

Mittelstand-Digital
**Zentrum
Hannover**

Anomalieerkennung mit KI

Grundlagen und Methoden

Uwe Hadler

Mittelstand-
Digital 

Gefördert durch:

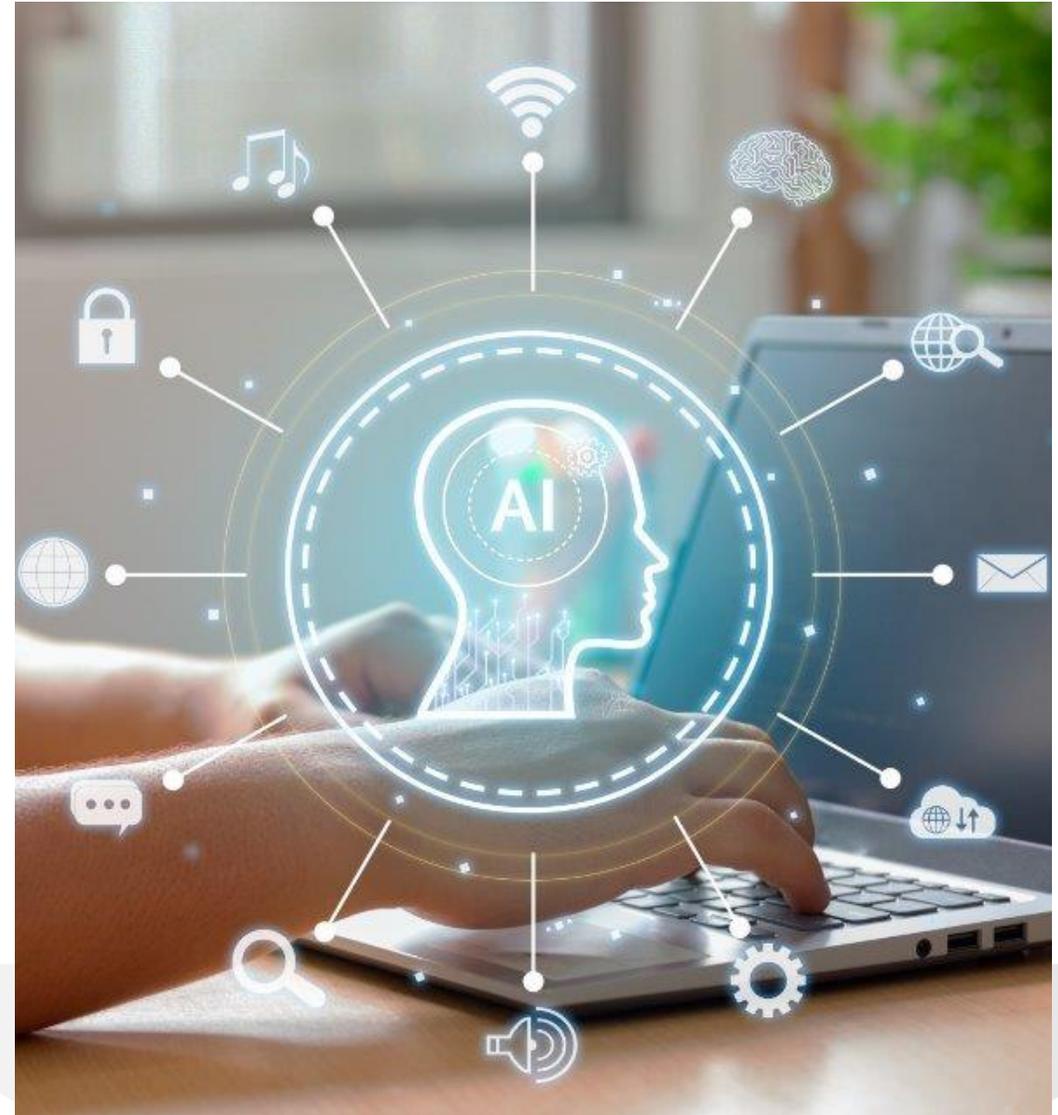


Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

Agenda

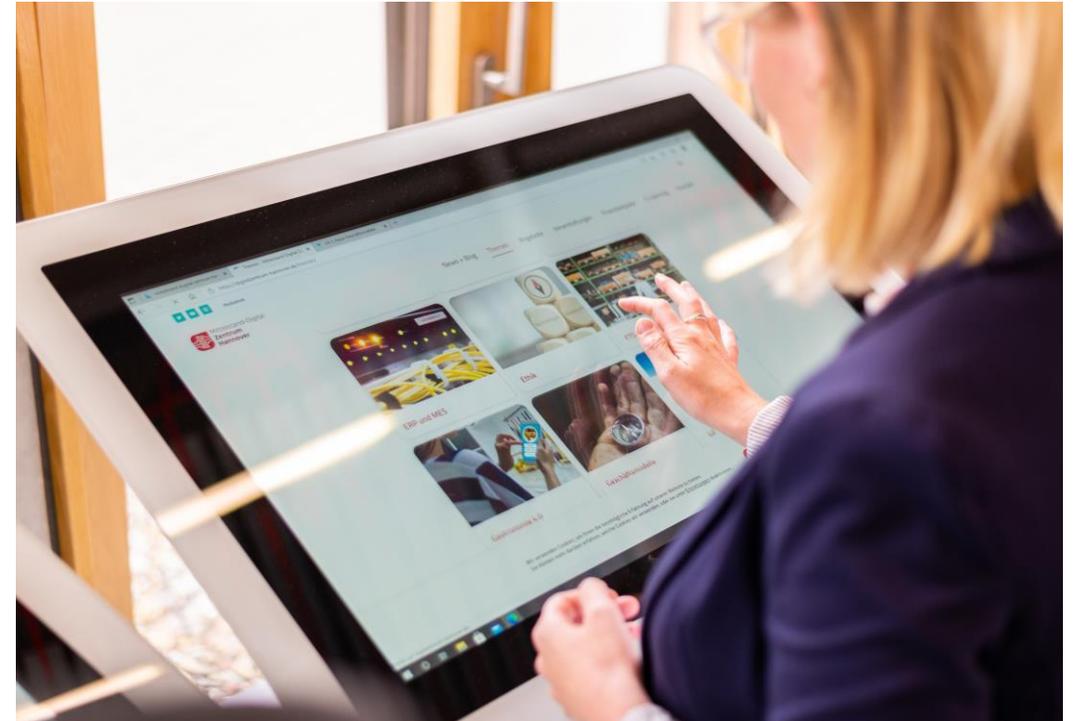
- TOP 1 **Einführung**
- TOP 2 Definition von **Anomalien**
- TOP 3 Anomalieerkennung: **Statistisch**
- TOP 4 **KI-basierte** Anomalieerkennung
