We will add current papers & tutorials!
- Standard background reading:
  - **C.M. Bishop, Pattern Recognition and Machine Learning (2006), Springer**
  - **K.P. Murphy, Machine Learning: a Probabilistic Perspective (2012), MIT Press**
  - **S. Rogers, M. Girolami, A First Course in Machine Learning (2016), CRC Press**
- Mathematics for machine learning background:
  - **Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong, Mathematics for Machine Learning,** `https://mml-book.github.io/`

TECHNISCHE UNIVERSITÄT DARMSTADT

# Background Reading

- Other resources
  - D. Barber, Bayesian Reasoning and Machine Learning (2012), Cambridge University Press (http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/090310.pdf)
  - T. Hastie, R. Tibshirani, and J. Friedman (2015), The Elements of Statistical Learning, Springer Verlag (https://web.stanford.edu/~hastie/Papers/ESLII.pdf)
  - R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification (2nd ed. 2001), Willey- Interscience
  - T.M. Mitchell, Machine Learning (1997), McGraw-Hill
  - R. Sutton, A. Barto. Reinforcement Learning - an Introduction, MIT Press (http://incompleteideas.net/book/RLbook2018.pdf)

TECHNISCHE UNIVERSITÄT DARMSTADT

# How does it fit in your course plan? 1/3

**VL *Statistical Machine Learning*** is a good preparation for advanced lectures:

- **VL** *Lernende Robot* (aka *Robot Learning*)
- **VL** *Probababilistic Graphical Models*
- **VL** *Statistical Relational AI*
- **IP** *Robot Learning 1, 2*

TECHNISCHE UNIVERSITÄT DARMSTADT

# How does it fit in your course plan? 2/3

**Related Classes:**

- Improve Foundations: *Data Mining and Machine Learning* (WiSe), *Robot Learning* (WiSe), *Deep Learning: Architectures and Methods* (WiSe)
- Useful Techniques: *Optimierung statischer und dynamischer Systeme*
- Applications of learning: *Computer Vision*

**Theses**: We always have B.Sc. or M.Sc. Theses on ML topics.

TECHNISCHE UNIVERSITÄT DARMSTADT

# How does it fit in your course plan? 3/3

**B.Sc. / M.Sc. Informatik**:

- Human Computer Systems (see Modulhandbuch)
- If you are strongly interested in machine learning you should take:
  - Statistical Machine Learning for HCS credit
  - Data Mining and Machine Learning for DKE credit
  - Robot Learning for CE credit
  - Computer Vision for Visual Computing

**M.Sc. in Autonome Systeme**

**M.Sc. in Visual Computing**: Area “Computer Vision & ML”

TECHNISCHE UNIVERSITÄT DARMSTADT

# Outline

TECHNISCHE UNIVERSITÄT DARMSTADT

# Why Machine Learning?

- “We are drowning in information and starving for knowledge.” - John Naisbitt

# Why Machine Learning?

- “We are drowning in information and starving for knowledge.” - John Naisbitt

- **Era of big data:**
  - In 2017 there are about 1.8 trillion webpages on the internet
  - 20 hours of video are uploaded to YouTube every minute
  - Walmart handles more than 1M transactions per hour and has databases containing more than 2.5 petabytes ($2.5 \times 10^{15}$) of information.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Why Machine Learning?

- "We are drowning in information and starving for knowledge." - John Naisbitt

- **Era of big data:**
  - In 2017 there are about 1.8 trillion webpages on the internet
  - 20 hours of video are uploaded to YouTube every minute
  - Walmart handles more than 1M transactions per hour and has databases containing more than 2.5 petabytes ($2.5 \times 10^{15}$) of information.

**No human being can deal with the data avalanche!**

TECHNISCHE UNIVERSITÄT DARMSTADT

# Why Machine Learning?

“I keep saying the sexy job in the next ten years will be **statisticians** and **machine learners**. People think I’m joking, but who would’ve guessed that computer engineers would’ve been the sexy job of the 1990s? The ability to take data – to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it – that’s going to be a hugely important skill in the next decades.”

