Statistical Machine Learning

Lecture 1: Organization & Introduction

Simone Schaub-Meyer & Marcus Rohrbach
Department of Computer Science
TU Darmstadt

Summer Term 2025

Outline

1. Organizational Aspects

2. Introduction to ML

3. Wrap-Up

Instructors

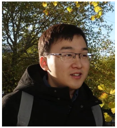
- **Simone Schaub-Meyer** is an assistant professor at TU Darmstadt leading the Image & Video Group and is affiliated with the Hessian Center for AI. Her research focuses on efficient, robust, and interpretable methods. She earned her Ph.D. from ETH Zurich in collaboration with Disney Research Zurich.

- **Marcus Rohrbach** is a professor at TU Darmstadt. He leads the Multimodal Reliable AI lab. His research focuses on computer vision, machine learning, and computational linguistics. Previously, he worked at Meta AI and UC Berkeley. He earned his Ph.D. from the Max Planck Institute for Informatics.

Teaching Assistants

- **Tobias Braun** is a PhD student in the Multimodal AI Lab at TU Darmstadt, working on multimodal reliable AI. You can contact Tobias at `tobias.braun@mai.tu-darmstadt.de`.

- **Jiayun Li** is a PhD student in the Interactive Robot Perception & Learning lab at TU Darmstadt, working on motion planning. You can contact Jiayun at `jiayun.li@tu-darmstadt.de`.



For Q&As, please use Moodle.
For personal matters, contact us directly at `smlstaff@visinf.tu-darmstadt.de`.

Website & Discussion Board

- Moodle:
  https://moodle.tu-darmstadt.de/course/view.php?id=42016
  - Lecture slides
  - Pointers to readings
  - Homework assignments + hand in
- Discussion board on Moodle
  - Please use it to ask questions of public interest.
  - You are encouraged to discuss with each other.
  - However: Please do not share solutions or give strong hints about the solutions to the homework problems.
  - Asking questions via the discussion boards on Moodle is the preferred form of communication with the course staff!

TECHNISCHE UNIVERSITÄT DARMSTADT

Course Language

...will be in **English**

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.

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.

Organization

- Lecture (13:25 – 15:05)
  - One lecture per week (exceptions announced on Moodle)
  - No lecture on the 4th of June (TU meet and move)
  - We will cover the foundational aspects of each topic
  - We might record and upload the lectures in the Moodle

TECHNISCHE UNIVERSITÄT DARMSTADT

Organization

- Lecture (13:25 – 15:05)
  - One lecture per week (exceptions announced on Moodle)
  - No lecture on the 4th of June (TU meet and move)
  - We will cover the foundational aspects of each topic
  - We might record and upload the lectures in the Moodle
- Exercises, homework, etc.
  - Exercise session before the lecture (12:35–13:20)
  - We will cover some practical aspects, show general problem-solving and discuss the homework assignments
  - Time for Q&A
  - Mostly: Time to work on your homework assignments

TECHNISCHE UNIVERSITÄT DARMSTADT

Exam & Bonus Points from Homework

There will be a written exam.

- The exam will be in English.
- Date: Tuesday, 2nd of September (exact information, see TUCaN)

Exam & Bonus Points from Homework

There will be a written exam.

- The exam will be in English.
- Date: Tuesday, 2nd of September (exact information, see TUCaN)

Homework Exercises:

- Homework exercises are crucial for the exam!
- The points from the homework will count as bonus toward the final grade.
- Up to 1.0 higher grade (maxing out the departmental rules)
- Until next week: Register in Moodle with groups of 3 students.

