

# Cross-Validierung

Evaluierung mit wenig Daten

**Mit mehr Daten werden die Modelle  
oft besser und die  
Evaluierungsergebnisse  
zuverlässiger.**

## Das Problem bei Datenmangel

### Möglichst gutes Modell

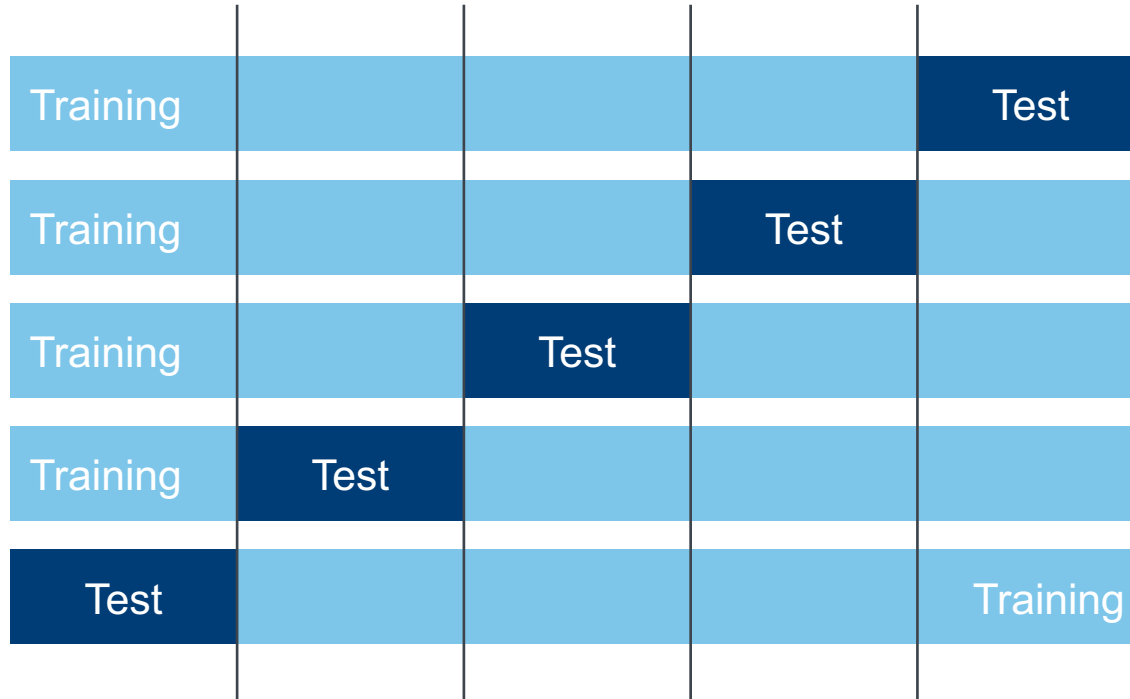
- So viele Daten zum Trainieren wie möglich

### Möglichst gute Einschätzung der Performanz

- So viele Daten zum Testen wie möglich
- Sonst evtl. per Zufall gute/schlechte Ergebnisse

# 5-fache Kreuzvalidierung

5  
Teilmengen



Möglichst viele Daten zum Trainieren

⇒ Bessere Modelle

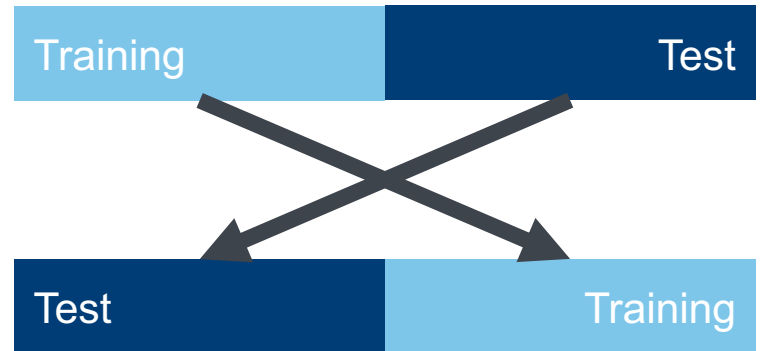
Evaluierung auf verschiedenen Teilmengen

⇒ Mittelwert über Evaluierungsergebnisse

⇒ Objektiver, weniger Zufallseffekte

## Warum heißt das Kreuzvalidierung?

- Evaluierung „über Kreuz“
- Bei 2-facher Kreuzvalidierung:  
offensichtlich 😊
- Übrigens: Der englische Begriff „Cross-Validierung“ ist beim Maschinellen Lernen gebräuchlicher als der deutsche



## n-fache Cross-Validierung

- Je größer  $n$ , desto mehr Daten bleiben zum Training
  - Aber desto mehr Modelle müssen trainiert und evaluiert werden
- ⇒ Kompromiss finden!

