



Praktische Zeitreihenklassifikation mit Sktime

Prof. Dr. Alexander Schiendorfer



Almotion Bavaria

A I N I N



Agenda



- Wrap-up und Fragen
- Tipps für die Praxis

15
min

- Grundbegriffe Zeitreihen
- Machine Learning

15
min

15
min

- Faltungsbasierte Ansätze
 - Rocket
 - CNNs

15
min

- Nützliche Tools in Sktime
- Transformationen & Pipelines

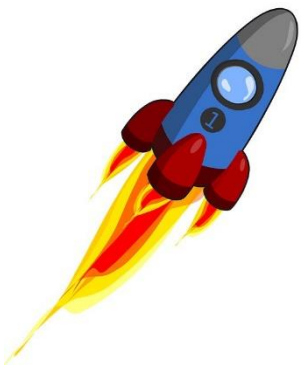
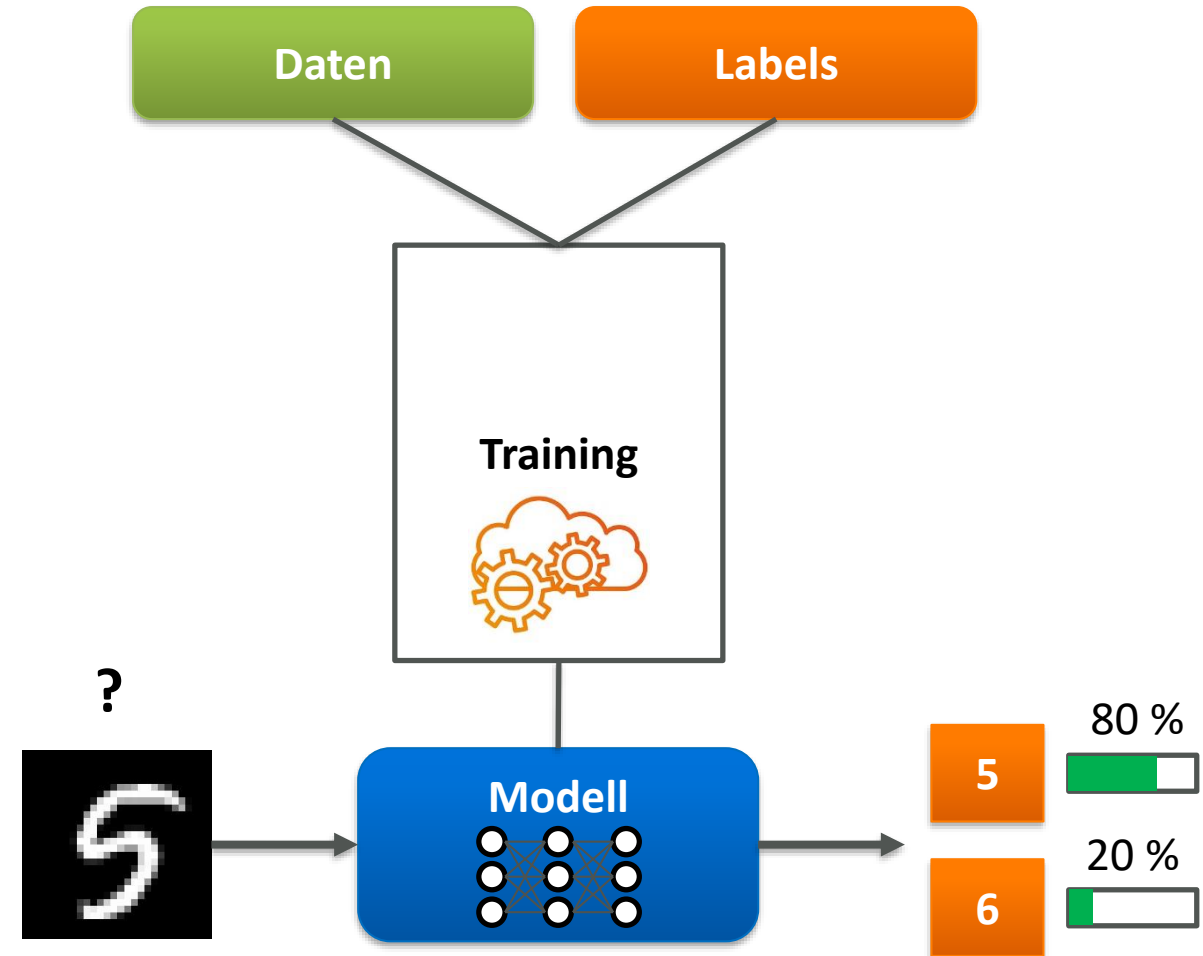
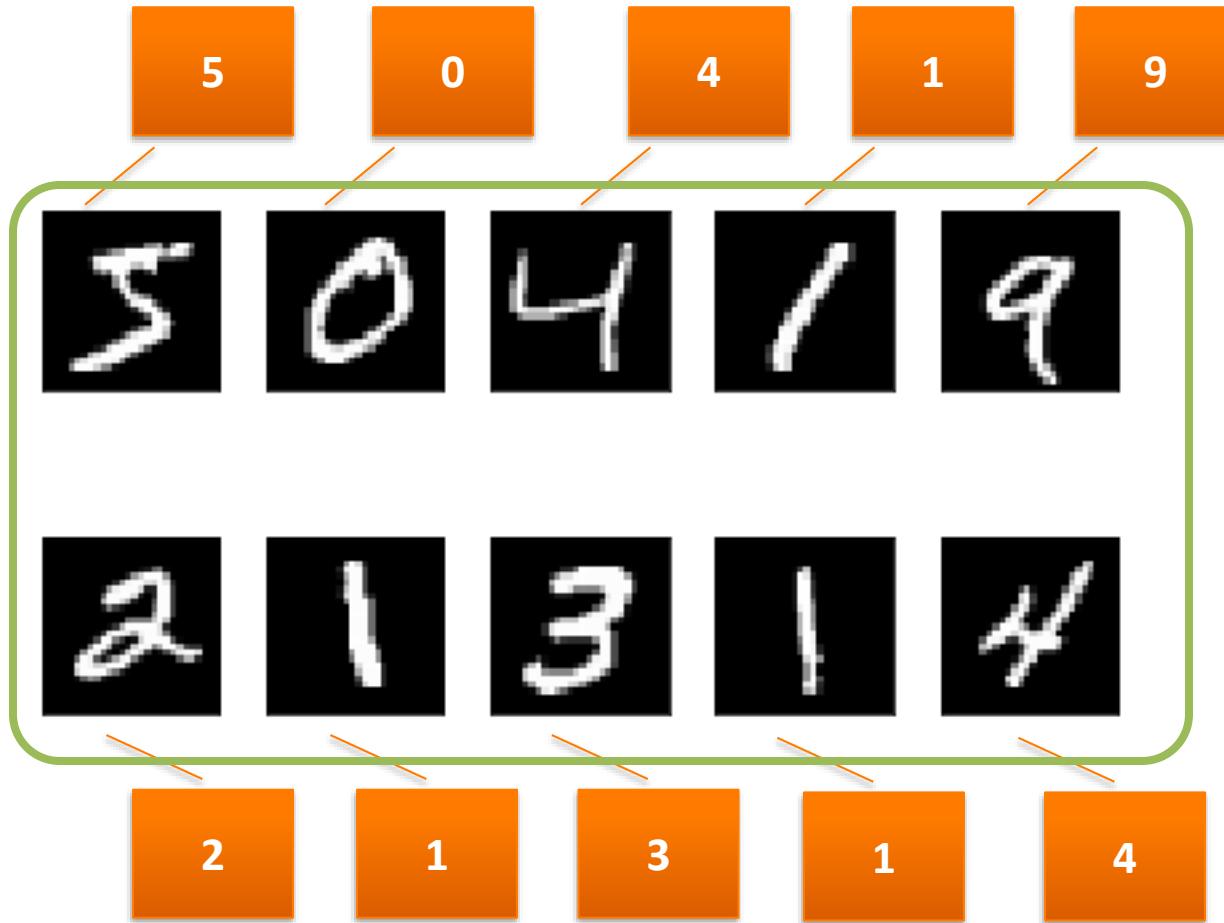