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

Homework Assignments

- There will be **4** homework assignments.
- Each assignment will contain:
  - A few multiple choice questions
  - A few essay questions
  - Some programming exercises
- Homework assignment 1 will be released on 30.04.2025

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
- Homework assignment 1 will be released on 30.04.2025
- 2 homework assignments will be checked in person
  - Groups randomly selected
  - Probably Tuesdays, more information in the exercise session

TECHNISCHE UNIVERSITÄT DARMSTADT

Background Reading

- We will add current papers & tutorials.
- Standard background reading:
  - ***C.M. Bishop, Pattern Recognition and Machine Learning (2006), Springer,*** https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf
  - ***K.P. Murphy, Probabilistic Machine Learning: An Introduction (2022), MIT Press,*** https://probml.github.io/pml-book/book1.html
  - ***S. Rogers, M. Girolami, A First Course in Machine Learning (2016), CRC Press***
- Mathematics background for machine learning:
  - ***M.P. Deisenroth, A. Aldo Faisal, and C.S. Ong, Mathematics for Machine Learning (2020), Cambridge University Press*** 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?

VL *Statistical Machine Learning* prepares you for advanced courses:

- **VL** *Reinforcement Learning*
- **VL** *Lernende Roboter* (aka *Robot Learning*)
- **IP** *Robot Learning*
- **VL** *Intelligente Robotermanipulation*
- **VL** *Probababilistic Graphical Models*
- **VL** *Computer Vision I and II*
- **VL** *Multimodal Artificial Intelligence*
- **VL** *Deep Learning for NLP*

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

Questions about course organization

Time for your questions regarding the course organization!

Outline

1. Organizational Aspects

2. Introduction to ML

3. Wrap-Up

Why Machine Learning?

TECHNISCHE UNIVERSITÄT DARMSTADT

- We generate much more data than we can analyze manually
  - "We are drowning in information and starving for knowledge."
    – John Naisbitt, author
  - 500 hours of video are uploaded to YouTube every minute (2022)
  - Intel estimates that a single self driving car will generate up to 4 TB of data per day

Why Machine Learning?

- We generate much more data than we can analyze manually
  - “We are drowning in information and starving for knowledge.” – John Naisbitt, author
  - 500 hours of video are uploaded to YouTube every minute (2022)
  - Intel estimates that a single self driving car will generate up to 4 TB of data per day

No human being can deal with this data avalanche!

TECHNISCHE UNIVERSITÄT DARMSTADT

Why Machine Learning?

- We generate much more data than we can analyze manually
  - "We are drowning in information and starving for knowledge."
    – John Naisbitt, author
  - 500 hours of video are uploaded to YouTube every minute (2022)
  - Intel estimates that a single self driving car will generate up to 4 TB of data per day

No human being can deal with this data avalanche!

!! ML is inherently data driven, data is at the core of ML. The goal of ML is to design general-purpose methodologies to extract valuable patterns from data, ideally without much domain-specific expertise.

Why Machine Learning?

Machine learning is a subfield of **artificial intelligence** that gives computers the ability to learn without explicitly being programmed. **Artificial intelligence systems** are used to perform complex tasks in a way that is similar to how humans solve problems.

Why Machine Learning?

Machine learning is a subfield of **artificial intelligence** that gives computers the ability to learn without explicitly being programmed. **Artificial intelligence systems** are used to perform complex tasks in a way that is similar to how humans solve problems.

“Machine learning is changing, or will change, every industry, and leaders need to **understand** the basic principles, the potential, and the limitations”
–Aleksander Madry, director of the MIT Center for Deployable Machine Learning.

TECHNISCHE UNIVERSITÄT DARMSTADT

Why Machine Learning?

- It's amazing what modern Machine Learning methods can do!
  - Image Generation (DALL-E, Midjourney, ...)
  - Text Generation (ChatGPT, ...)
  - Intelligent Robots (ANYmal, Boston Dynamics, ...)

TECHNISCHE UNIVERSITÄT DARMSTADT

Why Machine Learning?

- It's amazing what modern Machine Learning methods can do!
  - Image Generation (DALL-E, Midjourney, ...)
  - Text Generation (ChatGPT, ...)
  - Intelligent Robots (ANYmal, Boston Dynamics, ...)
- High demand on the job market & Great career opportunities

Why Machine Learning?

- It's amazing what modern Machine Learning methods can do!
  - Image Generation (DALL-E, Midjourney, ...)
  - Text Generation (ChatGPT, ...)
  - Intelligent Robots (ANYmal, Boston Dynamics, ...)
- High demand on the job market & Great career opportunities

In this class we will cover most of the topics you will need to get started in Machine Learning, like Projects, Thesis, Job Interviews

Machine Learning

What is ML? What is its goal?

