

# Machine Learning

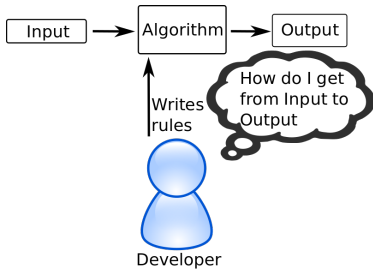
# Was ist ML-KA?

- ▶ Kurz für **Machine Learning Karlsruhe**
- ▶ **Hochschulgruppe** seit 15. Oktober 2015
- ▶ **Ziel:** Wissen über ML Verbreiten / Mehren
- ▶ **Idee:** Forum für interessierte Studenten bilden, organisation in kleinen Gruppen
- ▶ **Umsetzung** bisher
  - ▶ Paper Discussion Group (PDG, wöchentlich)
  - ▶ Gesellschaftliche Implikationen vom ML (GIML, 2-wöchentlich)
  - ▶ Teilnahme (und Preisträger) der Herbsttagung der Gesellschaft für Datenanalyse und Numerische Klassifikation
- ▶ Mehr auf <https://ml-ka.de>

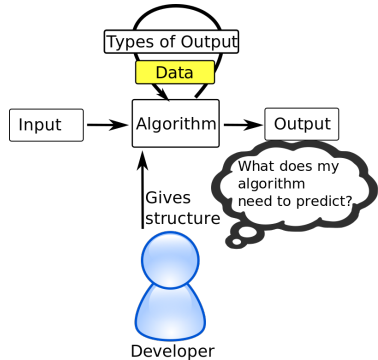
# Wer ist ML-KA?

- ▶ Vorstand:
  - ▶ Martin Thoma ([info@martin-thoma.de](mailto:info@martin-thoma.de))
  - ▶ Marvin Teichmann
  - ▶ Marvin Schweizer
- ▶ Mitglieder: Überwiegend aktuelle Studenten des KIT
  - ▶ 20 Mitglieder (Stand: 3. Februar 2016)
  - ▶ 213 Facebook Mitglieder (Stand: 26. Mai 2016)

# Was ist Machine Learning?

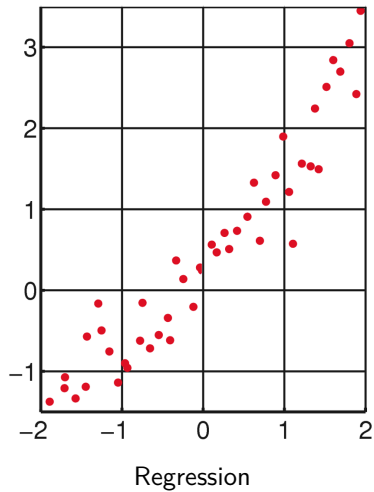


Traditional Development Model

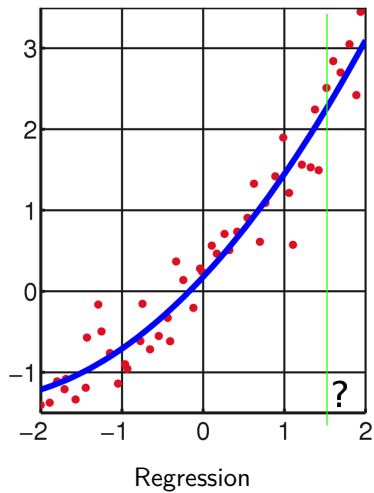


ML Model

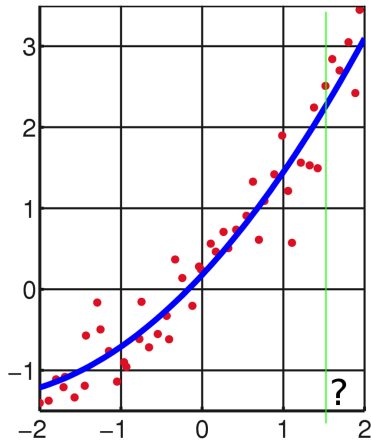
# Problemtypen



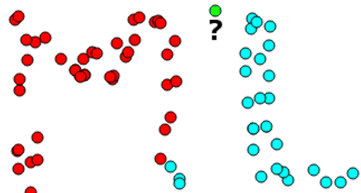
# Problemtypen



# Problemtypen

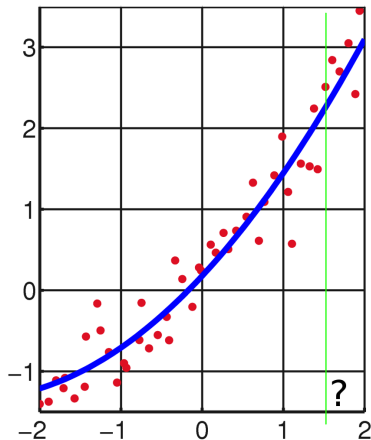


Regression

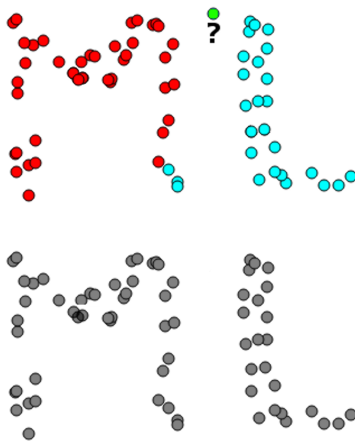


Klassifikation (überwacht)