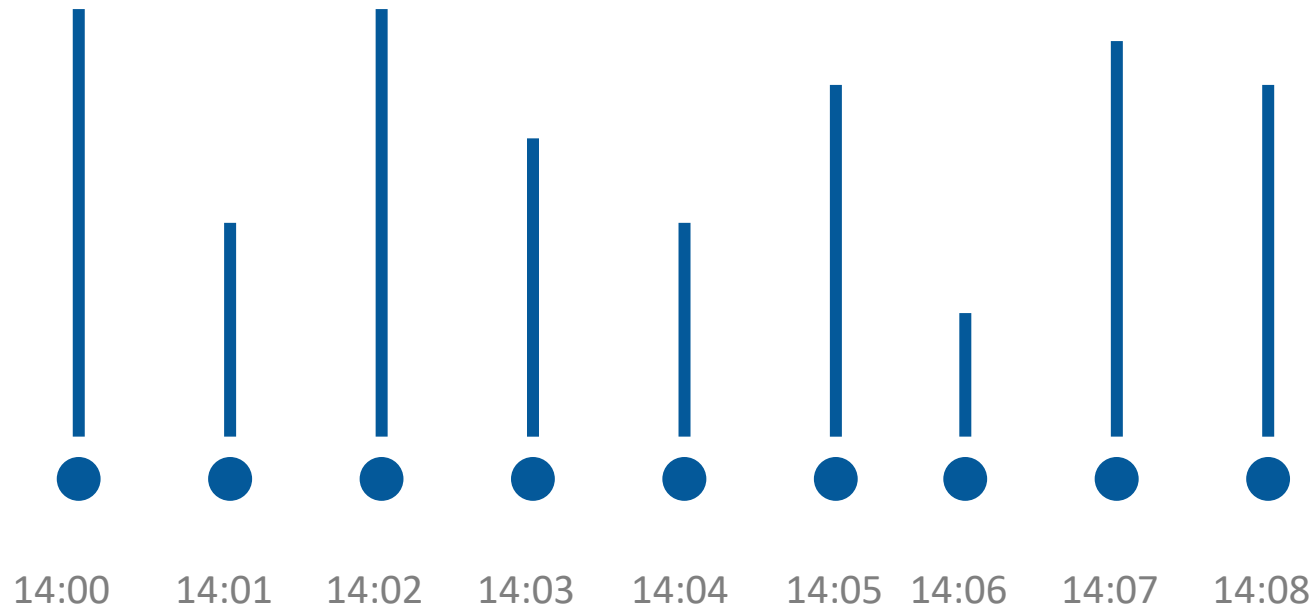
Image by [OpenClipart-Vectors](#) at [pixabay](#)

Grundbegriffe Zeitreihen und ML

Überwachtes Machine-Learning



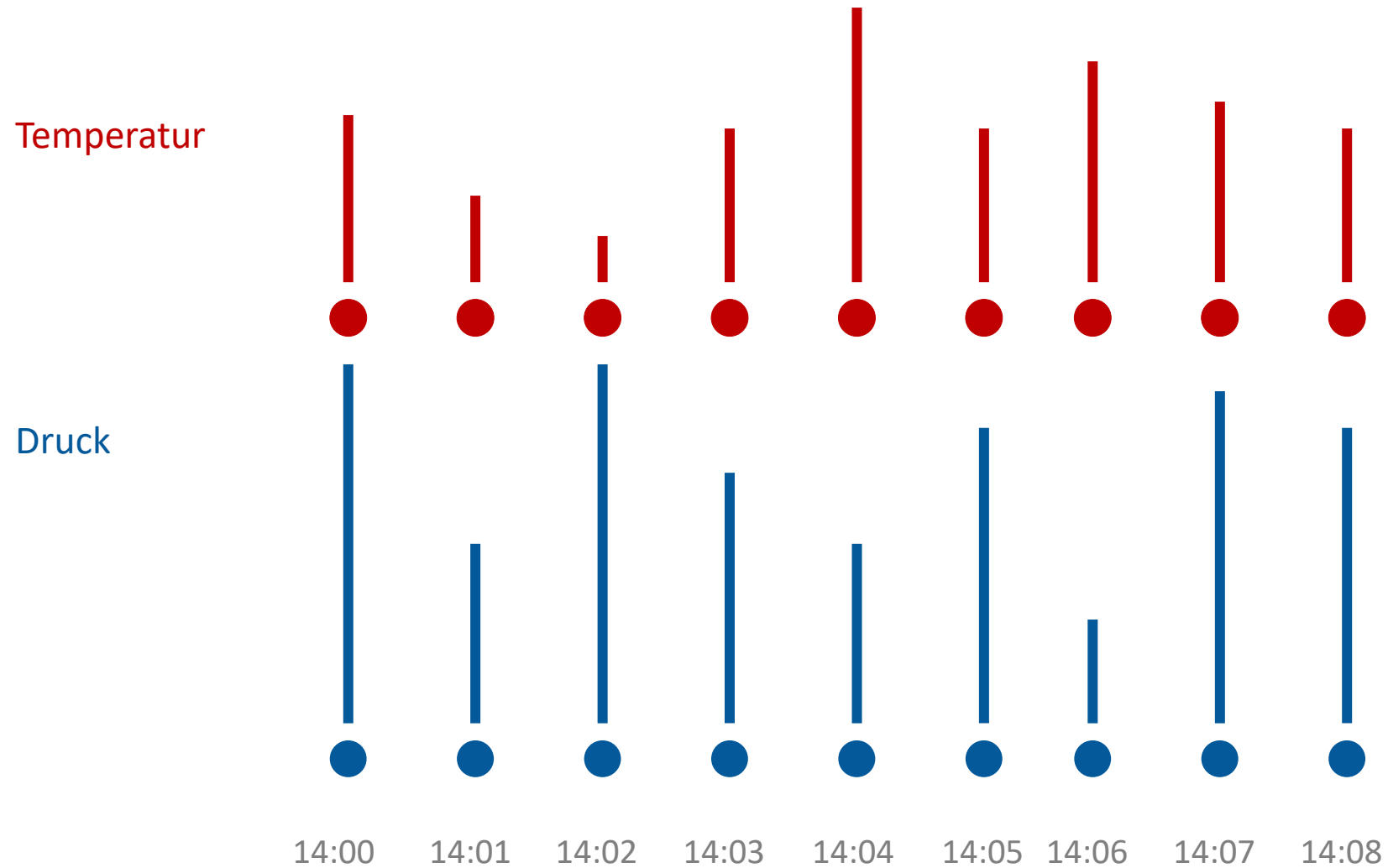
- Fixe Eingabegrößen
 - „tabellarisches Format“ bzw. Bildabmessungen
- Zielwerte als Ausgaben
 - Regression -> numerisch
 - Klassifikation -> kategoriell



Beispiele

- Messgrößen eines Fertigungsprozesses (Temperatur, Druck, Einsinktiefen, Zerspanungskräfte ...)
- Absatzzahlen
- Stromverbräuche / Einspeisung PV
- Finanzdaten (Aktienkurse)
- Klima / Meteorologie
- ...