Mit und ohne gelabelte Daten
- TOP 5 **Diskussion**



Was wir Ihnen bieten

- Technologie- und Wissenstransfer
- Unterstützung beim Einsatz von digitalen Lösungen und KI-Anwendungen im Betrieb
- Flächendeckende Angebote in Niedersachsen
- Branchen: v. a. Produktion und Handwerk
- Kostenfreie Angebote für alle Unternehmen

individuell, unabhängig, vor Ort



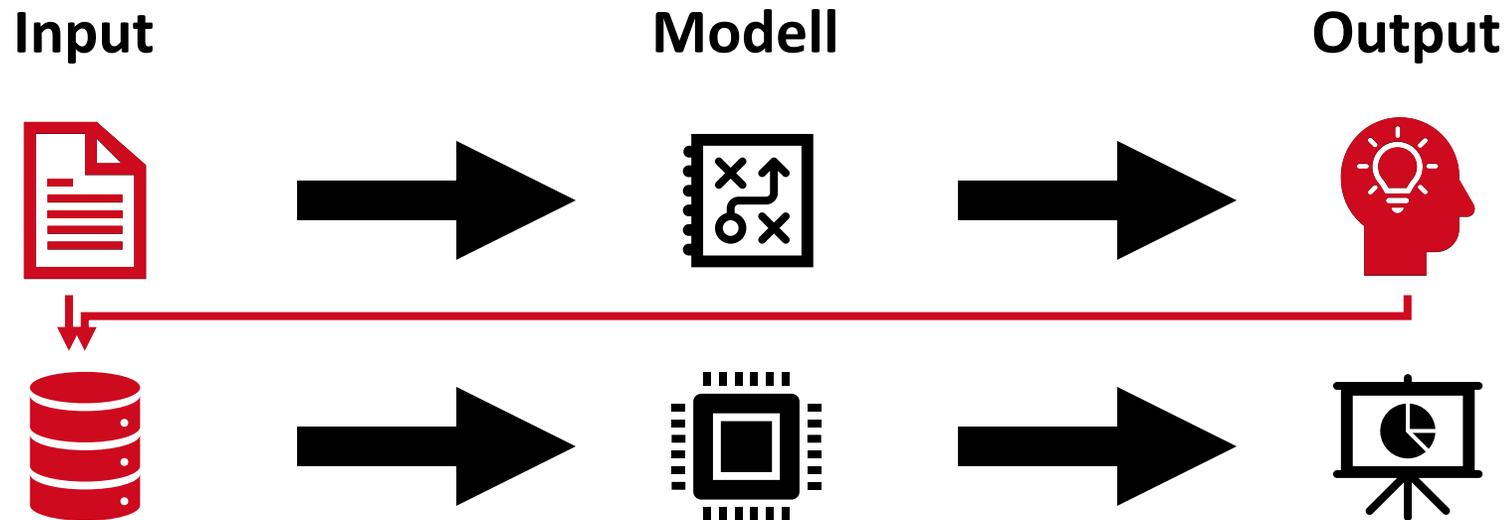
Quelle: Mittelstand-Digital Zentrum Hannover



