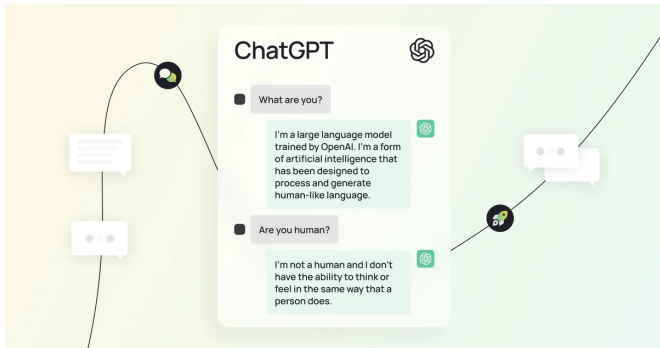


Künstliche Intelligenz (Sommersemester 2024)

# Kapitel 02: Machine Learning – Grundlagen

Prof. Dr. Adrian Ulges



*"Machine intelligence is the last invention that humanity will ever need to make."*

(Nick Bostrom, "Superintelligence")

# Maschinelles Lernen (ML): Bereiche aktueller Erfolge

image: [4]



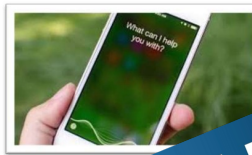
autonomous vehicles



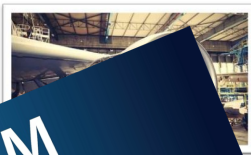
dialog systems



speech recognition



predictive maintenance



medical diagnosis

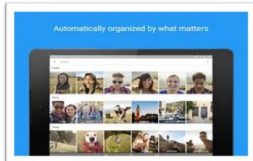


image recognition



machine



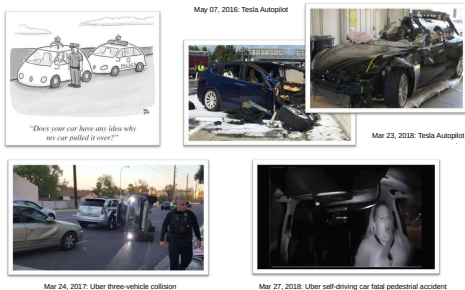
## 'Klassische' Anwendungsfelder

- ▶ **Computer Vision:** Handschrifterkennung, Objekterkennung, ...
- ▶ **Nutzermodellierung:** Suchmaschinen, Empfehlungssysteme, Targeting, ...
- ▶ **NLP:** Informationsentnahme, Stimmungsanalyse, Spam, ...

## Sonstige Anwendungsfelder ...

- ▶ Restaurant-Umsatzprognose  
*Vorhersage des Jahresumsatzes von zu eröffnenden Restaurants*
- ▶ Fahrertelematik-Analyse (AXA)  
*anhand von GPS-Routen den Fahrer eines Autos identifizieren*
- ▶ Wal-Erkennung  
*Walgesänge in Audio erkennen, Kollisionen mit Schiffsverkehr verhindern*
- ▶ ...

# ML: Misserfolge image: [4]



Tay Chatbot on Twitter (Microsoft, Mar 23, 2016):



- ▶ ML kommt zunehmend in **sicherheitskritischen Anwendungen** zum Einsatz: selbstfahrende Fahrzeuge, Gesundheit, Pharmazie, Aktienhandel ...
- ▶ Die regulatorischen Auswirkungen sind enorm (*Die Genauigkeit ist begrenzt*)!
- ▶ ML-Modelle sollten **sicher, fair, transparent, ressourceneffizient, datenschutzgerecht** sein.



*"The field of study that gives computers the ability to learn without being explicitly programmed."*

(Arthur Samuel (1959))

---

*"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))

*“Jedes If-Statement ist eine potenzielle Anwendung für maschinelles Lernen”*

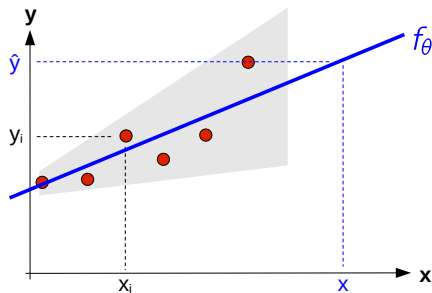
(Thomas M. Breuel (2004))

- ▶ Ein Computersystem soll eine nicht-triviale Entscheidung treffen, z.B. **Spam-Filterung**.
- ▶ Warum nicht die Entscheidungslogik **hart codieren**?