Machine Learning

What is ML? What is its goal?

According to T. Mitchell (1997):

- 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.
  in other words...

TECHNISCHE UNIVERSITÄT DARMSTADT

Machine Learning

What is ML? What is its goal?

According to T. Mitchell (1997):

- 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.
  in other words…
- Develop a machine / an algorithm that learns to perform a task from experience.

Machine Learning

Why? What for?

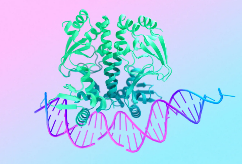
Machine Learning

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

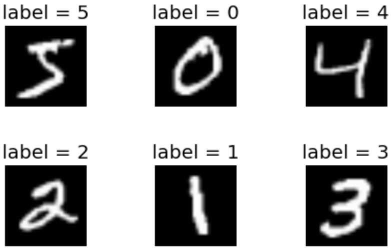
Machine Learning Impact & Successes

- **AlphaFold 3:** Predicts protein interactions with DNA/RNA — transforming biology and drug discovery.
- **Nobel Prize 2024:** Awarded to Hinton and Hopfield for their groundbreaking work in neural networks.
- **Autonomous Driving:** Tesla FSD V12.3 shows major real-world improvements.
- **ChatGPT / DeepSeek:** Powers human-like conversation, coding, tutoring, and creativity at scale.



Machine Learning Examples

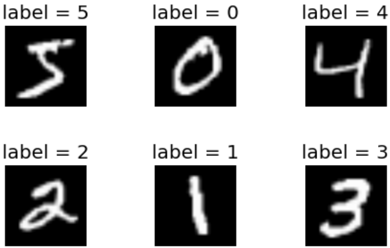
Recognition of handwritten digits



- These digits are given to us as small digital images

Machine Learning Examples

Recognition of handwritten digits

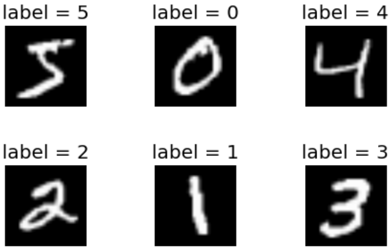


- These digits are given to us as small digital images
    - We have to build a “machine” to decide which digit it is