Hal Varian, Chief Economist at Google, 2009

# Job Perspective

"A significant constraint on realizing value from big data will be a shortage of talent, particularly of people with deep expertise in statistics and machine learning."

Big data: The next frontier for innovation, competition, and productivity, 2011, McKinsey Global Institute

# Machine Learning

What is ML? What is its goal?

# Machine Learning

What is ML? What is its goal?

- Develop a machine / an algorithm that learns to perform a task from past experience.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Machine Learning

What is ML? What is its goal?

- Develop a machine / an algorithm that learns to perform a task from past experience.

Why? What for?

TECHNISCHE UNIVERSITÄT DARMSTADT

# Machine Learning

What is ML? What is its goal?

- Develop a machine / an algorithm that learns to perform a task from past experience.

Why? What for?

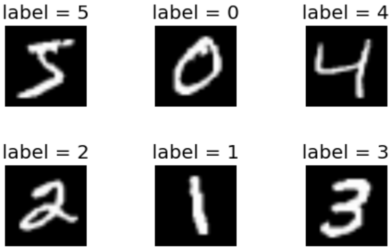
- Fundamental component of every intelligent and / or autonomous system
- Discovering "rules" and patterns in data
- Automatic adaptation of systems
- Attempting to understand human / biological learning

# Machine Learning in Action

# Machine Learning Examples

TECHNISCHE UNIVERSITÄT DARMSTADT

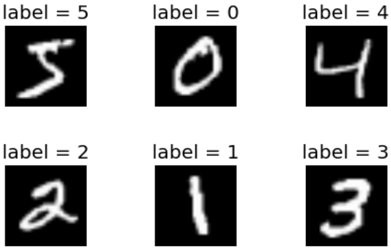
## Recognition of handwritten digits




- These digits are given to us as small digital images

# Machine Learning Examples

## Recognition of handwritten digits

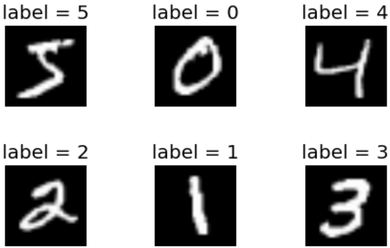



- These digits are given to us as small digital images
  - We have to build a “machine” to decide which digit it is

TECHNISCHE UNIVERSITÄT DARMSTADT

# Machine Learning Examples

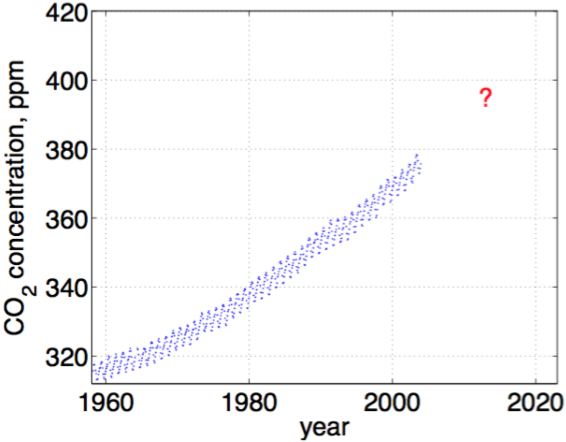
## Recognition of handwritten digits




- These digits are given to us as small digital images
  - We have to build a “machine” to decide which digit it is
  - **Obvious challenge**: There are many different ways in which people handwrite

