# Machine Learning I

Lucas Fabian Naumann, David Stein, Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

https://mlcv.cs.tu-dresden.de/courses/25-winter/ml1/

Winter Term 2025/2026

# Machine Learning I

https://mlcv.cs.tu-dresden.de/courses/25-winter/ml1/

- Course consisting of
  - lectures in TRE/PHYS/E on Fridays, 9:20–10:50
  - exercise groups starting October 21st

| | |
|---|---|
| In VMB/0302/U | on Tuesdays, 16:40–18:10 |
| In APB/E001/U | on Thursdays, 16:40–18:10 |
| In APB/E001/U | on Fridays, 14:50–16:20 |
| In APB/E001/U | on Fridays, 16:40–18:10 |
| Online (see OPAL) | on Thursdays, 16:40–18:10 |

  - self-study
  - final examination (covering lectures and exercises).
- Registration:
  - All participating students need to register through OPAL
  - All participating students enrolled in the study program Computational Modeling and Simulation need to register additionally via CampusNet.
- No recordings/reproductions of the lectures or exercises!

# Machine Learning I

**Machine Learning** is an area of computer science focused on the research of mathematical models and algorithms for analyzing, understanding and interpreting data, and for deciding and acting based on data. It

- poses challenging problems
- combines insights and methods from
  - mathematics (esp. combinatorics, optimization, probability theory, statistics)
  - computer science (esp. algorithms, complexity, software engineering)
- provides an opportunity for applying analytical and engineering skills
- has impact on applications (scientific, medical, robotic, consumer)

This introductory course focuses on machine learning **problems** and **algorithms**. It consists of three main parts:

**Supervised learning**

- Binary decision trees
- Disjunctive normal forms
- Linear functions
- Composite functions (i.e. deep learning)
- Transformers (with attention)

Exercises only!

**Unsupervised learning**

- Partitioning (i.e. clustering of sets)
- Ordering
- Classifying with multiple labels

**Structured learning**

- Conditional graphical models
- Message passing algorithms
- Pseudo-Boolean optimization

**Linear and integer optimization for machine learning**

Prerequisites:

- Mathematics
  - Linear algebra
  - Multivariate calculus (basics)
  - Probability theory (basics)
- Computer Science
  - Algorithms and data structures (basics)
  - Theoretical computer science (basics of complexity theory)

# Machine Learning I

- Textbooks:
  - Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press 2012
  - Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Jonathan Taylor. An Introduction to Statistical Learning. Springer 2023
  - Christopher M. Bishop, Hugh Bishop. Deep Learning: Foundations and Concepts. Springer 2024
  - Marc Peter Deisenroth. Mathematics for Machine Learning. Cambridge University Press 2020
- Scholarly journals (selection):
  - Journal of Machine Learning Research (JMLR)
  - Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
  - Transactions on Machine Learning Research (TMLR)
- Academic conferences (selection):
  - International Conference on Machine Learning (ICML)
  - Neural Information Processing Systems (NeurIPS)
  - International Conference on Learning Representations (ICLR)

https://mlcv.cs.tu-dresden.de/teaching.html

Related courses we are offering this term:

- **Computer Vision I**
  BEY/0E40/H (George-Bähr-Str. 1), Mondays, 11:10–12:40
- **Research Projects**
  APB-2026, Mondays, 14:50–16:20. Kick-off meeting: October 20th
  Creditable to INF-25-MA-FP, INF-PM-FPG, INF-PM-FPA and CMS-PRO

Notation:

- Considering that $0 = \emptyset$ and for any $m \in \mathbb{N}$: $m = \{0, \dots, m-1\}$, we may write $j \in m$ instead of $j \in \{0, \dots, m-1\}$.
- For any finite set $A$, let $|A|$ denote the number of elements of $A$.
- For any set $A$, let $2^A$ denote the power set of $A$.
- For any set $A$ and any $m \in \mathbb{N}_0$, let $\binom{A}{m}$ denote the set of all $m$-elementary subsets of $A$, i.e. $\binom{A}{m} = \{B \in 2^A \colon |B| = m\}$.
- For any sets $A, B$, let $B^A$ denote the set of all maps from $A$ to $B$. Moreover, let $AB$ be shorthand for the ordered pair $(A, B)$.
- For any $f \in B^A$, any $a \in A$ and any $b \in B$, we may write $b = f(a)$ or $b = f_a$ instead of $(a, b) \in f$.
- Let $\langle \cdot, \cdot \rangle$ denote the standard inner product, and let $\| \cdot \|$ denote the $l_2$-norm.
- Given any set $J$ and, for any $j \in J$, a set $S_j$, we denote by $\prod_{j \in J} S_j$ the Cartesian product of the family $\{S_j\}_{j \in J}$, i.e.

$$\prod_{j \in J} S_j = \left\{ f \colon J \to \bigcup_{j \in J} S_j \;\middle|\; \forall j \in J \colon f(j) \in S_j \right\} \quad . \tag{1}$$