## Probleme

- ▶ Hoher **initialer Verandaufwand**.
- ▶ Schwierig, das **bestmögliche** Programm zu erreichen.
- ▶ **Überprüfung** der Optimalität ist schwierig.
- ▶ Code ist extrem schwierig zu **updaten/warten**.
- ▶ Das Verfolgen von **Datendrift** ist schwierig, wenn z.B. Spammer ihre Strategien ändern.
- ▶ Es gibt keine Möglichkeit, das **Feedback der Benutzer** zu berücksichtigen.





- ▶ **Ziel:** Vorhersage des **Gewichts einer Person** in der Zukunft!
- ▶ **Gegeben:** Stichprobe  $x_1, \dots, x_n$  (*Zeitpunkte*) mit sogenannten "Labels"  $y_1, \dots, y_n$  (*dem jeweiligen Gewicht der Person*).
- ▶ **Vorgehensweise** (*lineare Regression*):
  - ▶ Wir fitten eine Linie  $f_\theta$  auf die Punkte.
  - ▶ Gegeben einen Zeitpunkt  $x$ , verwenden wir  $\hat{y} := f_\theta(x)$  als Prognose des Gewichts.
- ▶ **Ist dies maschinelles Lernen?**



- Wir definieren unsere Linie als eine **Funktion**  $f$  mit **Parametern**  $\theta = (a, b)$

$$f_{\theta}(x) = a \cdot x + b$$

- Wir messen die **Qualität** einer bestimmten Linie  $f_{\theta}$  mit einer **Zielfunktion**  $\mathcal{L}$ :

$$\mathcal{L}(\theta) = \sum_{i=1}^n \left( f_{\theta}(x_i) - y_i \right)^2$$

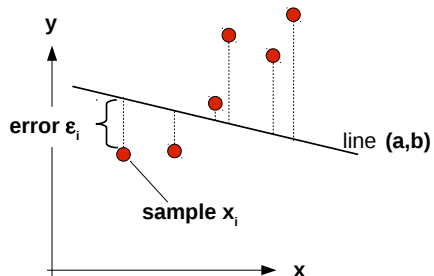
- Die **beste Linie** ist diejenige, die  $\mathcal{L}$  **minimiert**:

$$\theta^* = \arg \min_{\theta \in \mathbb{R}^2} \mathcal{L}(\theta)$$

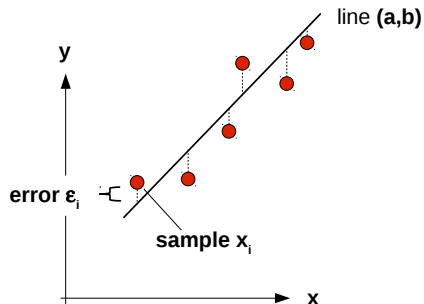
- Wir setzen die partiellen **Ableitungen**  $(\partial \mathcal{L} / \partial a, \partial \mathcal{L} / \partial b)$  gleich null. Es ergibt sich:

$$a^* = \left( \sum_i y_i x_i - \bar{y} \sum_i x_i \right) / \left( \sum_i x_i^2 - \bar{x} \sum_i x_i \right)$$
$$b^* = \frac{1}{n} \sum_i y_i - a^* \cdot \frac{1}{n} \sum_i x_i$$

schlechte Lösung  $\theta$ : Die Fehler  $\epsilon_1^2, \dots, \epsilon_n^2$  sind hoch.



gute Lösung  $\theta$ : Die Fehler  $\epsilon_1^2, \dots, \epsilon_n^2$  sind niedrig.