**Bei der n-fachen Cross-Validierung teilt man die Daten in n Teilmengen auf, die sich nicht überlappen. Man hält dann jede dieser Teilmengen einmal für die Evaluierung zurück, evaluiert also n mal. Das zu testende Modell wird jeweils auf den restlichen n-1 Teilmengen trainiert.**

**Für das Gesamt-Evaluierungsergebnis bildet man den Mittelwert der Evaluierungsergebnisse der n Modelle.**

## Ein cross-validiertes Modell verwenden

- n Modelle -  
Welches Modell  
sollte denn nun  
eingesetzt werden?
  - Finales Modell auf  
**allen** Daten  
trainieren
- ⇒ Cross-Validierung  
nur zur objektiven  
Einschätzung der  
Performanz



Training

Test



Training

Test



Training

Test



Training

Test



Test

Training



Training

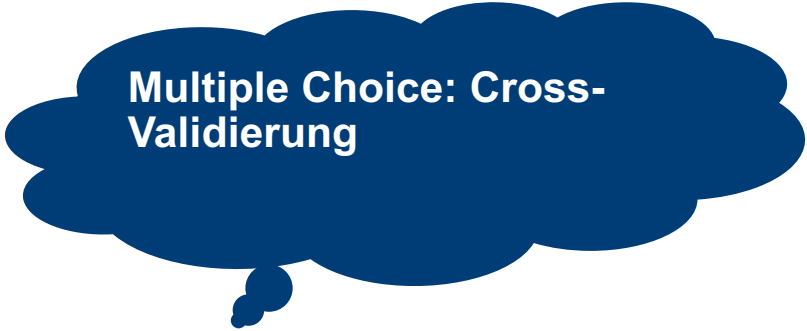


**Wenn man ein Modell einsetzen möchte, das man vorher per Cross-Validierung evaluiert hat, trainiert man das finale Modell am besten auf allen Daten neu.**

**Die Evaluierungsergebnisse aus der Cross-Validierung geben dann eine Einschätzung für die (Mindest-) Performanz dieses finalen Modells.**



**Drag the words: Cross-Validierung**



**Multiple Choice: Cross-Validierung**

## Dr. Antje Schweitzer

Universität Stuttgart  
Institut für Maschinelle Sprachverarbeitung



**Universität Stuttgart**

Institut für Maschinelle Sprachverarbeitung  
Institut für Software Engineering



**IHK** Industrie- und Handelskammer  
Reutlingen

Reutlingen | Tübingen | Zollernalb



**IHK** Region Stuttgart



**IHK** Industrie- und Handelskammer  
Karlsruhe



# Lizenzbestimmungen

“Cross-Validierung“ von Antje Schweitzer, KI B<sup>3</sup> / Uni Stuttgart

Das Werk - mit Ausnahme der folgenden Elemente:

- Logos der Verbundpartner und des Förderprogramms
- im Quellenverzeichnis aufgeführte Medien

ist lizenziert unter:

 [CC BY 4.0 \(https://creativecommons.org/licenses/by/4.0/deed.de\)](https://creativecommons.org/licenses/by/4.0/deed.de)

(Namensnennung 4.0 International)

## Quellenverzeichnis

Titelfoto: [Suraj Kardile \(https://unsplash.com/@suraj09\)](https://unsplash.com/@suraj09), ohne Titel, auf [Unsplash \(https://unsplash.com/photos/Rt2DsRJP9IQ\)](https://unsplash.com/photos/Rt2DsRJP9IQ), ist lizenziert unter [Unsplash-Lizenz \(https://unsplash.com/license\)](https://unsplash.com/license).

Bildausschnitt verändert.