Was ist **Künstliche Intelligenz**?



Was ist **Künstliche Intelligenz**?



- **Künstliche Intelligenz** bezeichnet die Fähigkeit eines Systems, **aus Daten zu lernen**
- Menschen lernen aus **Erfahrung**, durch Lebenserfahrung, Ausbildung, ...
- Trainingsdaten sind **einzig** Erfahrungsschatz einer KI (*und meist anwendungsspezifisch!*)
- Grundlage: komplexe **statistische** Verfahren, **neuronale Netze** (angelehnt an Menschen)

Source: LightFieldStudios / envato elements



Was ist eine **Anomalie**?

Was ist eine **Anomalie**?

Anomalien (oder Ausreißer) sind Datenpunkte, die sich **signifikant von der Mehrheit** der Daten **unterscheiden**.

Wie wird das **definiert**?

$$\mathbb{P}(x \in Normal) \approx 0$$

- Was ist "normal"?
- Wie können wir das testen?

Was können wir aus Anomalien **schließen**?

Anomalien sind ein Indikator für...

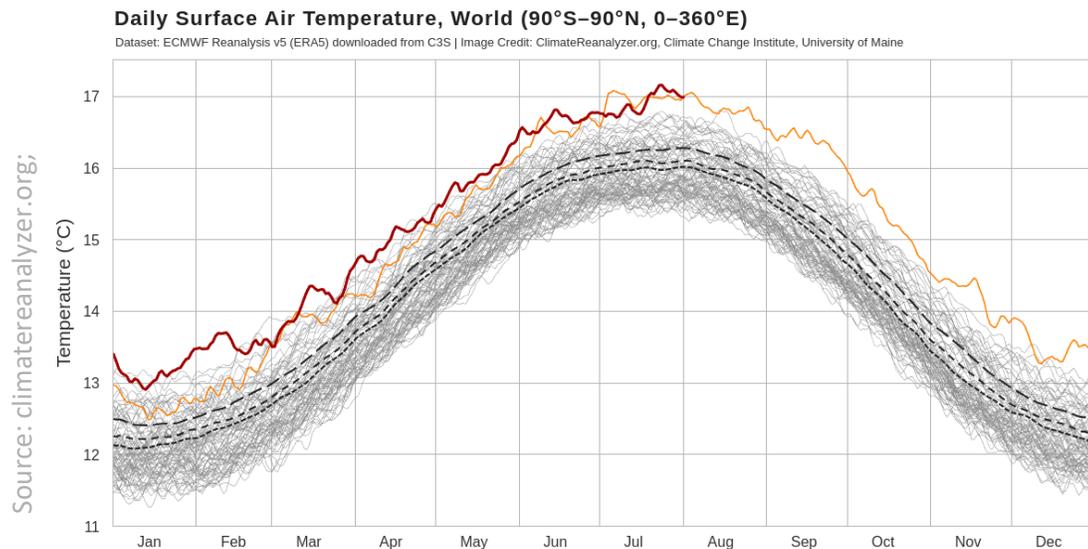
- Fehler
- Neue Aktivitäten
- Betrügerische Aktivitäten

In diesem Vortrag...

Wie können wir Anomalien erkennen? ✓

Was machen wir, wenn eine Anomalie gefunden wurde? ✗

Fokus auf Zeitreihendaten



Warum ist Anomalieerkennung dort **schwer**?

Erkennung von Ausreißern in Zeitreihen ist machbar, Telemetriedaten sind allerdings meistens...