# Machine Learning Examples

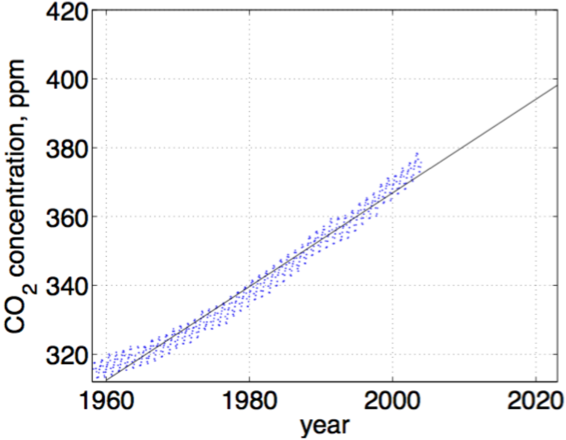
## CO2 prediction



TECHNISCHE UNIVERSITÄT DARMSTADT

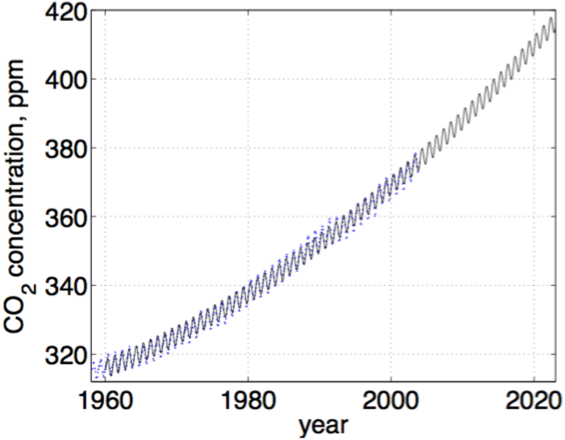
# Machine Learning Examples

## CO2 prediction



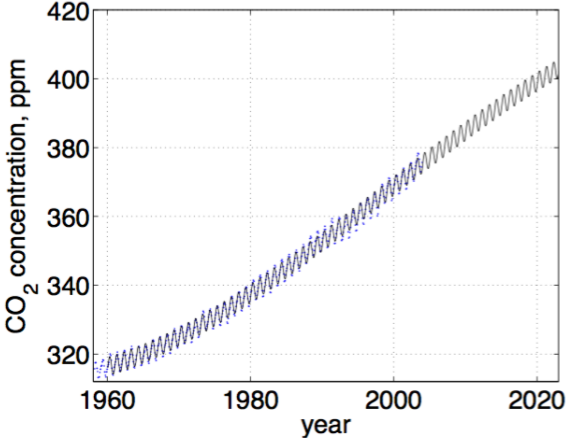
# Machine Learning Examples

## CO2 prediction



# Machine Learning Examples

## CO2 prediction



# Machine Learning Examples

- **Email filtering**

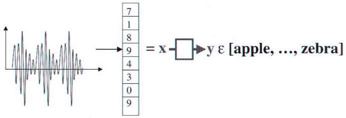
x ε [a-z]+ → y ε [important, spam]

TECHNISCHE UNIVERSITÄT DARMSTADT

# Machine Learning Examples

- **Email filtering**


- **Speech recognition**

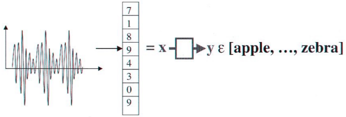


# Machine Learning Examples

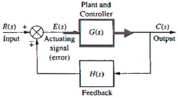
- **Email filtering**

  x ε [a-z]+ → y ε [important, spam]

- **Speech recognition**

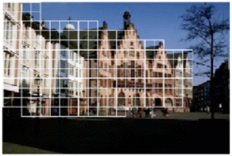


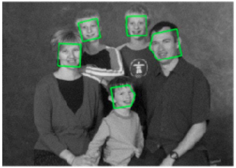

- **Vehicle control**



# Machine Learning Impact & Successes

- Recognition of speech, letters, faces, ...
- Autonomous vehicle navigation
- Games
  - Backgammon world-champion
  - Chess: Deep-Blue vs. Kasparov
  - Go: AlphaGo, AlphaGo Zero
- Google
- Finding new astronomical structures
- Fraud detection (credit card applications)
- ...