- ▶ Wir nennen die Punkte  $(x_1, y_1), \dots, (x_n, y_n)$  die **Trainingsdaten**.
- ▶ Die „wahren“ Werte  $y_i$  werden auch als die **Labels**, **Targets** oder **Grundwahrheit** (engl. “ground truth”) bezeichnet.
- ▶ Wir nennen unsere Linie  $f_\theta(x) = a \cdot x + b$  das **Modell**.
- ▶ Wir bezeichnen den Prozess der Schätzung der **Modellparameter**  $\theta = (a, b)$  als **Training** oder **Fitting**.
- ▶ Eine typische Trainingsstrategie besteht darin, eine **Zielfunktion** (engl. “objective function” oder “loss”)  $\mathcal{L}$  zu **optimieren**, oft der Form:

$$\mathcal{L}(\theta) = \frac{1}{n} \sum_i \ell(f_\theta(x_i), y_i)$$

## Anmerkungen

- ▶ Praktische Modelle haben deutlich mehr Parameter (*GPT-3.5*:  $\#\theta = 175$  Mrd.).
- ▶ Im obigen Beispiel haben wir die Lösung manuell abgeleiten können (wir sagen: es gibt eine *analytische* Lösung). In der Praxis ist  $\mathcal{L}$  meist **schwieriger zu optimieren**, und die Optimierung wird per **lokaler Suche** durchgeführt.

# ML ist multi-variat! image: [3]



- ▶ ML soll also Prognosen  $y$  über Eingabeobjekte  $x$  treffen.
- ▶ In der Praxis sind  $x$  und  $y$  keine Skalare, sondern **Vektoren**  $x$  und  $y$ !

PassengerId	Survived	Class	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
1	0	3	Brund, Mr. Owen Harris	male	22	1	0	OA5 21171	7.25		S
2	1	1	Corey, Mrs. John Bradley (Florence Briggs Thayer)	female	38	1	0	OPC 17369	71.2833	C86	C
3	1	1	Hickman, Miss. Laina	female	26	0	0	OSTON 35101282	70.95		S
4	1	1	Wheeler, Mrs. James (Lucy) A. McLeod	female	35	0	0	513400354	53.1	C123	S
5	0	3	Allen, Mr. William Henry	male	35	0	0	374503 25			S
6	0	3	Brown, Mr. James	male	0	0	0	5081 71 8355			S
7	0	3	McCarthy, Mr. Timothy J	male	54	0	0	17463 51 8625	5.48		S
8	0	3	Panson, Master. Costa Locurat	male	2	3	1	348829	21.015		S
9	0	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27	0	2	347462 11 1323	5		S
10	0	3	Nasser, Mrs. Nicholas (Adela Achmet)	female	14	0	0	337 36 10 6708			C
11	0	3	Stenstrom, Mrs. Marguerite Rut								
12	0	3	Bonnell, Mrs. Elizabeth								
13	0	3	Standbrook, Mr. William Henry								
14	0	3	Andersson, Mr. Anders Johan								
15	0	3	Vesterlund, Mrs. Hilda Amanda Adalfrida								
16	0	3	Hewlett, Mrs. Mary D (Kingscott)								
17	0	3	Wiles, Master. Eugene								
18	0	3	Williams, Mr. Charles Eugene								
19	0	3	Vander Planke, Mrs. Julius (Emelia Maria Vanderplanke)								
20	0	3	Masabiani, Mrs. Gabria								
21	0	3	Fynney, Mr. Joseph J								
22	0	3	Oweney, Mr. Lawrence								
23	0	3	McGowan, Mrs. Anna "Annie"								
24	0	3	Trapp, Mr. William Thompson								
25	0	3	Panson, Mrs. Tertong Daniels								
26	0	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia Johansson)								
27	0	3	Ernst, Mr. Famed Chahab								
28	0	3	Forsman, Mr. Charles Alexander								
29	0	3	O'Connor, Mrs. Ellen "Nellie"								
30	0	3	Tobinoff, Mr. Lulu								
31	0	3	Uusimurto, Dan. Manuel E								
32	0	3	Sponcer, Mrs. William Augustus (Mato Espen)								
33	0	3	Gibson, Mrs. Mary Agatha								
34	0	3	Whetton, Mr. Edward H								
35	0	3	Meyer, Mr. Edgar Joseph								
36	0	3	Huovinen, Mr. Alexander Oskar								
37	0	3	Marens, Mr. Herman								
38	0	3	Carr, Mr. Ernest Charles								
39	0	3	Vander Planke, Mrs. Augusta Maria								
40	0	3	Wilde-Yarnell, Mrs. Jennie								
41	0	3	Jahns, Mrs. Johan Johanna Persdotter (Larsson)								
42	0	3	Tuan, Mrs. William John Robert (Dorothy Ann Wainwright)								
43	0	3	Kraeff, Mr. Theodor								