Multivariate Zeitreihen



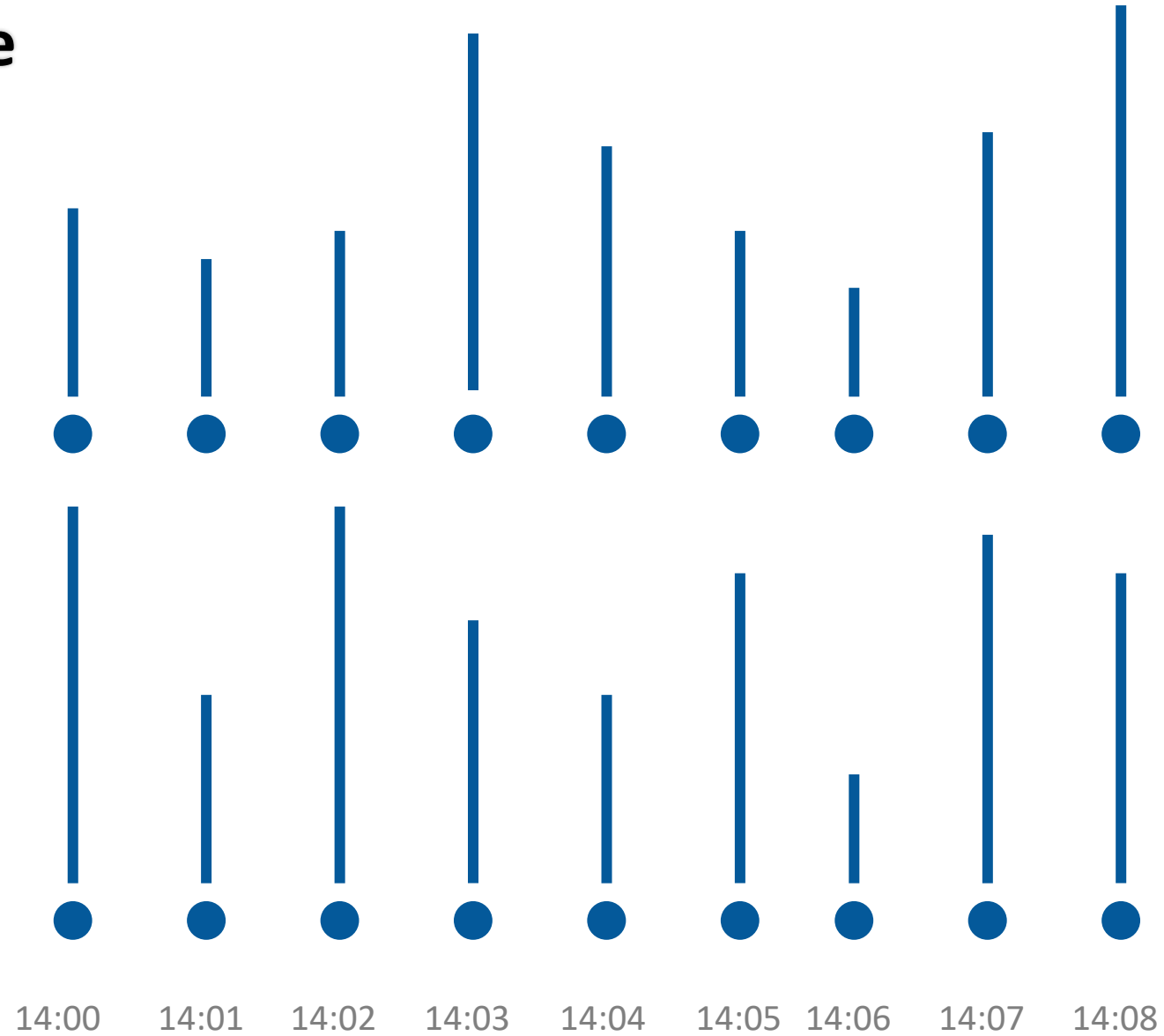
Korreliert,
unterschiedliche
Messgrößen, gleiche
Prozessinstanz

Panel-Datensätze



25.06.

24.06.



Unabhängige
Instanzen des
gleichen Prozess

Multivariate Panel-Datensätze



25.06.

Temperatur

24.06.