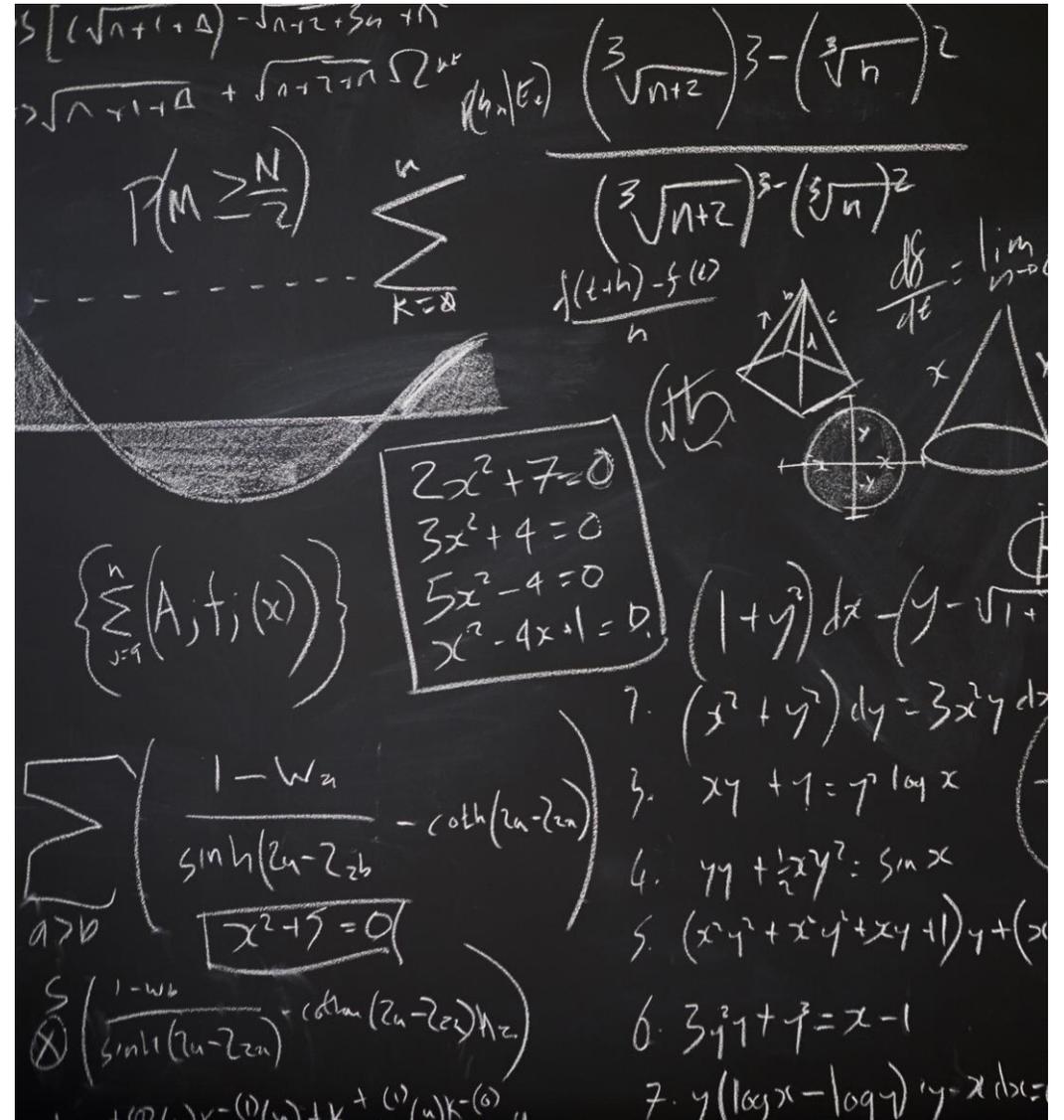
- **Mehrdimensional**
- **Korreliert**
- **multimodal**

... und bilden sehr komplexe Muster ab.