TECHNISCHE UNIVERSITÄT DARMSTADT

Machine Learning Examples

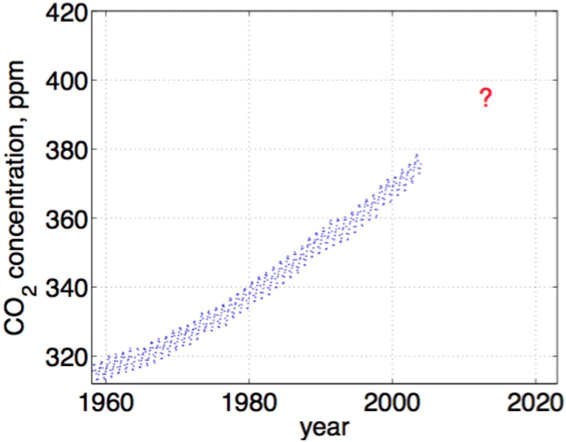
Recognition of handwritten digits



- 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

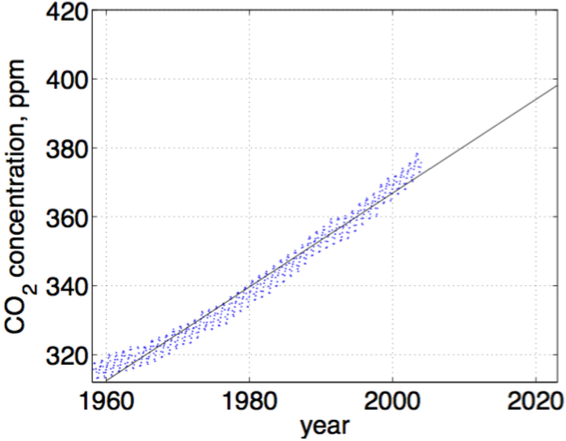
CO2 prediction



TECHNISCHE UNIVERSITÄT DARMSTADT

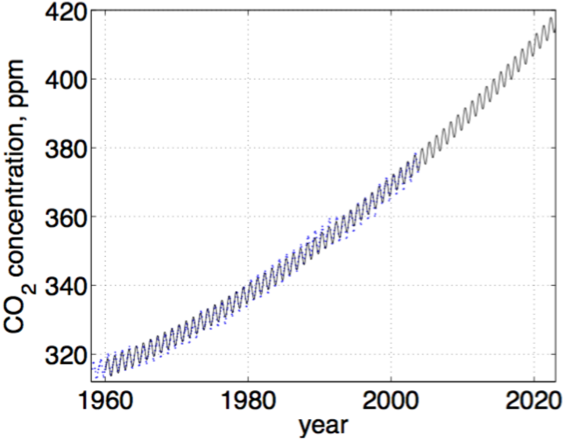
Machine Learning Examples

CO2 prediction



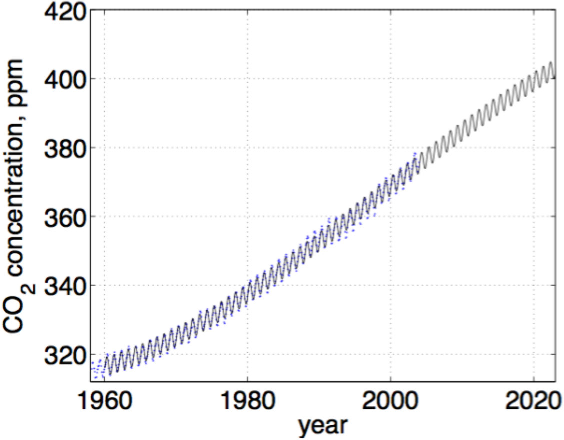
Machine Learning Examples

CO2 prediction



Machine Learning Examples

CO2 prediction



Machine Learning Examples

- **Email filtering**

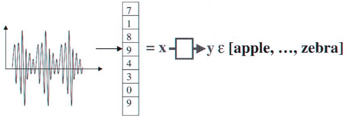
$x \ \varepsilon\ [a\text{-}z]+ \longrightarrow \boxed{} \longrightarrow y \ \varepsilon\ [\text{important, spam}]$

Machine Learning Examples

- **Email filtering**

$x \in [a\text{-}z]+ \longrightarrow y \in [\text{important}, \text{spam}]$

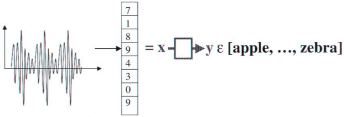
- **Speech recognition**



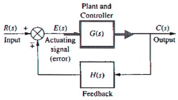
Machine Learning Examples

- **Email filtering**

- **Speech recognition**



- **Vehicle control**



TECHNISCHE UNIVERSITÄT DARMSTADT

Machine Learning

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

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.

$$f : I \rightarrow O$$

TECHNISCHE UNIVERSITÄT DARMSTADT

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.

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

TECHNISCHE UNIVERSITÄT DARMSTADT

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.

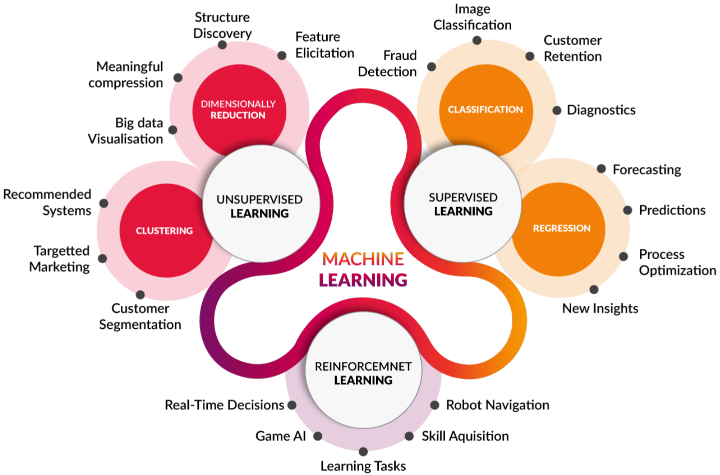
$$f : I \rightarrow O$$

  - 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”)