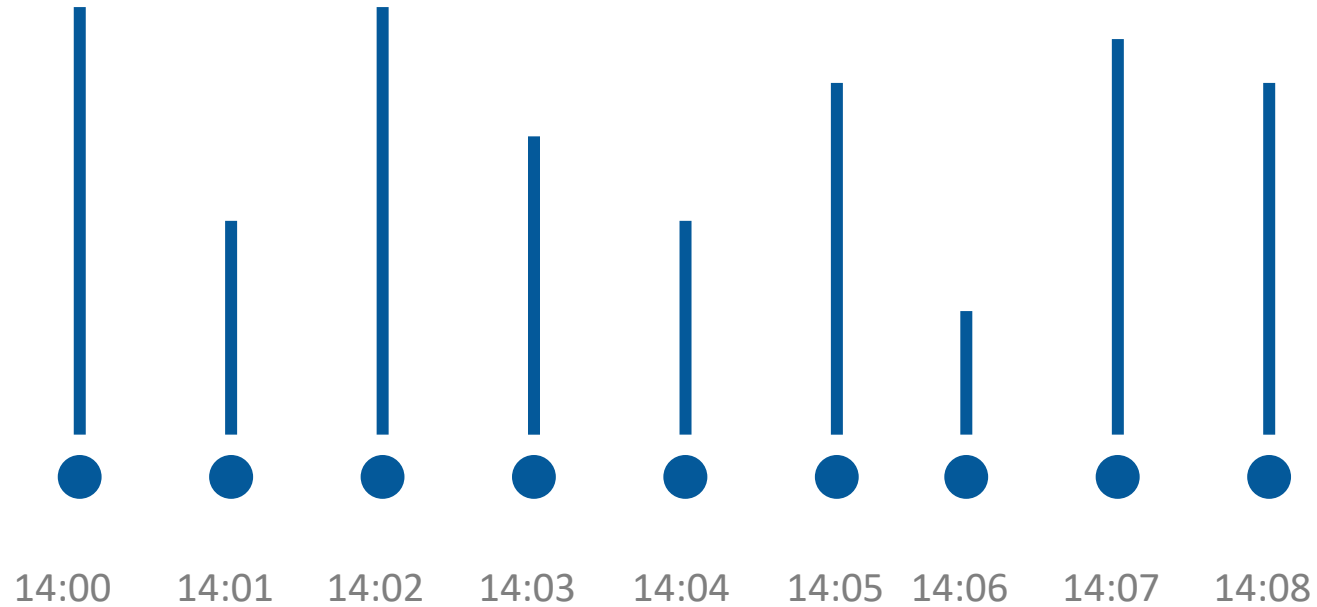
Druck

14:00 14:01 14:02 14:03 14:04 14:05 14:06 14:07 14:08

Typische Herausforderungen: Fehlende Messwerte

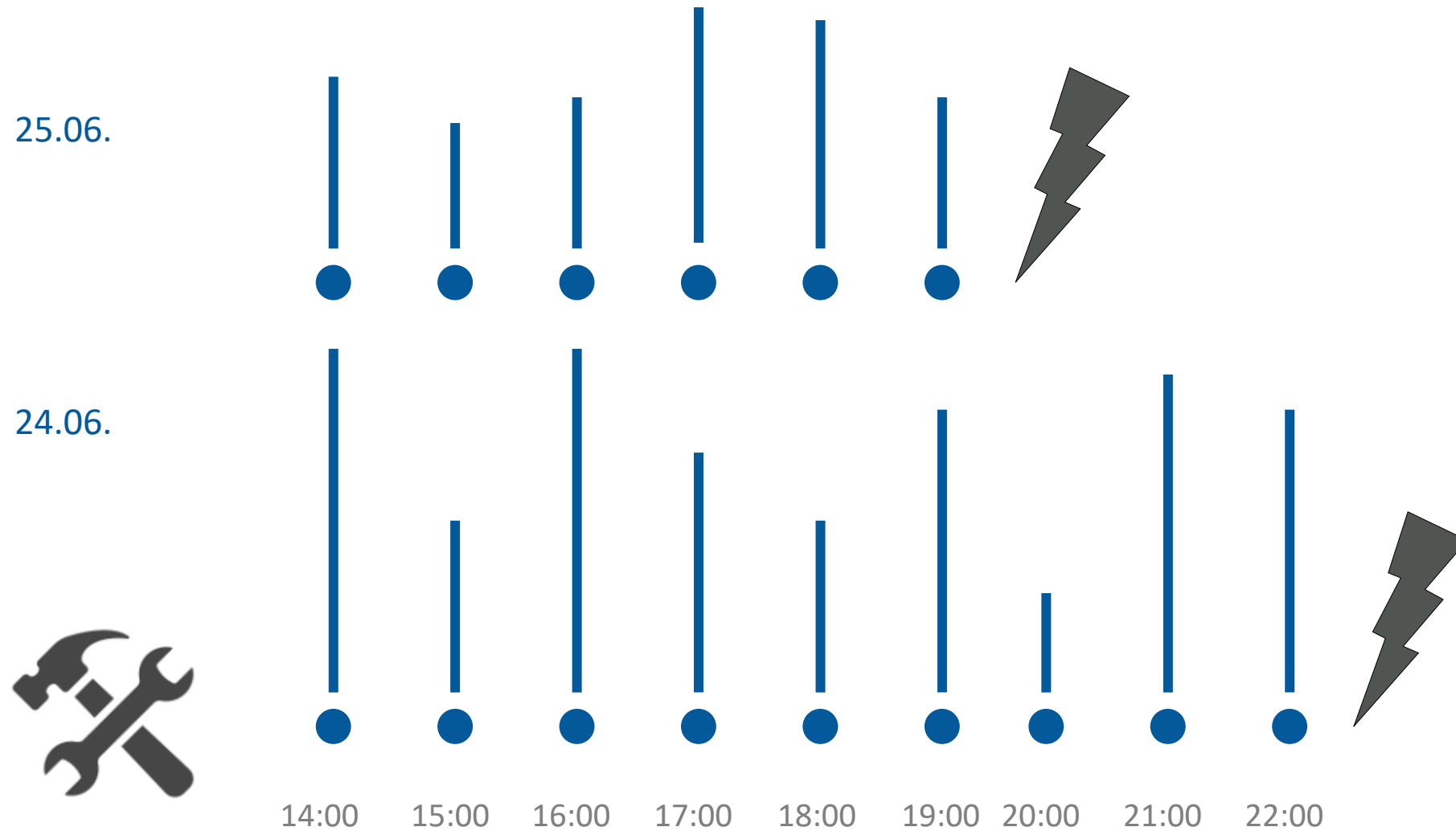


Aufzeichnung nur
bei Ereignissen



Unterschiedliche
Sensoraufösungen

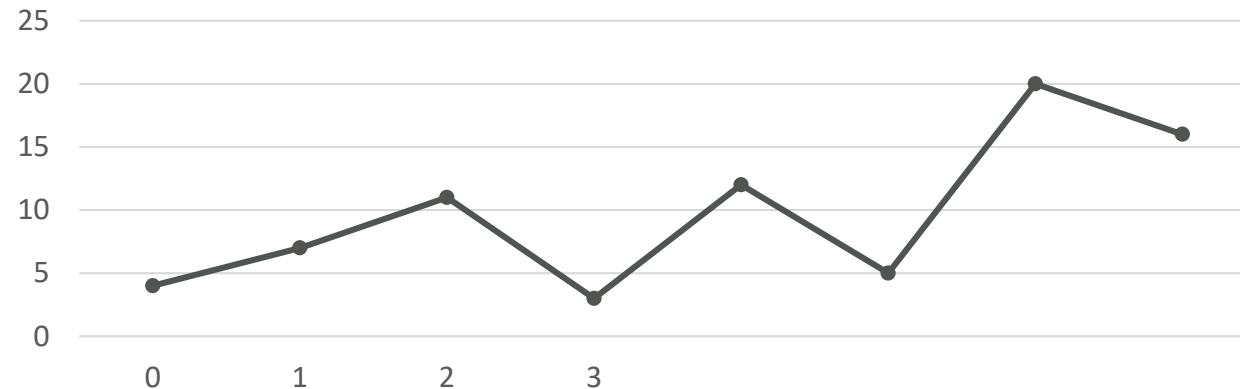
Typische Herausforderungen: Unterschiedliche Längen



Beispiel: Zeitreihen in Sktime laden

```
# sktime uses data in dimensions (m, d, t) where
# m ... is the number of distinct instances in the panel
# d ... is the number of features/variables that are measured
# t ... is the number of time steps

test_series = np.array([4, 7, 11, 3, 12, 5, 20, 16])
test_series = np.reshape(test_series, (1, 1, len(test_series)))
X = from_3d_numpy_to_nested(test_series)
```