# Problemtypen



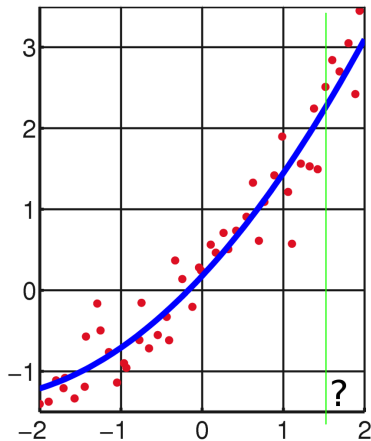
Regression



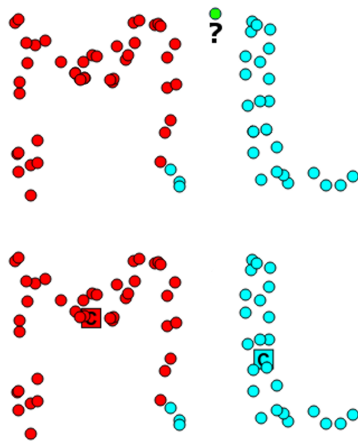
und Clustering (unüberwacht)



# Problemtypen




























Regression



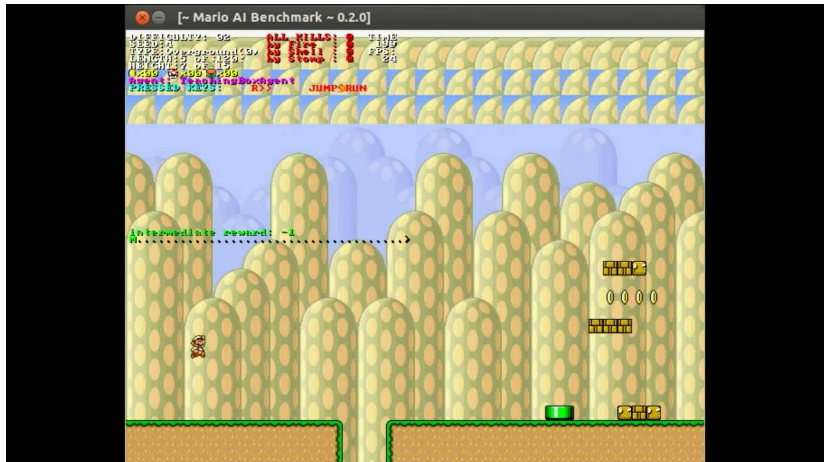
und Clustering (unüberwacht)

# Was ist Machine Learning?

Collaborative Filtering

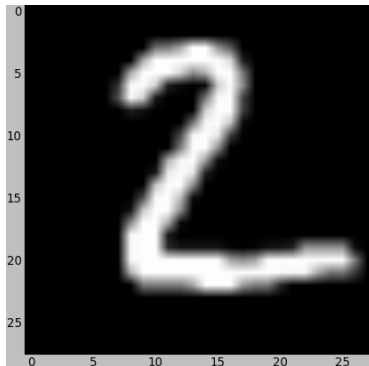
# Was ist Machine Learning?



Reinforcement Learning (RL)

# MNIST - Ziffern klassifizieren

- ▶ Klassen: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
- ▶ 60 000 Trainingsdaten, 10 000 Testdaten auf [yann.lecun.com/exdb/mnist](http://yann.lecun.com/exdb/mnist)
- ▶ Algorithmen zur Klassifizierung: **SVMs** (Support Vector Machines), **CNNs** (Convolutional Neural Networks),  $k$  Nearest Neighbors (siehe [tinyurl.com/knn-interact](http://tinyurl.com/knn-interact))



Datensatz der Klasse "2";  
28 px × 28 px

# Wie löst man das?

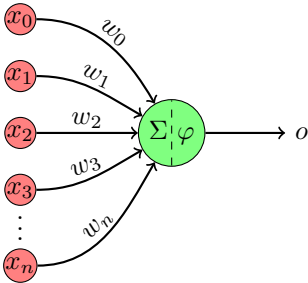
- **Situation:** Daten im  $\mathbb{R}^{28 \times 28}$ , Lösungen in  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

# Wie löst man das?

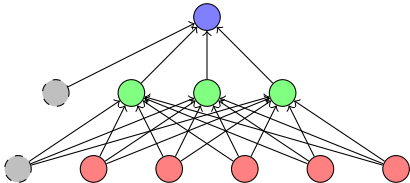
- ▶ **Situation:** Daten im  $\mathbb{R}^{28 \times 28}$ , Lösungen in  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
- ▶ ... oder  $[0, 1]^{10}$ , wenn wir eine Wahrscheinlichkeitsverteilung wollen

# Wie löst man das?

- ▶ **Situation:** Daten im  $\mathbb{R}^{28 \times 28}$ , Lösungen in  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
- ▶ ... oder  $[0, 1]^{10}$ , wenn wir eine Wahrscheinlichkeitsverteilung wollen
- ▶ **Lösung:** Neuronale Netze!