Anomalieerkennung

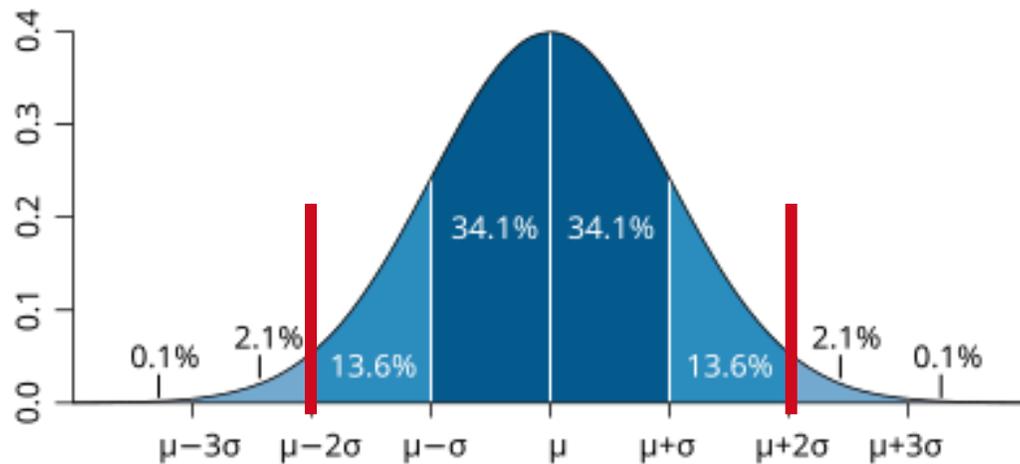
Traditionelle Methoden



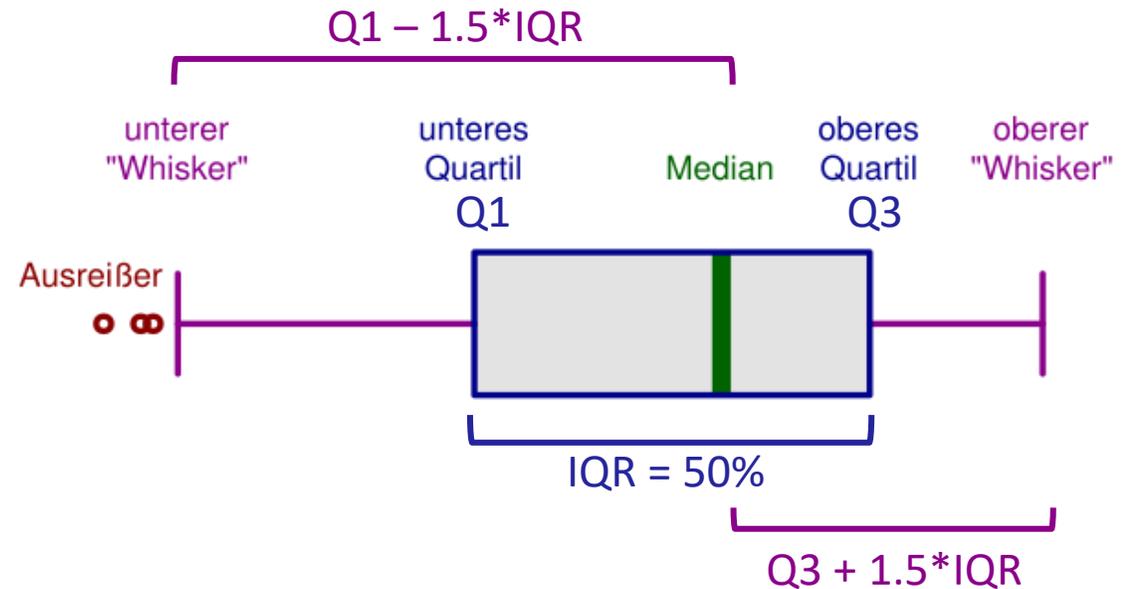
Schwellenwerte – Z-Werte und IQR

Unter der Annahme, dass die zu untersuchenden Daten aus einer konsistenten Verteilung stammen, können manuell gewählte oder statistisch berechnete Schwellenwerte genutzt werden, um Ausreißer zu erkennen.