 Open in Colab

https://www.sktime.org/en/latest/examples/loading_data.html

Lernaufgaben auf Zeitreihen

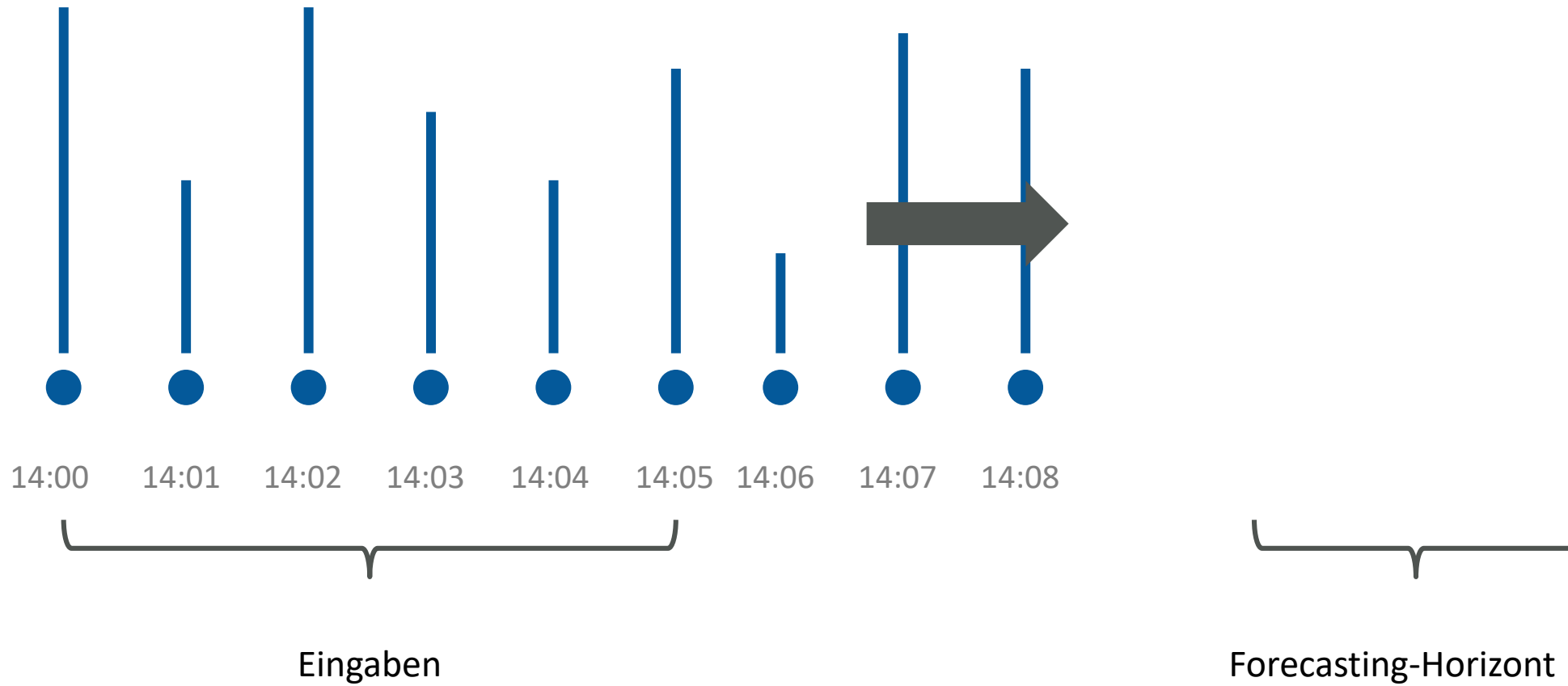
Forecasting



Absatzprognose

Stromverbrauch

Temperaturentwicklung



Klassifikation

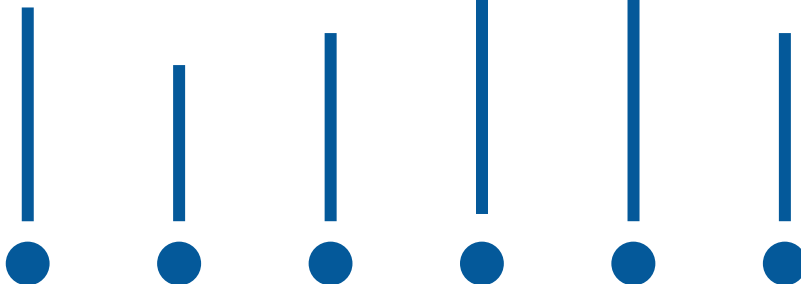


Fehlerklassifikation

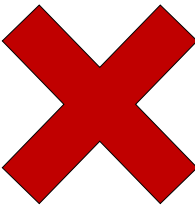
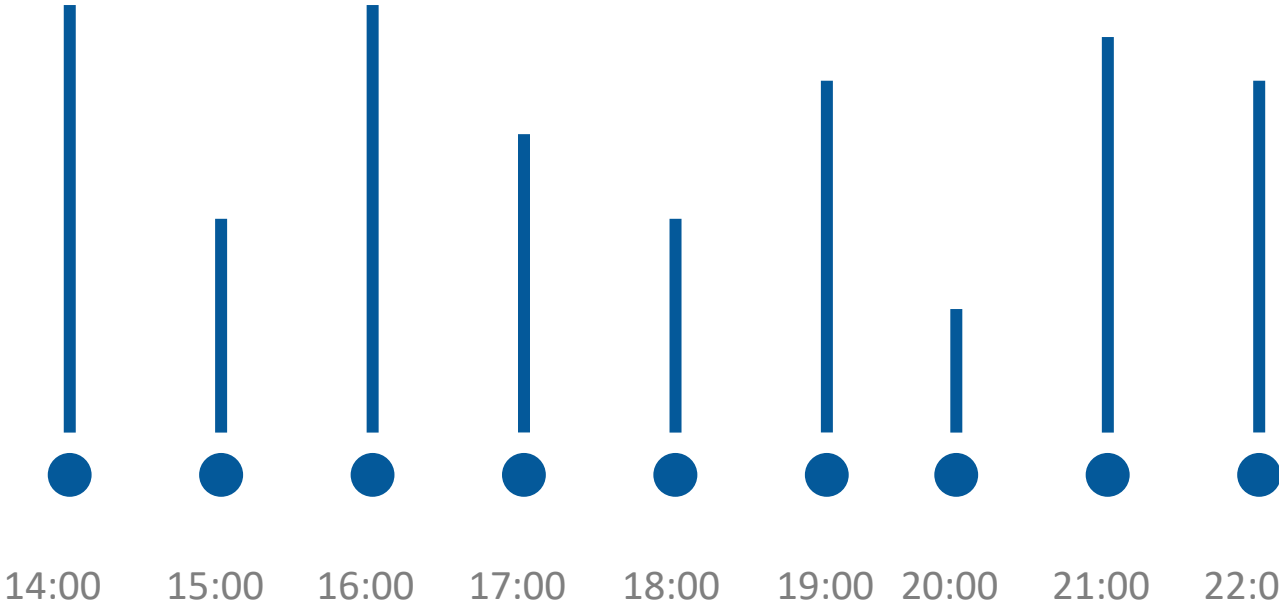
Wartung ja/nein

Aktivitätserkennung

25.06.



24.06.



Regression



25.06.

Lebensdauer

Qualitätseinstufung

Kosten

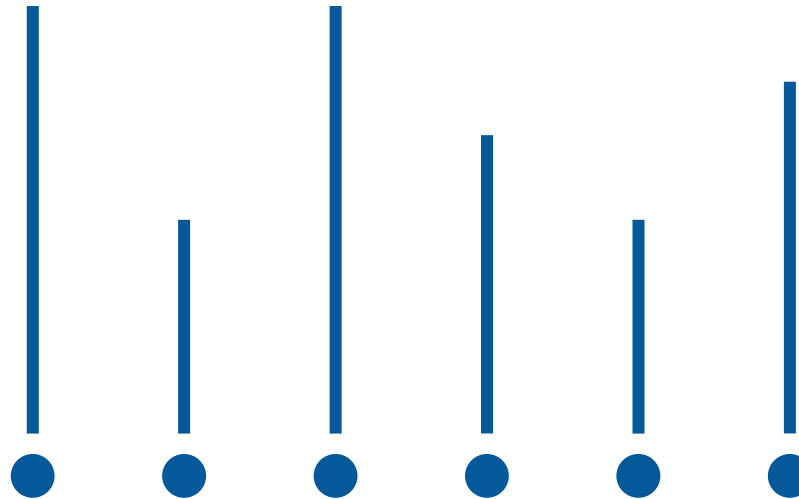


Noch

150

Schweißdurchgänge

24.06.



Noch

250

Schweißdurchgänge



14:00

15:00

16:00

17:00

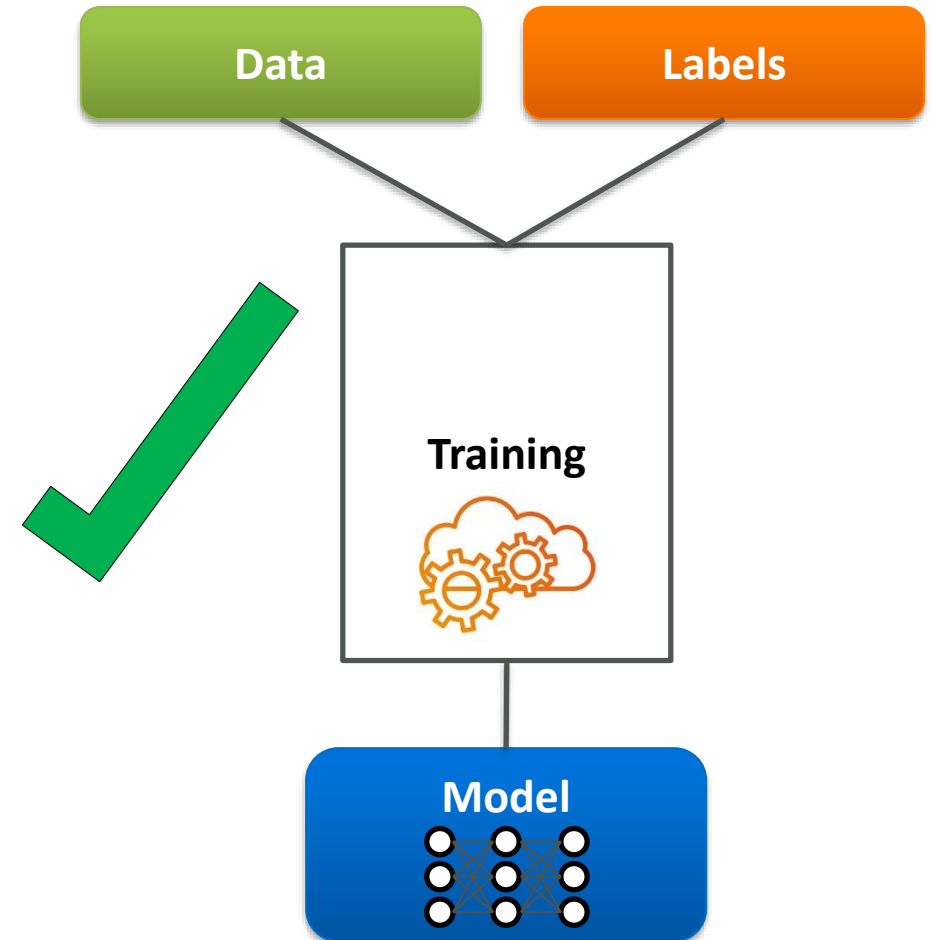
18:00

19:00

Transformation von Zeitreihen: Naive Kodierung

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
4	7	11	3	12	5	20	16	8
14:00	14:01	14:02	14:03	14:04	14:05	14:06	14:07	14:08