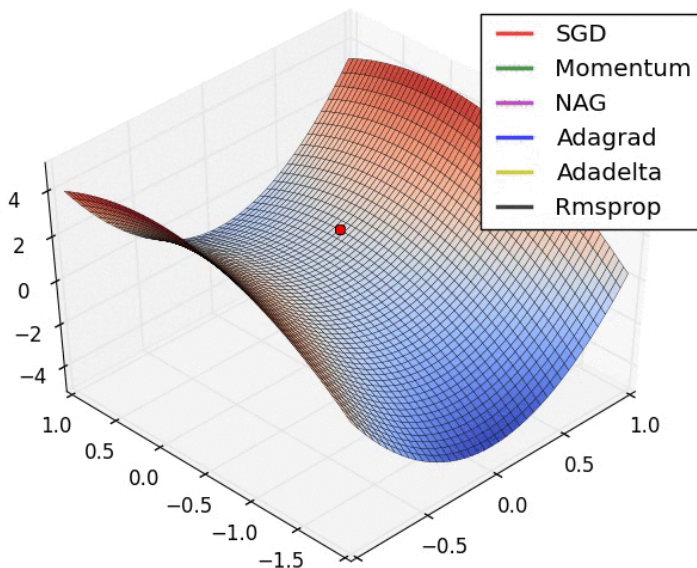




Wie finden wir die Gewichte?

Wie finden wir die Gewichte?  
Gradientenabstieg  
(Mehrdimensionale Ableitung)

# Gradientenabstieg



Quelle: <http://imgur.com/a/Hqolp>

# ImageNet

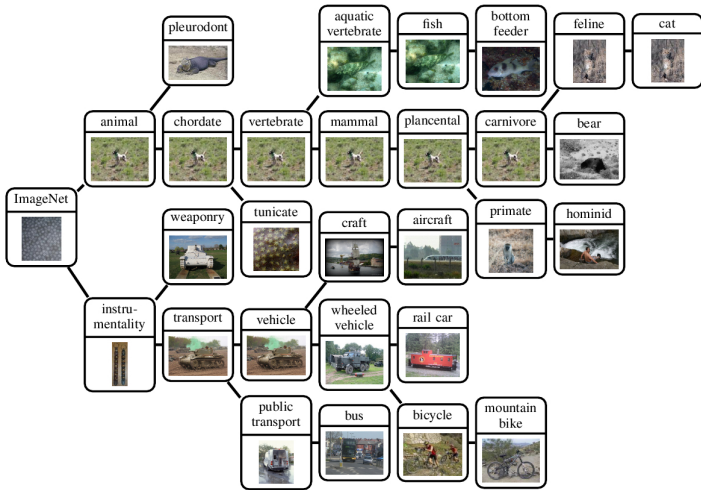
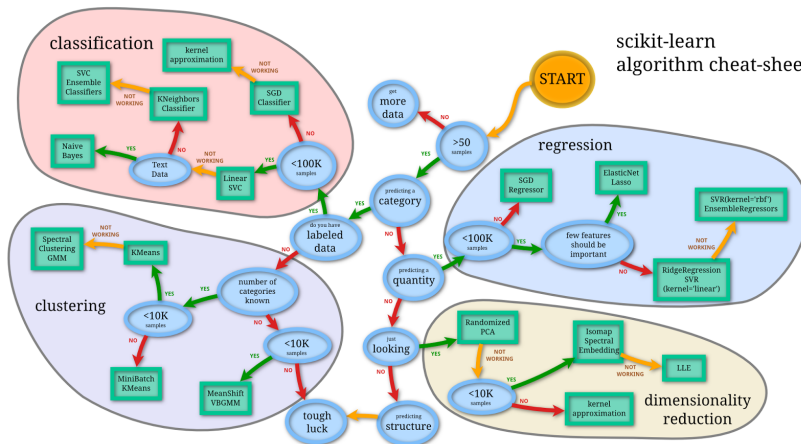


Image by Thomas Deselaers

21 841 Synsets, 14 197 122 Bilder



# Weitere Tools

- ▶ TensorFlow ([Tutorials](#))
- ▶ [TensorBox](#) basiert auf TensorFlow, Lokalisierung (Computer Vision)
- ▶ [Keras](#): Sehr einfach zu bedienen, abstrahiert von TensorFlow
- ▶ Datenvisualisierung
  - ▶ [Pandas](#)
  - ▶ [Seaborn](#)