Z-Werte: Entfernung zum Mittelwert in Standardabweichung



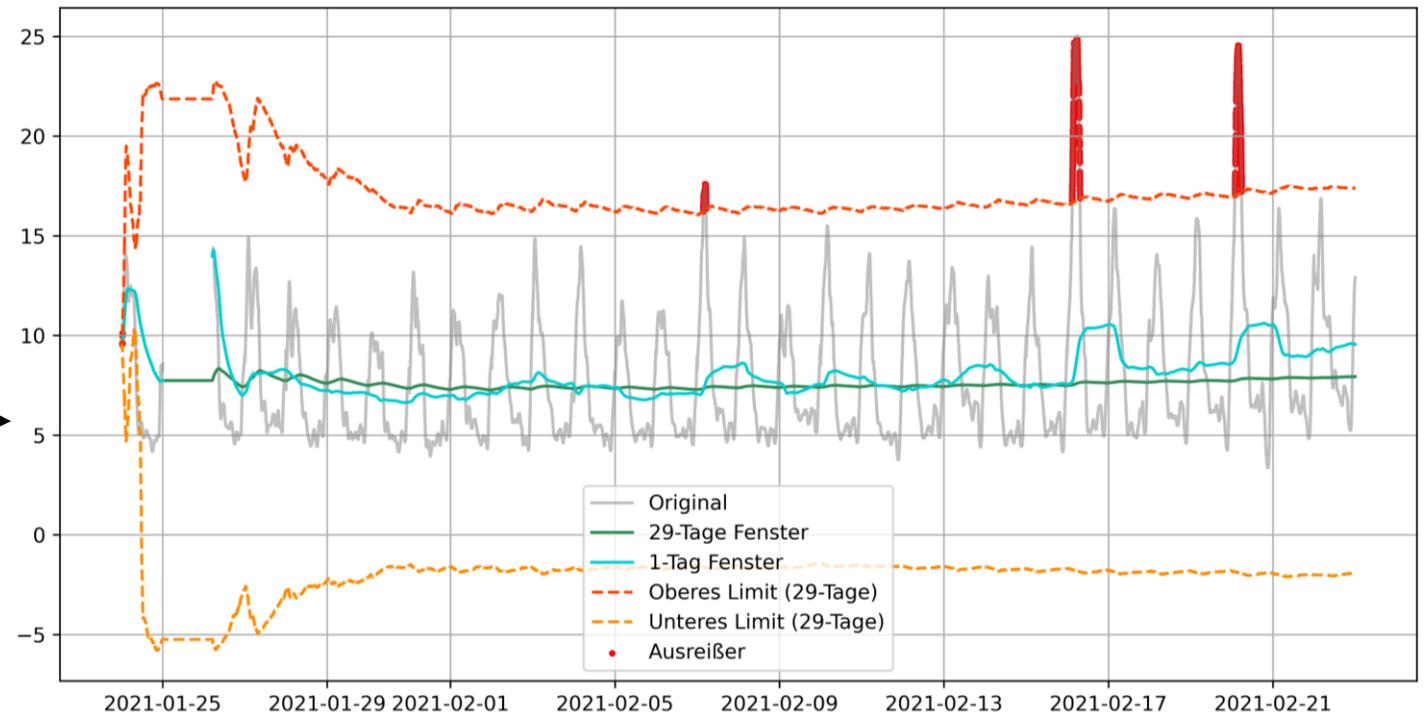
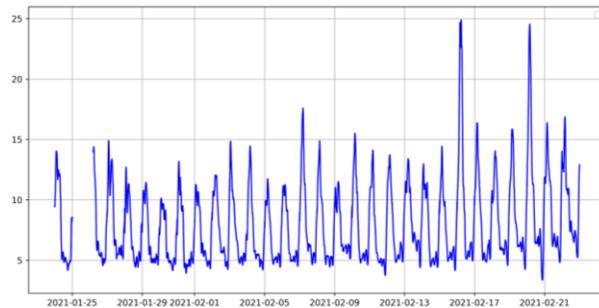
Interquartilsabstand (IQR): außerhalb des oberen und unteren Whiskers