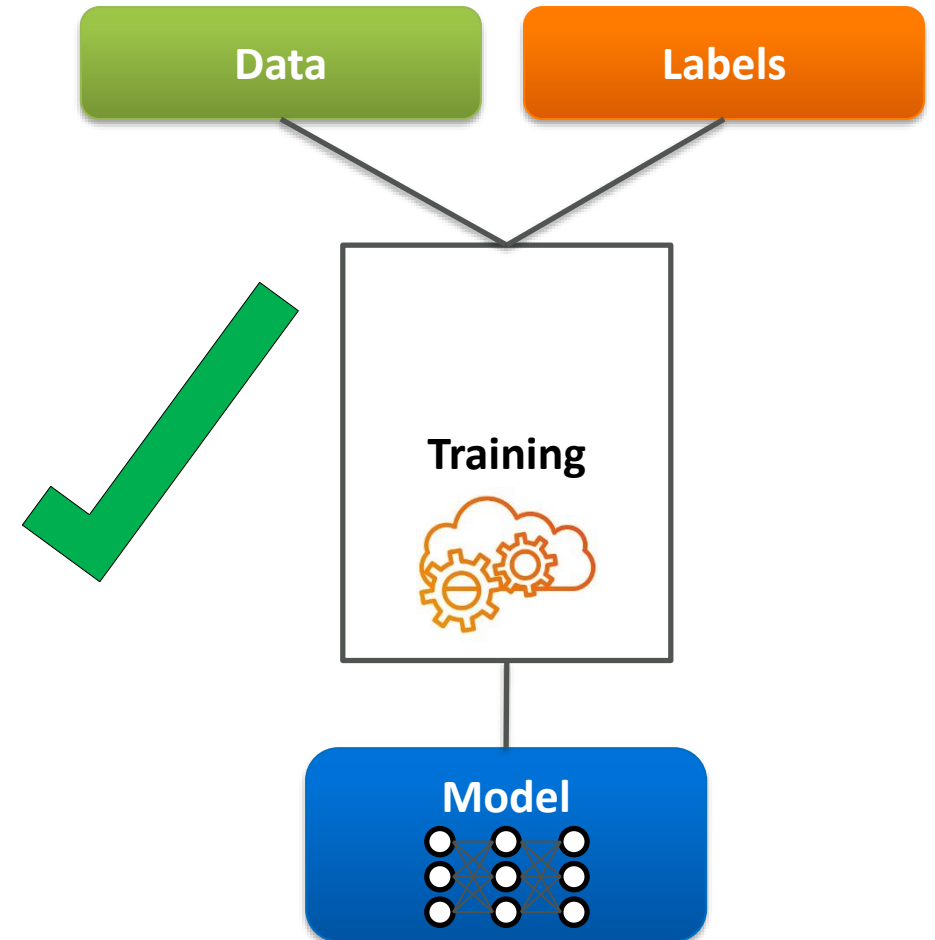
- Kann funktionieren.
- Erfordert gleiche Zeitreihenlänge
- Nutzt Reihenfolge nur bedingt



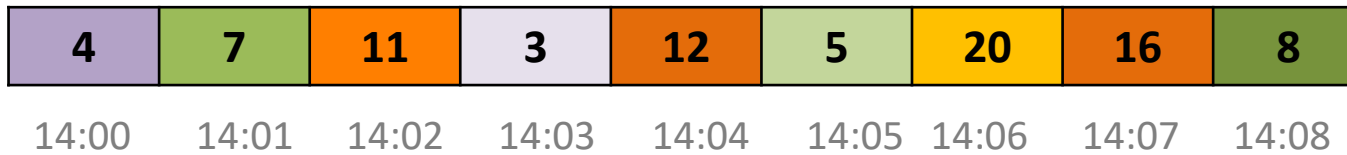
Transformation von Zeitreihen: Lagged Features



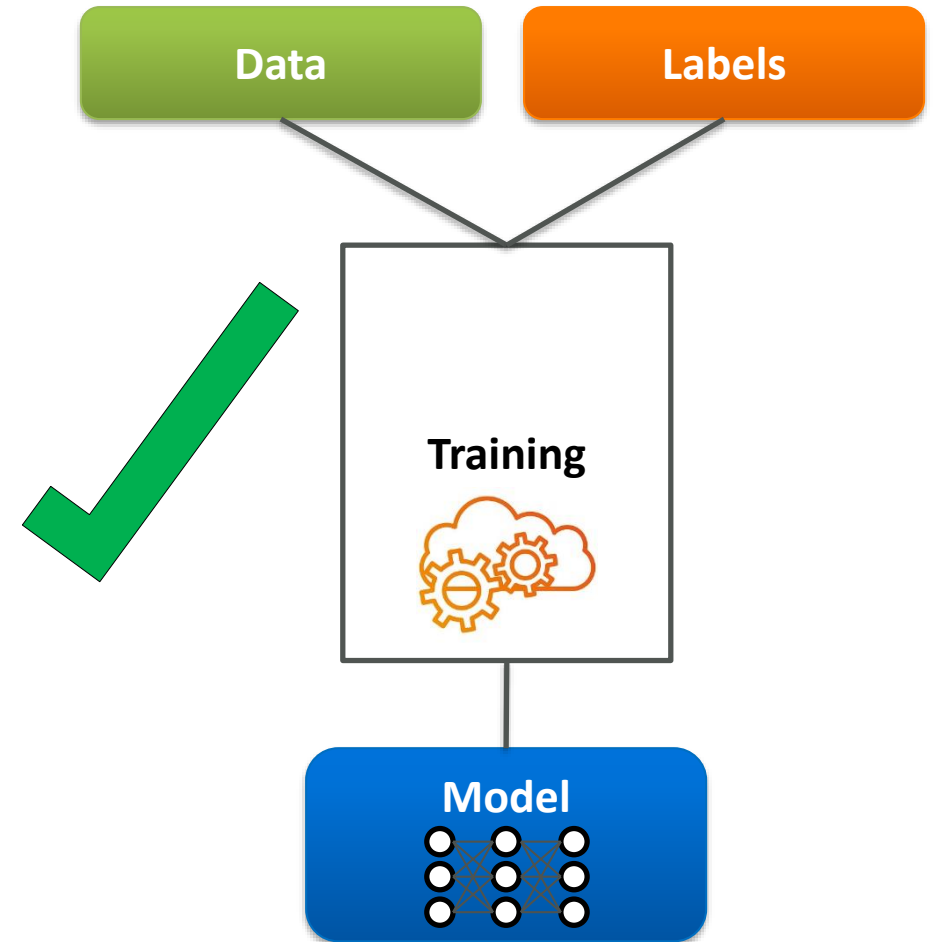
4	7	11	3	12
7	11	3	12	5
11	3	12	5	20
3	12	5	20	16
12	5	20	16	8