# Machine Learning



- Develop a machine / an algorithm that learns to perform a task from past experience.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Machine Learning

- Develop a machine / an algorithm that learns to perform a task from past experience.

- Put more abstractly:

# Machine Learning

- Develop a machine / an algorithm that learns to perform a task from past experience.

- Put more abstractly:
  - Our task is to learn a mapping from input to output.

$$f : I \rightarrow O$$

# Machine Learning

- Develop a machine / an algorithm that learns to perform a task from past experience.
- Put more abstractly:
  - Our task is to learn a mapping from input to output.

$$f : I \to O$$

  - Put differently, we want to predict the output from the input.

$$y = f(x; \theta)$$

# Machine Learning

- Develop a machine / an algorithm that learns to perform a task from past experience.
- Put more abstractly:
  - Our task is to learn a mapping from input to output.

$$f : I \to O$$

  - Put differently, we want to predict the output from the input.

$$y = f(x; \theta)$$

  - Input: $x \in I$ (images, text, sensor measurements, ...)
  - Output: $y \in O$
  - Parameters: $\theta \in \Theta$ (what needs to be "learned")

# Classification vs Regression

## Classification

- Learn a mapping into a discrete space, e.g.

# Classification vs Regression

## Classification

- Learn a mapping into a discrete space, e.g.
  - $O = \{0, 1\}$
  - $O = \{0, 1, 2, 3, \ldots\}$
  - $O = \{\text{verb}, \text{noun}, \text{adjective}, \ldots\}$
- Examples:

# Classification vs Regression

## Classification

- Learn a mapping into a discrete space, e.g.
  - $O = \{0, 1\}$
  - $O = \{0, 1, 2, 3, \ldots\}$
  - $O = \{\text{verb}, \text{noun}, \text{adjective}, \ldots\}$
- Examples:
  - Spam / not spam
  - Digit recognition
  - Part of Speech tagging

# Classification vs Regression

**Regression**

- Learn a mapping into a continuous space, e.g.

# Classification vs Regression

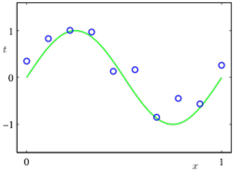
**Regression**

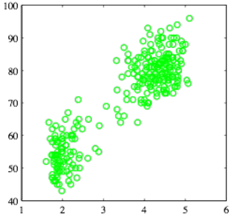
- Learn a mapping into a continuous space, e.g.
  - $O = \mathbb{R}$
  - $O = \mathbb{R}^3$
- Examples

# Classification vs Regression

## Regression

- Learn a mapping into a continuous space, e.g.
  - $O = \mathbb{R}$
  - $O = \mathbb{R}^3$
- Examples
  - Curve fitting, Financial Analysis, Housing prices, ...

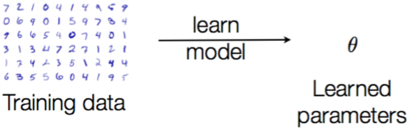


TECHNISCHE UNIVERSITÄT DARMSTADT

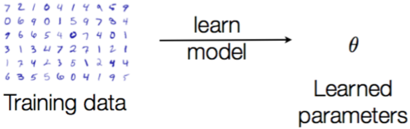
# General Paradigm

## Training

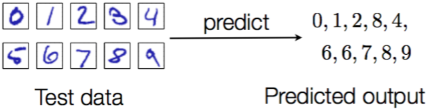


# General Paradigm

## Training

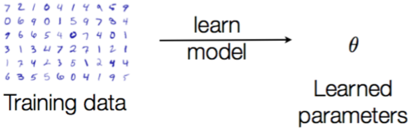



## Testing

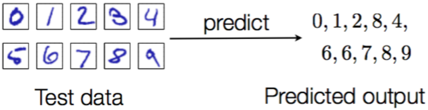


# General Paradigm

## Training




## Testing




The test dataset needs to be different than the training dataset!