Source: Wikimedia Commons

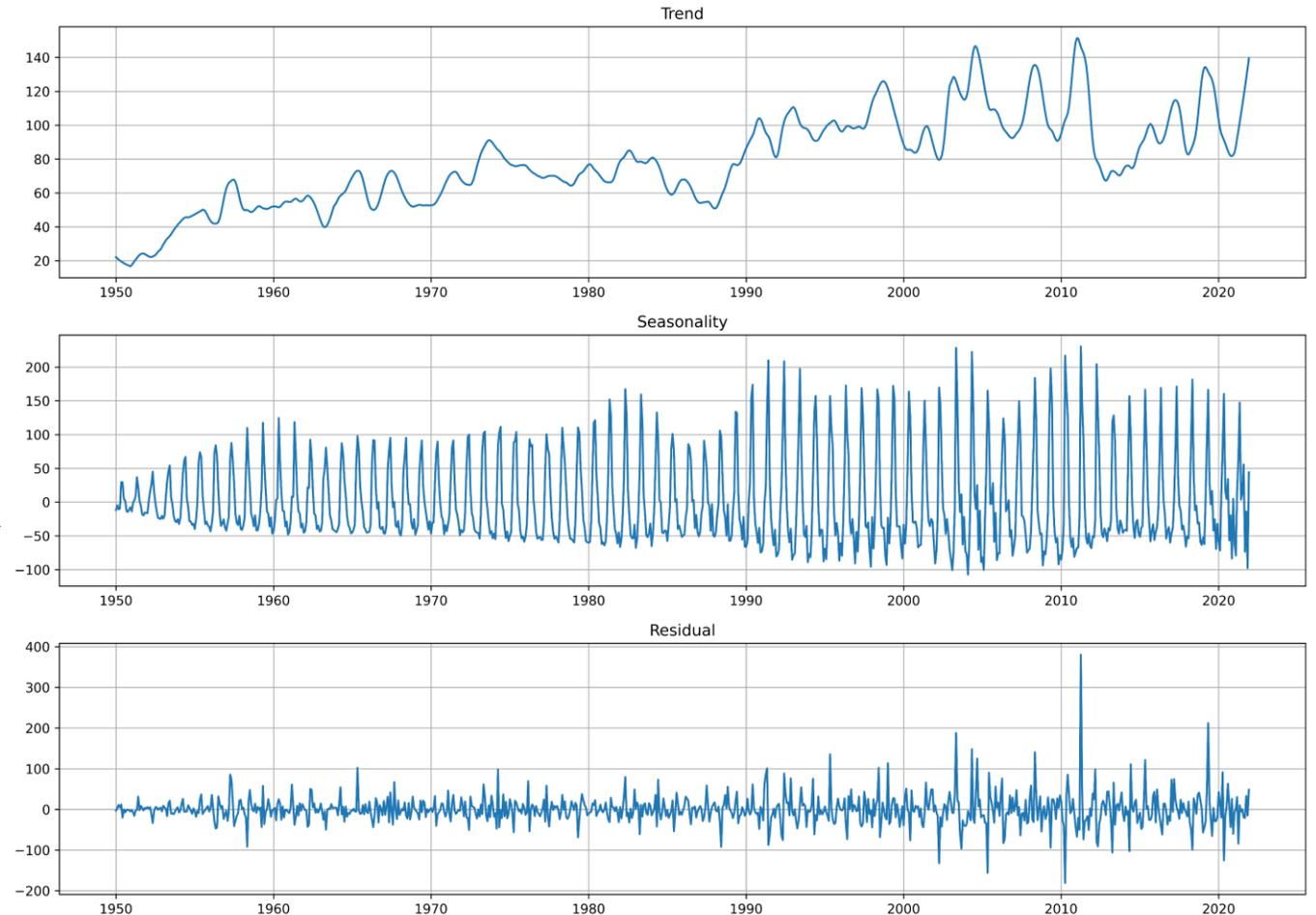
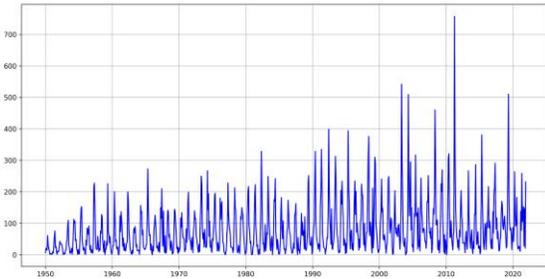
Anomalieerkennung in **Zeitreihendaten** mit statistischen Methoden

In Zeitreihendaten können einzelne Beobachtungen oder Zeiträume als anomal erkannt werden. Zur Nutzung statistischer Methoden werden oft **fortlaufende Fenster** oder Dekompositionsverfahren eingesetzt.



Anomalieerkennung in **Zeitreihendaten** mit statistischen Methoden

In Zeitreihendaten können einzelne Beobachtungen oder Zeiträume als anomal erkannt werden. Zur Nutzung statistischer Methoden werden oft fortlaufende Fenster oder **Dekompositionsverfahren** eingesetzt.





Anomalieerkennung

KI-basierte Methoden



Vor- und Nachteile von **KI-** **basierten Methoden**



- **Skalierbarkeit** auf (sehr) große Datenmengen
- Lernt eigenständig aus **komplexen Mustern** in den Daten
- Kann **Interaktionen** zwischen Variablen abbilden und ausnutzen



- Benötigt große **Datenmengen**, um volles Potential auszuschöpfen
- Erhöhter **Zeit- und Ressourcenaufwand**
- Entscheidungsfindung nicht immer **interpretierbar**



Source: iLexx / envato elements



Überwachtes Lernen in der Anomalieerkennung

Überwachtes Lernen stellt Anomalieerkennung als **Klassifikationsproblem** dar. Aus historischen Daten wird ein Model von normalen und anomalen Beobachtungen gelernt, für neue Punkte wird eine Klassen vorhergesagt.