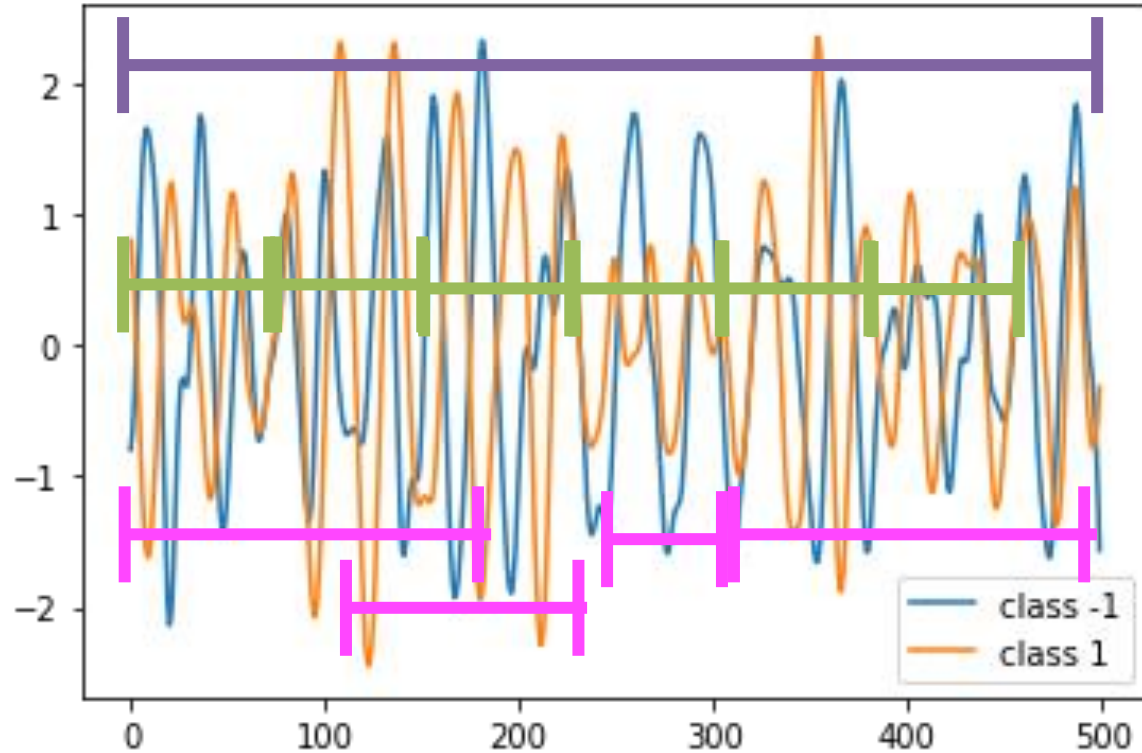
Transformation von Zeitreihen: Feature-Extraktion



Länge	Min	Max	Median
9	4	20	8



Wie können wir Zeitreihen und ML zusammen bringen?



FordA Datensatz von UCR

- Motorenlärm aufgezeichnet durch Sensor
- Klasse 1 *OK*, Klasse -1 *nicht OK*

■ Globale Features

- Min, Max
- Durchschnitt, Median
- Transformationen (z.B. Fourier)