But ideally from the same underlying distribution.

# What data do we have for training?

- Data with labels (input / output pairs): supervised learning
  - Image with digit label
  - Sensory data for car with intended steering control

TECHNISCHE UNIVERSITÄT DARMSTADT

# What data do we have for training?

- Data with labels (input / output pairs): supervised learning
  - Image with digit label
  - Sensory data for car with intended steering control
- Data without labels: unsupervised learning
  - Automatic clustering (grouping) of sounds
  - Clustering of text according to topics
  - Density Estimation
  - Dimensionality Reduction

# What data do we have for training?

- Data with labels (input / output pairs): supervised learning
  - Image with digit label
  - Sensory data for car with intended steering control
- Data without labels: unsupervised learning
  - Automatic clustering (grouping) of sounds
  - Clustering of text according to topics
  - Density Estimation
  - Dimensionality Reduction
- Data with and without labels: semi-supervised learning

TECHNISCHE UNIVERSITÄT DARMSTADT

# What data do we have for training?

- Data with labels (input / output pairs): supervised learning
  - Image with digit label
  - Sensory data for car with intended steering control
- Data without labels: unsupervised learning
  - Automatic clustering (grouping) of sounds
  - Clustering of text according to topics
  - Density Estimation
  - Dimensionality Reduction
- Data with and without labels: semi-supervised learning
- No examples: learn-by-doing
  - Reinforcement Learning

# Some Key Challenges

- We need generalization!
  - We cannot simply memorize the training set.
- What if we see an input that we haven't seen before?
  - Different shape of the digit image (unknown writer)
  - "Dirt" on the picture, etc.
  - We need to learn what is important for carrying out our task.
- This is one of the most crucial points that we will return to many times.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Generalization

## How do we achieve generalization?

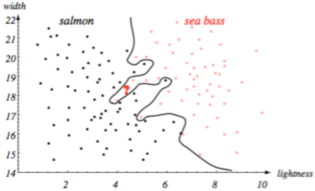



**FIGURE 1.5.** Overly complex models for the fish will lead to decision boundaries that are complicated. While such a decision may lead to perfect classification of our training samples, it would lead to poor performance on future patterns. The novel test point marked **?** is evidently most likely a salmon, whereas the complex decision boundary shown leads it to be classified as a sea bass. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Generalization

## How do we achieve generalization?

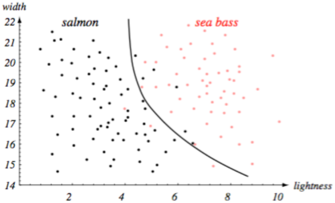



**FIGURE 1.6.** The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Occam's Razor

# Generalization

## How do we achieve generalization?

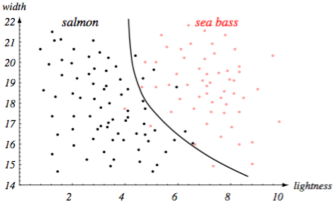



**FIGURE 1.6.** The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Occam's Razor

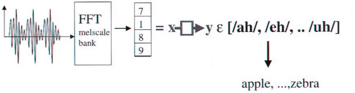
We should not make the model overly complex!

# Prominent example of overfitting...

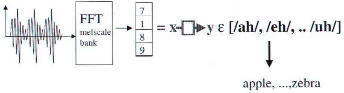
# Some Key Challenges

- Input:



# Some Key Challenges

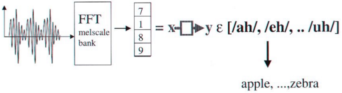
- Input:




- Features

# Some Key Challenges

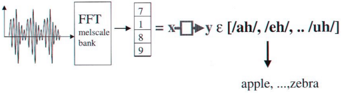
- Input:




- Features
  - Choosing the “right” features is very important.
  - Coding and use of domain knowledge.
  - May allow for invariance (e.g., volume and pitch of voice).

