Machine Learning I

Bjoern Andres (Lectures)
Shengxian Zhao and Jérome Thiessat (Exercises)

Machine Learning for Computer Vision
TU Dresden

Winter Term 2021/2022

- Online course consisting of
  - Live video lectures with Q&A on **Fridays, 14:50–16:20**
  - Live discussion of assignments with Q&A, from October 25th:
    - **Mondays, 11:10–12:40**
    - **Mondays, 16:40–18:10**
    - **Thursdays, 14:50–16:20**
  - Assignments, self-study and moderated discussion in a forum
- **Course website**:
  https://mlcv.inf.tu-dresden.de/courses/21-winter/ml1/index.html
- All students need to register via **OPAL**. All students of the study program CMS need to register additionally via **SELMA**.
- Contents of the exercises will be part of the examination
- Textbooks:
  - Barber, Bayesian Reasoning and Machine Learning
  - Bishop, Pattern Recognition and Machine Learning
  - Shai, Understanding Machine Learning

Machine Learning is a branch of computer science and a scientific community that *studies* and *develops* mathematical models and algorithms for understanding and interpreting data, as well as for deciding and acting wrt. data.

- Poses challenging problems
- Combines insights and methods from
  - Mathematics (esp. optimization, probability theory, statistics)
  - Computer Science (esp. algorithms, complexity, software engineering)
- Provides an opportunity for applying analytical and engineering skills
- Has impact on applications (medical, robotic, consumer)
- Grows dynamically
  - Excellent career opportunities (start-up companies, established corporations, public sector)

Machine Learning is a branch of computer science and a scientific community that *studies* and *develops* mathematical models and algorithms for understanding and interpreting data, as well as for deciding and acting wrt. data.

- Leading scholarly journal:
  - Journal of Machine Learning Research (JMLR)
- Leading academic conferences:
  - International Conference on Machine Learning (ICML)
  - Neural Information Processing Systems (NeurIPS)
  - International Conference on Learning Representations (ICLR)
- Closely related scientific communities:
  - Learning theory (e.g. ALT, COLT)
  - Artificial Intelligence (e.g. IJCAI, AAAI, UAI, AISTATS)

Contents

- **Supervised learning**
  - Disjunctive normal forms
  - Binary decision trees
  - Linear functions
  - Artificial neural networks
- **Semi-supervised and unsupervised learning**
  - Partitioning
  - Clustering
  - Ordering
- **Structured learning**
  - Conditional graphical models
- **Density estimation**
- **Embedding**
- **Applications**

Prerequisites

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

Notation

- We write “iff” as shorthand for “if and only if”.

Notation

- We write "iff" as shorthand for "if and only if".
- For any finite set A, we denote by $|A|$ the number of elements of A.

Notation

- We write “iff” as shorthand for “if and only if”.
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.

Notation

- We write "iff" as shorthand for "if and only if".
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.
- For any set A and any $m \in \mathbb{N}$, we denote by $\binom{A}{m}$ the set of all m-elementary subsets of A, i.e. $\binom{A}{m} = \{B \in 2^A \colon |B| = m\}$.

Notation

- We write "iff" as shorthand for "if and only if".
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.
- For any set A and any $m \in \mathbb{N}$, we denote by $\binom{A}{m}$ 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, we denote by B^A the set of all maps from A to B

Notation

- We write "iff" as shorthand for "if and only if".
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.
- For any set A and any $m \in \mathbb{N}$, we denote by $\binom{A}{m}$ 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, we denote by B^A the set of all maps from A to B
- For any map $f \in B^A$, any $a \in A$ and any $b \in B$, we may write $b = f(a)$ or $b = f_a$ instead of $(a, b) \in f$

- We write "iff" as shorthand for "if and only if".
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.
- For any set A and any $m \in \mathbb{N}$, we denote by $\binom{A}{m}$ 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, we denote by B^A the set of all maps from A to B
- For any map $f \in B^A$, any $a \in A$ and any $b \in B$, we may write $b = f(a)$ or $b = f_a$ instead of $(a, b) \in f$
- 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.

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

Notation

- We write "iff" as shorthand for "if and only if".
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.
- For any set A and any $m \in \mathbb{N}$, we denote by $\binom{A}{m}$ 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, we denote by B^A the set of all maps from A to B
- For any map $f \in B^A$, any $a \in A$ and any $b \in B$, we may write $b = f(a)$ or $b = f_a$ instead of $(a, b) \in f$
- 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.

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

- We denote by $\langle \cdot, \cdot \rangle$ the standard inner product, and by $\| \cdot \|$ the l_2-norm.

- We write “iff” as shorthand for “if and only if”.
- For any finite set A, we denote by $|A|$ the number of elements of A.
- For any set A, we denote by 2^A the power set of A.
- For any set A and any $m \in \mathbb{N}$, we denote by $\binom{A}{m}$ 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, we denote by B^A the set of all maps from A to B
- For any map $f \in B^A$, any $a \in A$ and any $b \in B$, we may write $b = f(a)$ or $b = f_a$ instead of $(a, b) \in f$
- 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.

$$\prod_{j \in J} S_j = \left\{ f \colon J \to \bigcup_{j \in J} S_j \,\middle|\, \forall j \in J \colon f(j) \in S_j \right\} \qquad (1)$$

- We denote by $\langle \cdot, \cdot \rangle$ the standard inner product, and by $\| \cdot \|$ the l_2-norm.
- For any $m \in \mathbb{N}$, we define $[m] = \{0, \ldots, m-1\}$.