■ Fenster

- gleiche aggregierte Features

■ Unregelmäßige Intervalle

Sktime-Demo



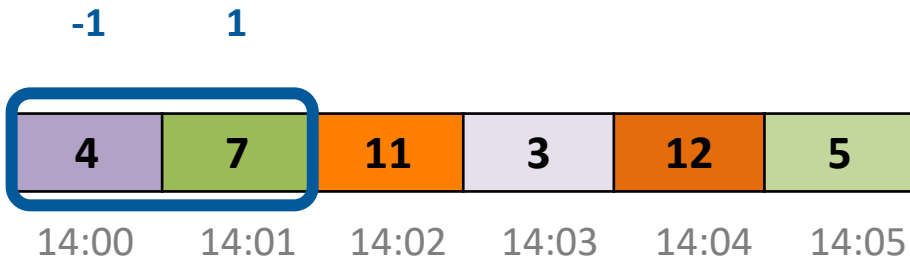
Faltungsmodelle

Diskrete Faltungen (Convolutions) als Merkmalsextraktoren

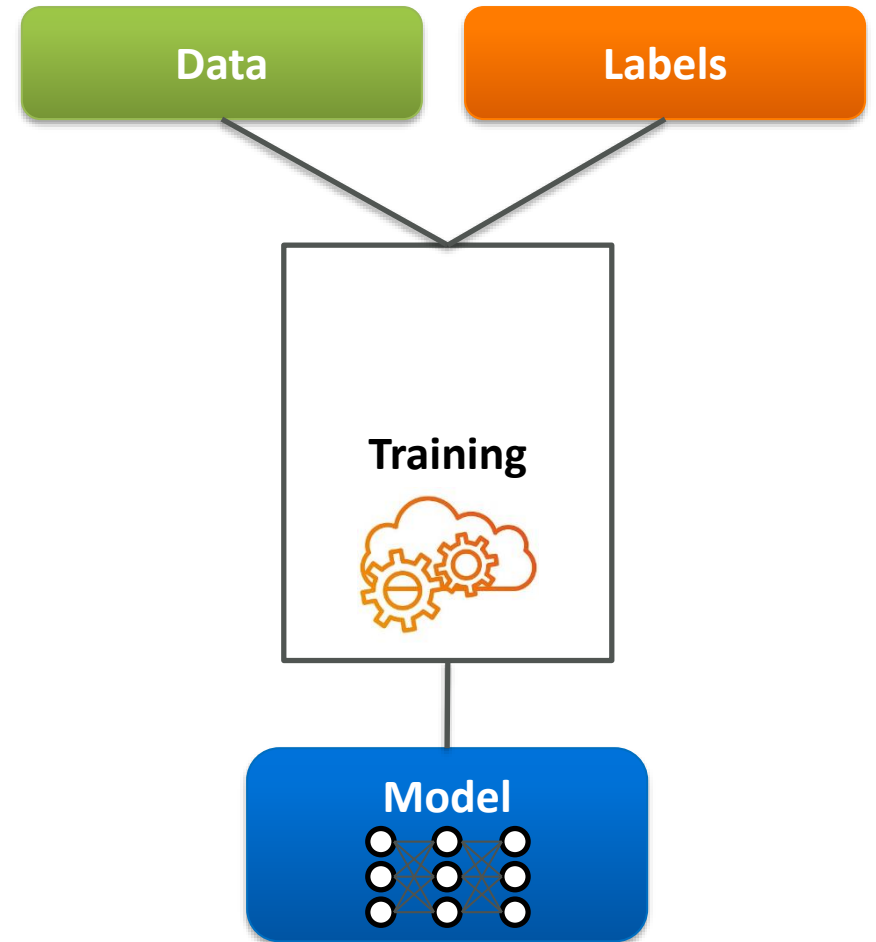


Filterkern

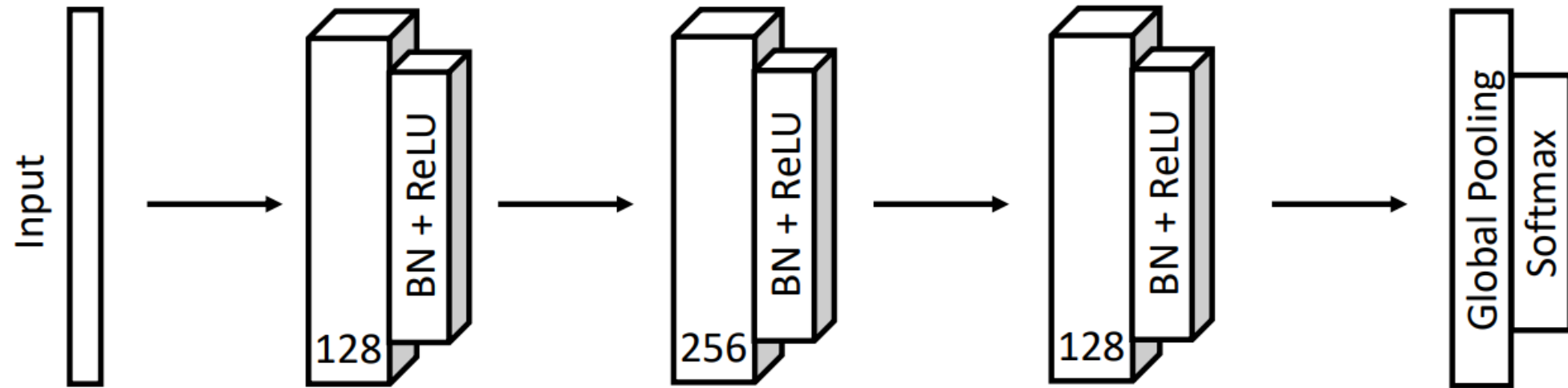
-1	1
----	---



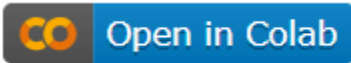
3	4	-8	9	-7
---	---	----	---	----



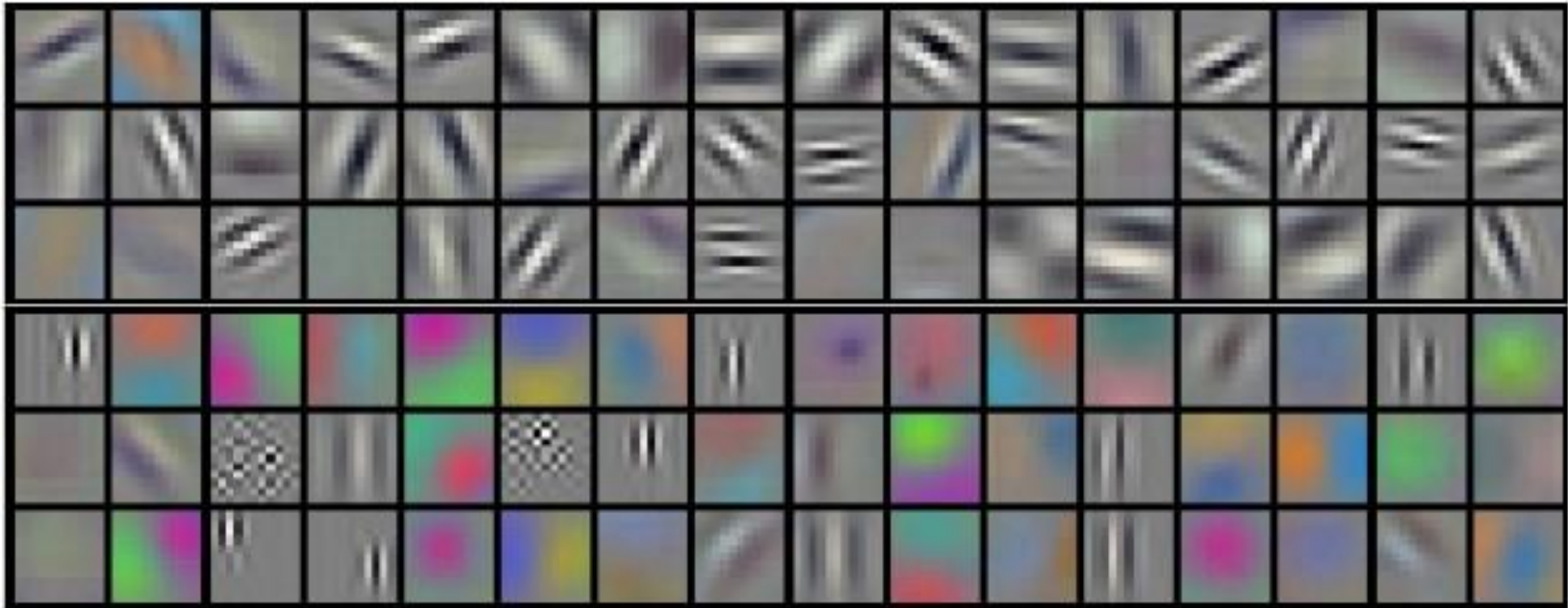
Typischer Einsatz in neuronalen Netzen



Time series classification from scratch with deep neural networks: A strong baseline; Wang, Zhiguang and Yan, Weizhong and Oates, Tim; 2017 IJCNN



Beispiel für trainierte Faltungskerne



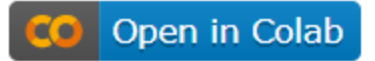
[Krizhevsky et al., 2012]

Rocket-Algorithmus (RandOm Convolutional KErnel Transform)



- 108 Datensätze in UCR:
 - Rocket 2 Stunden
 - Mini-Rocket 8 Minuten
 - Stand der Technik 14 Stunden

Grundidee



1. Verwende zufällig gewählte Faltungskerne (Gewichte, Shift, usw)
2. Extrahiere Merkmale durch Faltung
3. Verwende genau 2 Pooling-Operationen über resultierende Reihe (pro Faltungskern)
 - Max-Wert
 - Proportion positiver Werte
4. Trainiere herkömmlichen Klassifizierer

ROCKET: Exceptionally fast and accurate time series classification using random convolutional kernels

Angus Dempster, François Petitjean, Geoffrey I. Webb

Tipps & Tricks

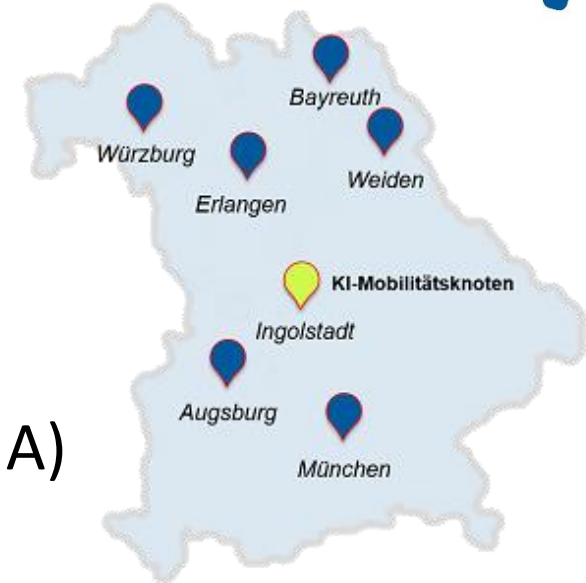
- Wichtige Ereignisse per Zeitstempel genau speichern
 - Defekt, planmäßige Wartung, etc.
 - Möglichst ergonomisch für Mitarbeitende (Smartwatch o.ä.)
- Einordnung des richtigen Zeitreihenproblems ist wichtig
 - Forecasting am weitestend verbreitet, aber
 - Prädiktive Wartung aber eher Regression / Klassifikation
- Fangen Sie früh mit der Datenaufzeichnung an
 - Vorlauf mehrere Monate vor KI / Data Science Projekt

Key facts „Almotion Bavaria“ aimotion.de

Bayerischer KI-Mobilitätsknoten an der TH



Bayerische Staatsregierung



- 20 neue Forschungsprofessuren (halbes Lehrdeputat + 1 wiss. MA)
 - 10 aus der Hightech-Agenda Bayern <https://aimotion.de>
 - 10 aus Stiftungsmitteln (u.a. Audi, Stadt Ingolstadt, Klinikum, etc.) <https://ainin.de/>
 - *Schwerpunkte*: Autonome Mobilität (Fahren, Fliegen), **KI-gestützte Produktion**
 - *Zielsetzung*: Aufbau von bis zu 120 Wissenschaftler:innen durch Drittmittelforschung
- Geplant: Einzug in den Digitalbau (Kavalier Dalwigk)
 - 4.000 m² Fläche, Seminarräume, KI-Labore
 - Ausbau Informatik, KI im Mobilitätsknoten