# Some Key Challenges

- Input:




- Features
  - Choosing the “right” features is very important.
  - Coding and use of domain knowledge.
  - May allow for invariance (e.g., volume and pitch of voice).
- Curse of Dimensionality:
  - If the features are too high-dimensional, we will run into trouble
  - Dimensionality reduction.

# Some Key Challenges

- How do we measure performance?
  - 99% correct classification in speech recognition: What does that really mean?
  - We understand the meaning of the sentence? We understand every word? For all speakers?

TECHNISCHE UNIVERSITÄT DARMSTADT

# Some Key Challenges

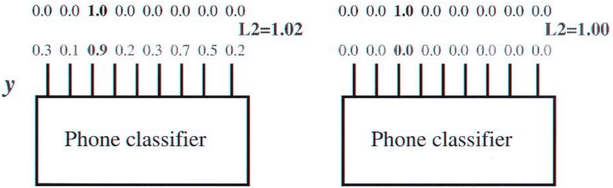
- How do we measure performance?
  - 99% correct classification in speech recognition: What does that really mean?
  - We understand the meaning of the sentence? We understand every word? For all speakers?
- Need more concrete numbers:
  - % of correctly classified letters
  - average distance driven (until accident...)
  - % of games won
  - % correctly recognized words, sentences, etc.

# Some Key Challenges

- How do we measure performance?
  - 99% correct classification in speech recognition: What does that really mean?
  - We understand the meaning of the sentence? We understand every word? For all speakers?
- Need more concrete numbers:
  - % of correctly classified letters
  - average distance driven (until accident...)
  - % of games won
  - % correctly recognized words, sentences, etc.
- Training vs. testing performance!

# Some Key Challenges

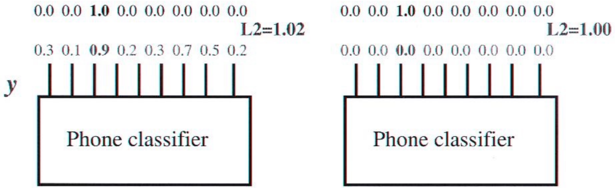
- We also need to define the right error metric:




- Which is better?

# Some Key Challenges

- We also need to define the right error metric:




- Which is better?
- Euclidean distance (L2 norm) might be useless.

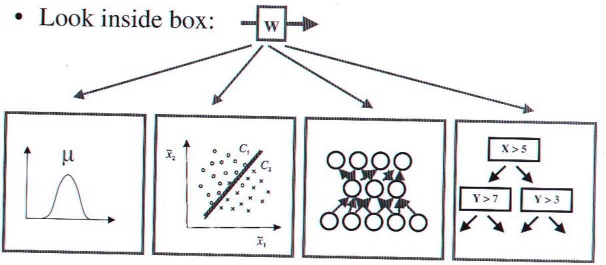
# Some Key Challenges

Which is the right model?

# Some Key Challenges

Which is the right model?

- The learned parameters (**w**) can mean a lot of different things:
  - May characterize the family of functions or the model space
  - May index the hypothesis space
  - **w** can be a vector, adjacency matrix, graph, ...



TECHNISCHE UNIVERSITÄT DARMSTADT

# Some Key Challenges

Even if we have solved the other problems, computation is usually quite hard:

# Some Key Challenges

Even if we have solved the other problems, computation is usually quite hard:

- Learning often involves some kind of optimization

# Some Key Challenges

Even if we have solved the other problems, computation is usually quite hard:

- Learning often involves some kind of optimization
- Find (search) best model parameters

# Some Key Challenges

Even if we have solved the other problems, computation is usually quite hard:

- Learning often involves some kind of optimization
- Find (search) best model parameters
- Often we have to deal with thousands, millions, billions, ..., of training examples

TECHNISCHE UNIVERSITÄT DARMSTADT

# Some Key Challenges

Even if we have solved the other problems, computation is usually quite hard:

- Learning often involves some kind of optimization
- Find (search) best model parameters
- Often we have to deal with thousands, millions, billions, ..., of training examples
- Given a model, compute the prediction efficiently

# Why is machine learning interesting (for you)?

- Machine learning is a challenging problem that is far from being solved.