Types of Machine Learning



Source: `https://www.cognub.com/index.php/cognitive-platform/`

TECHNISCHE UNIVERSITÄT DARMSTADT

Classification vs Regression

Classification

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

TECHNISCHE UNIVERSITÄT DARMSTADT

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:

TECHNISCHE UNIVERSITÄT DARMSTADT

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.

TECHNISCHE UNIVERSITÄT DARMSTADT

Classification vs Regression

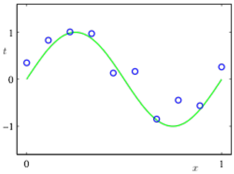
Regression

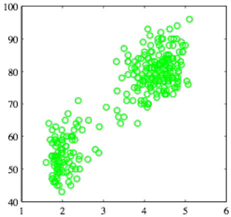
- Learn a mapping into a continuous space, e.g.
  - $O = \mathbb{R}$
  - $O = \mathbb{R}^3$
- Examples

Classification vs Regression

Regression

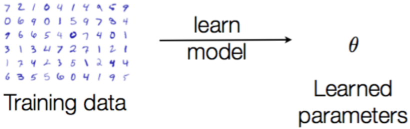
- Learn a mapping into a continuous space, e.g.
  - $O = \mathbb{R}$
  - $O = \mathbb{R}^3$
- Examples
  - Curve fitting, Financial Analysis, Housing prices, ...





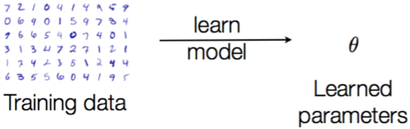
General Paradigm

Training

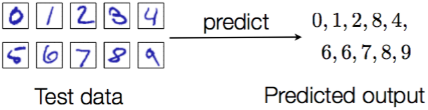


General Paradigm

Training



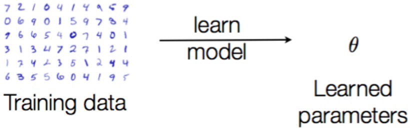
Testing



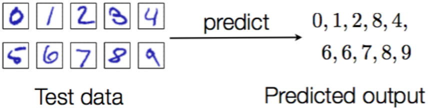
TECHNISCHE UNIVERSITÄT DARMSTADT

General Paradigm

Training



Testing



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

But ideally from the same underlying distribution.

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

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

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

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: learning-by-doing
  - Reinforcement learning

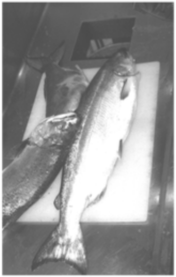
TECHNISCHE UNIVERSITÄT DARMSTADT

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.

Generalization

How do we achieve generalization?



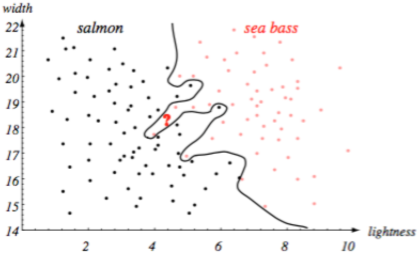


FIGURE 1.5. Overly complex models for the fish will lead to decision boundaries that are complicated. While such a decision may lead to perfect classification of our training samples, it would lead to poor performance on future patterns. The novel test point marked **?** is evidently most likely a salmon, 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?

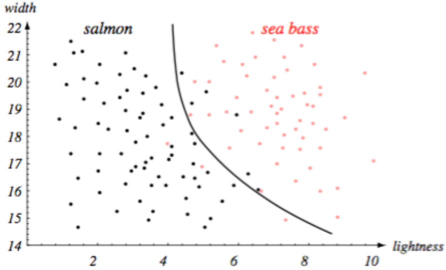


FIGURE 1.6. The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001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

Occam's Razor

TECHNISCHE UNIVERSITÄT DARMSTADT

Generalization

How do we achieve generalization?