- ▶ Unser Modell wird zu einer **multivariaten Funktion**  $f_{\theta} : \mathcal{X} \rightarrow \mathcal{Y}$ , wobei  $\mathcal{X} \subseteq \mathbb{R}^d$  und  $\mathcal{Y} \subseteq \mathbb{R}^{d'}$ .
- ▶ Wir bezeichnen die Einträge des **Merkmalsvektors**  $\mathbf{x}$ , z.B. Geschlecht, Alter ..., als **Merkmale** (engl. "features"), und nennen  $\mathcal{X}$  den **Merkmalsraum** (engl. "feature space").

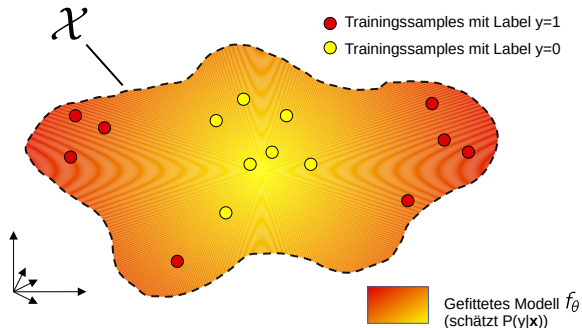
# ML ist multi-variater (cont'd)



## Anmerkungen

- ▶ Im Allgemeinen können viele Merkmale für das Zielproblem **irrelevant** sein. Während des Trainings müssen ML-Modelle die relevanten auswählen.
- ▶ Ein Merkmal kann auch erst in **Kombination** mit anderen Merkmalen nützlich sein.

# ML: Der Merkmalsraum ( "Feature Space" )



- ▶ Merkmalsvektoren  $\mathbf{x}$  können als **Punkte** im Merkmalsraum  $\mathcal{X}$  interpretiert werden.
- ▶ Wir können uns ein Modell  $f_\theta : \mathcal{X} \rightarrow \mathbb{R}$  als einen **Farbverlauf** vorstellen, der jedem Punkt  $\mathbf{x}$  einen Wert  $f_\theta(\mathbf{x})$  zuweist.
- ▶ Im obigen **Beispiel** schätzt das Modell  $f_\theta$  die Klassenzugehörigkeit eines Objekts  $\mathbf{x}$ , d.h.  $f_\theta(\mathbf{x}) \approx P(Y=1 | X=\mathbf{x})$ .

# References I



- [1] **Brizzle born and Bread.**  
<https://www.flickr.com/photos/brizzlebornandbred/5292576151/> (retrieved: Oct 2016).
- [2] **Spam (Monty Python).**  
[https://en.wikipedia.org/wiki/Spam\\_\(Monty\\_Python\)](https://en.wikipedia.org/wiki/Spam_(Monty_Python)) (retrieved: Oct 2016).
- [3] **'Untergang der Titanic' Illustration von Willy Stöwer für die Zeitschrift Die Gartenlaube.**  
[https://de.wikipedia.org/wiki/RMS\\_Titanic](https://de.wikipedia.org/wiki/RMS_Titanic) (retrieved: Oct 2016).
- [4] **Damian Borth.**  
Machine Learning (M.Sc. Course), University St. Gallen, summer term 2022.  
(retrieved: Aug 2022).