# Why is machine learning interesting (for you)?

- Machine learning is a challenging problem that is far from being solved.
  - Our learning systems are primitive compared to us humans.
  - Think about what and how quickly a child can learn!

# Why is machine learning interesting (for you)?

- Machine learning is a challenging problem that is far from being solved.
  - Our learning systems are primitive compared to us humans.
  - Think about what and how quickly a child can learn!
- It combines insights and tools from many fields and disciplines:
  - Traditional artificial intelligence (logic, semantic networks, ...)
  - Statistics
  - Complexity theory
  - Artificial neural networks
  - Psychology
  - Adaptive control
  - ...

# Why is machine learning interesting (for you)?

- Allows you to apply theoretical skills that you may otherwise only use rarely.

TECHNISCHE UNIVERSITÄT DARMSTADT

# Why is machine learning interesting (for you)?

- Allows you to apply theoretical skills that you may otherwise only use rarely.
- Has lots of applications:
  - Computer vision
  - Computer linguistics
  - Search (think Google)
  - Digital “assistants”
  - Computer systems
  - Robotics
  - ...

TECHNISCHE UNIVERSITÄT DARMSTADT

# Why is machine learning interesting (for you)?

- It is a growing field:
  - Many major companies are hiring people with machine learning knowledge.
  - Learning machine learning is probably the most promising route to such a 80-160.000 Euro Job...
  - Lampert: “Most Computer Vision is just machine learning applied to pictures...”

# Why is machine learning interesting (for you)?

- It is a growing field:
  - Many major companies are hiring people with machine learning knowledge.
  - Learning machine learning is probably the most promising route to such a 80-160.000 Euro Job...
  - Lampert: “Most Computer Vision is just machine learning applied to pictures...”
- It is beating traditional hand-engineered methods in many tasks (e.g., Vision, Natural Language, ...)

# Why is machine learning interesting (for you)?

- It is a growing field:
  - Many major companies are hiring people with machine learning knowledge.
  - Learning machine learning is probably the most promising route to such a 80-160.000 Euro Job...
  - Lampert: "Most Computer Vision is just machine learning applied to pictures..."
- It is beating traditional hand-engineered methods in many tasks (e.g., Vision, Natural Language, ...)
- Because it is fun!

# Preliminary Syllabus (Subject to change!)

# Preliminary Syllabus (Subject to change!)

- Refresher of Statistics, Linear Algebra & Optimization (~ 2 Weeks)

# Preliminary Syllabus (Subject to change!)

- Refresher of Statistics, Linear Algebra & Optimization (~ 2 Weeks)
- Fundamentals (~ 3 weeks)
  - Bayes decision theory, maximum likelihood, Bayesian inference
  - Performance evaluation
  - Probability density estimation
  - Mixture models, expectation maximization

TECHNISCHE UNIVERSITÄT DARMSTADT

# Preliminary Syllabus (Subject to change!)

- Refresher of Statistics, Linear Algebra & Optimization (~ 2 Weeks)
- Fundamentals (~ 3 weeks)
  - Bayes decision theory, maximum likelihood, Bayesian inference
  - Performance evaluation
  - Probability density estimation
  - Mixture models, expectation maximization
- Linear Methods (~ 3-4 weeks)
  - Linear regression
  - PCA, robust PCA
  - Fisher linear discriminant
  - Generalized linear models

# Preliminary Syllabus

- Large-Margin Methods (~ 3-4 weeks)
  - Statistical learning theory
  - Support vector machines
  - Kernel methods

TECHNISCHE UNIVERSITÄT DARMSTADT

# Preliminary Syllabus

- Large-Margin Methods (~ 3-4 weeks)
  - Statistical learning theory
  - Support vector machines
  - Kernel methods
- Neural Networks (~ 3 weeks)
  - Neural Networks: From Inspiration to Application
  - Deep Learning: What is really different?