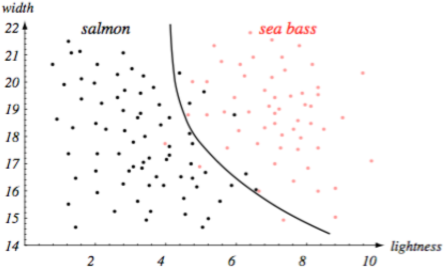


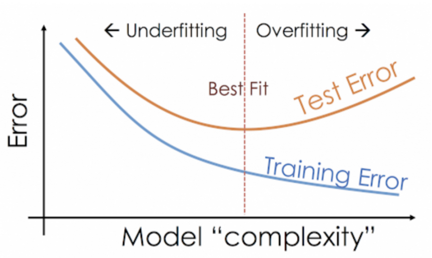
FIGURE 1.6. The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001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

Occam's Razor

We should not make the model overly complex!

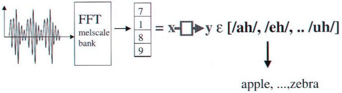
TECHNISCHE UNIVERSITÄT DARMSTADT

Example of overfitting...



Some Key Challenges

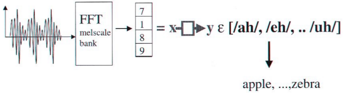
- Input:



TECHNISCHE UNIVERSITÄT DARMSTADT

Some Key Challenges

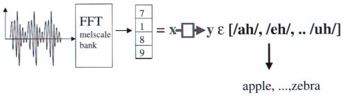
- Input:



- Features

Some Key Challenges

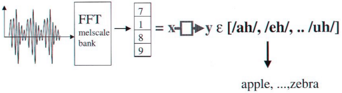
- Input:



- 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.

TECHNISCHE UNIVERSITÄT DARMSTADT

Some Key Challenges

- How do we measure accuracy?
  - 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?

Some Key Challenges

- How do we measure accuracy?
  - 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.

TECHNISCHE UNIVERSITÄT DARMSTADT

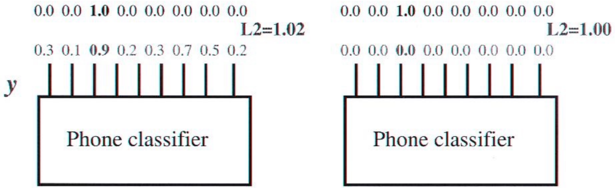
Some Key Challenges

- How do we measure accuracy?
  - 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 accuracy!

TECHNISCHE UNIVERSITÄT DARMSTADT

Some Key Challenges

- We also need to define the right error metric:

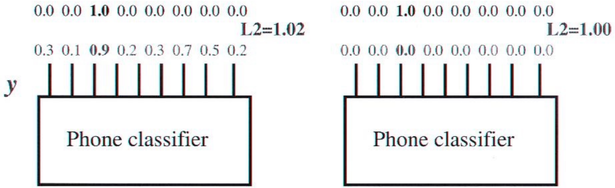


- Which is better?

TECHNISCHE UNIVERSITÄT DARMSTADT

Some Key Challenges

- We also need to define the right error metric:



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

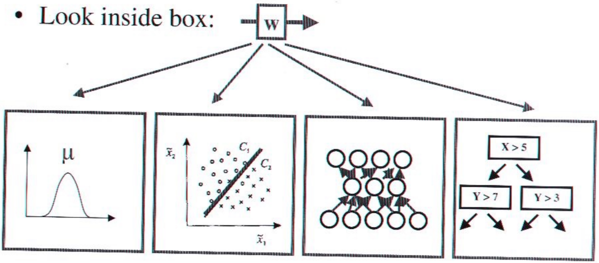
Some Key Challenges

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, ...



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

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

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

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

TECHNISCHE UNIVERSITÄT DARMSTADT

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!

TECHNISCHE UNIVERSITÄT DARMSTADT

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
  - Natural language processing
  - 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 one of the most promising routes to such a 80-160k€ job...