# Preliminary Syllabus

- Large-Margin Methods (~ 3-4 weeks)
    - Statistical learning theory
    - Support vector machines
    - Kernel methods
- Neural Networks (~ 3 weeks)
    - Neural Networks: From Inspiration to Application
    - Deep Learning: What is really different?
- Miscellaneous (~ 3 weeks)
    - Model averaging (bagging & boosting)
    - Graphical models (basic introduction)

# Credits

- These slides are essentially the slides of Jan Peters.
- Some parts of Jan's lecture material have been developed by Profs. Bernt Schiele, Stefan Roth and Stefan Schaal for the previous iterations of this course or similar classes.
- Many figures that I will use are directly taken out of the books by Chris Bishop and Duda, Hart & Stork and Kevin Murphy.

# Outline

TECHNISCHE UNIVERSITÄT DARMSTADT

# 3. Wrap-Up

You know now:

TECHNISCHE UNIVERSITÄT DARMSTADT

# 3. Wrap-Up

You know now:

- What Machine Learning is and what it is not.

# 3. Wrap-Up

You know now:

- What Machine Learning is and what it is not.
- Some of Machine Learning applications.

# 3. Wrap-Up

You know now:

- What Machine Learning is and what it is not.
- Some of Machine Learning applications.
- The different types of learning problems.

TECHNISCHE UNIVERSITÄT DARMSTADT

# 3. Wrap-Up

You know now:

- What Machine Learning is and what it is not.
- Some of Machine Learning applications.
- The different types of learning problems.
- What classification and regression are.

# 3. Wrap-Up

You know now:

- What Machine Learning is and what it is not.
- Some of Machine Learning applications.
- The different types of learning problems.
- What classification and regression are.
- The challenges in solving a problem with Machine Learning.

# Self-Test Questions

- What are some of Machine Learning applications?

# Self-Test Questions

- What are some of Machine Learning applications?
- When can we benefit from using Machine Learning methods?

# Self-Test Questions

- What are some of Machine Learning applications?
- When can we benefit from using Machine Learning methods?
- What are the different types of learning?

TECHNISCHE UNIVERSITÄT DARMSTADT

# Self-Test Questions

- What are some of Machine Learning applications?
- When can we benefit from using Machine Learning methods?
- What are the different types of learning?
- What is the difference between classification and regression? Can you give some examples of both tasks (and identify the domain and codomain)?

# Self-Test Questions

TECHNISCHE UNIVERSITÄT DARMSTADT

- What are some of Machine Learning applications?
- When can we benefit from using Machine Learning methods?
- What are the different types of learning?
- What is the difference between classification and regression? Can you give some examples of both tasks (and identify the domain and codomain)?
- What are the challenges when solving a Machine Learning problem?

# Self-Test Questions

TECHNISCHE UNIVERSITÄT DARMSTADT

- What are some of Machine Learning applications?
- When can we benefit from using Machine Learning methods?
- What are the different types of learning?
- What is the difference between classification and regression? Can you give some examples of both tasks (and identify the domain and codomain)?
- What are the challenges when solving a Machine Learning problem?
- What is generalization? What is overfitting?

# Homework

- Select some Machine Learning applications and check:
    - What type of learning is it?
    - Is it a classification or regression problem?
    - What challenges do you foresee when solving this problem using Machine Learning methods?

TECHNISCHE UNIVERSITÄT DARMSTADT

# Homework

- Select some Machine Learning applications and check:
  - What type of learning is it?
  - Is it a classification or regression problem?
  - What challenges do you foresee when solving this problem using Machine Learning methods?
- Reading assignment
  - Jordan Book, Linear Algebra chapter (online)
  - Pedro Domingos, *A few useful things to know about Machine Learning* (https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf)
  - Bishop ch. 1