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 one of the most promising routes to such a 80-160k€ job…
- It is beating traditional hand-engineered methods in many tasks (e.g., vision, natural language, ...)
  - To solve such problems, knowledge of ML *and* the specific domain is required.

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 one of the most promising routes to such a 80-160k€ job…
- It is beating traditional hand-engineered methods in many tasks (e.g., vision, natural language, …)
  - To solve such problems, knowledge of ML *and* the specific domain is required.
- Because it is fun!

Preliminary Syllabus

Preliminary Syllabus

- Refresher of Linear Algebra & Statistics (next week)

Preliminary Syllabus

- Refresher of Linear Algebra & Statistics (next week)
- Fundamentals Part I (~ 2 weeks)
  - Bayes decision theory, Decision Boundary, Risk Minimization
  - Probability density estimation
  - Mixture models, expectation maximization

Preliminary Syllabus

- Linear Methods (~ 3 weeks)
  - Classification, Fisher linear discriminant
  - Linear regression
  - Generalized linear models
  - PCA, robust PCA

TECHNISCHE UNIVERSITÄT DARMSTADT

Preliminary Syllabus

- Linear Methods (~ 3 weeks)
  - Classification, Fisher linear discriminant
  - Linear regression
  - Generalized linear models
  - PCA, robust PCA
- Fundamentals Part II (~ 1 week)
  - Clustering
  - Performance evaluation

TECHNISCHE UNIVERSITÄT DARMSTADT

Preliminary Syllabus

- Large-Margin Methods (~ 2 weeks)
  - Statistical learning theory
  - Support vector machines
  - Kernel methods

Preliminary Syllabus

- Large-Margin Methods (~ 2 weeks)
  - Statistical learning theory
  - Support vector machines
  - Kernel methods
- Refresher of Optimization (~ 1 week)

Preliminary Syllabus

- Large-Margin Methods (~ 2 weeks)
  - Statistical learning theory
  - Support vector machines
  - Kernel methods
- Refresher of Optimization (~ 1 week)
- Neural Networks (~ 2 week)
  - Neural Networks: From Inspiration to Application
  - Deep Learning: What is really different?

Credits

- These slides have been originally developed by Prof. Bernt Schiele and later modified by Profs. Jan Peters, Kristian Kersting, Stefan Schaal, Stefan Roth and myself for previous iterations of this course or similar classes.
- Many figures are directly taken out of the books by Chris Bishop and Duda, Hart & Stork and Kevin Murphy.

TECHNISCHE UNIVERSITÄT DARMSTADT

Outline

3. Wrap-Up

You know now:

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 examples of Machine Learning applications.

3. Wrap-Up

You know now:

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

3. Wrap-Up

You know now:

- What Machine Learning is and what it is not.
- Some examples 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 examples 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?

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)?

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)?
- What are the challenges when solving a Machine Learning problem?

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)?
- What are the challenges when solving a Machine Learning problem?
- What is generalization? What is overfitting?

TECHNISCHE UNIVERSITÄT DARMSTADT

Outlook

Change of order:

- 12:35 – 13:20: Exercise Session discussing homeworks (pre/post discussion) and Q&A
- 13:25 – 15:05: Lecture
- Both in S311/08

Outlook

Change of order:

- 12:35 – 13:20: Exercise Session discussing homeworks (pre/post discussion) and Q&A
- 13:25 – 15:05: Lecture
- Both in S311/08
- Exceptions;
  - No lecture on the 4th of June (TU meet and move)
  - 16th of July in room S101/A03

Outlook

Change of order:

- 12:35 – 13:20: Exercise Session discussing homeworks (pre/post discussion) and Q&A
- 13:25 – 15:05: Lecture
- Both in S311/08
- Exceptions;
  - No lecture on the 4th of June (TU meet and move)
  - 16th of July in room S101/A03
- Next week: Refresher Linear Algebra & Statistics

Before You Go

Please remember:

Register for an exercise group on Moodle
by next week, in teams of 3 students.

This is important for participating in the homework exercises.