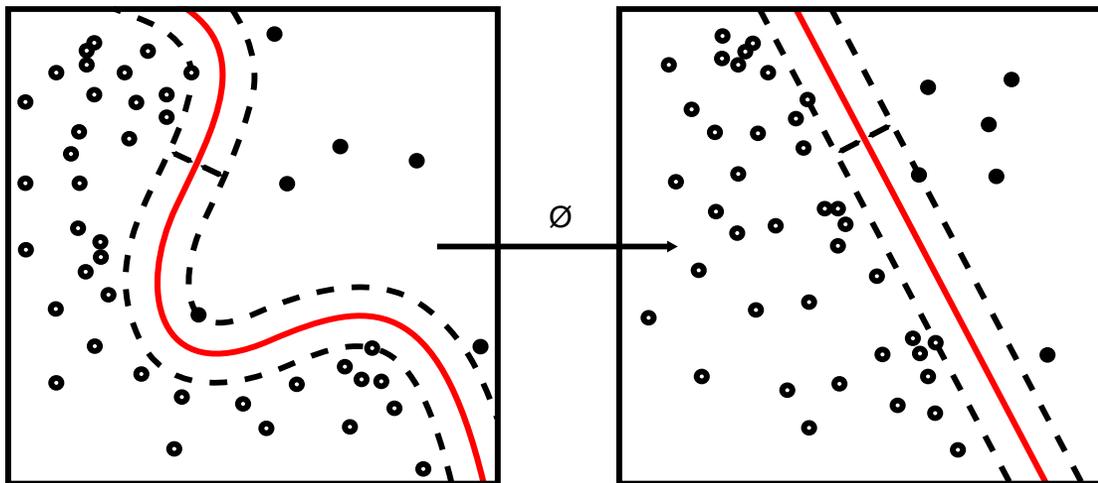
Was ist "normal"?

- Gelabelte Daten sind Voraussetzung für überwachtes Lernen
- Klares Verständnis im Voraus nötig
- Zeitaufwand zum Beschriften der Daten
- Label jeder Beobachtung im Datensatz stellt klar, ob normal oder anomal

ID	x1	x2	..	xN	Anomal
676f6f64	0
73686974	1
...	

Überwachtes Lernen in der Anomalieerkennung

Überwachtes Lernen stellt Anomalieerkennung als **Klassifikationsproblem** dar. Aus historischen Daten wird ein Model von normalen und anomalen Beobachtungen gelernt, für neue Punkte wird eine Klassen vorhergesagt.



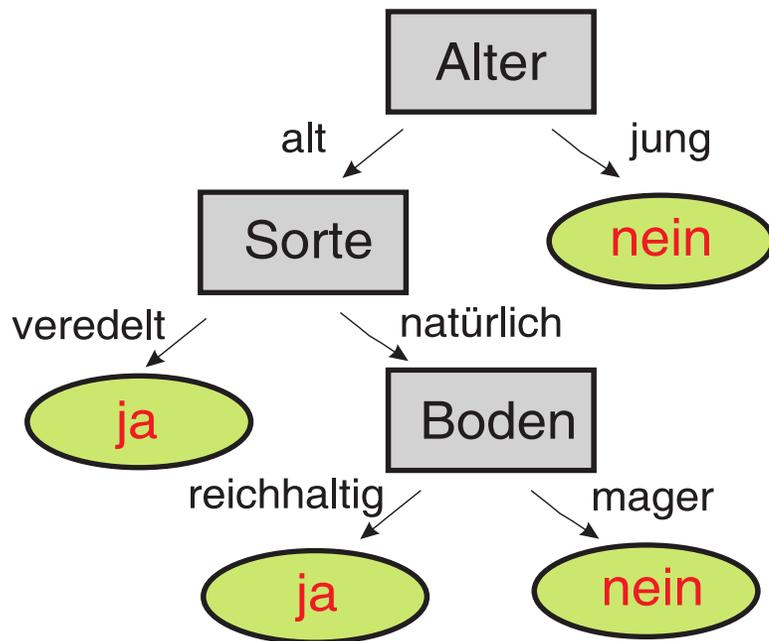
Wie können wir das testen?

Mit praktisch allen KI-Klassifikationsalgorithmen.
Typisch sind hier z.B. ...

- **Support Vector Machine (SVM)**

Überwachtes Lernen in der Anomalieerkennung

Überwachtes Lernen stellt Anomalieerkennung als **Klassifikationsproblem** dar. Aus historischen Daten wird ein Model von normalen und anomalen Beobachtungen gelernt, für neue Punkte wird eine Klassen vorhergesagt.



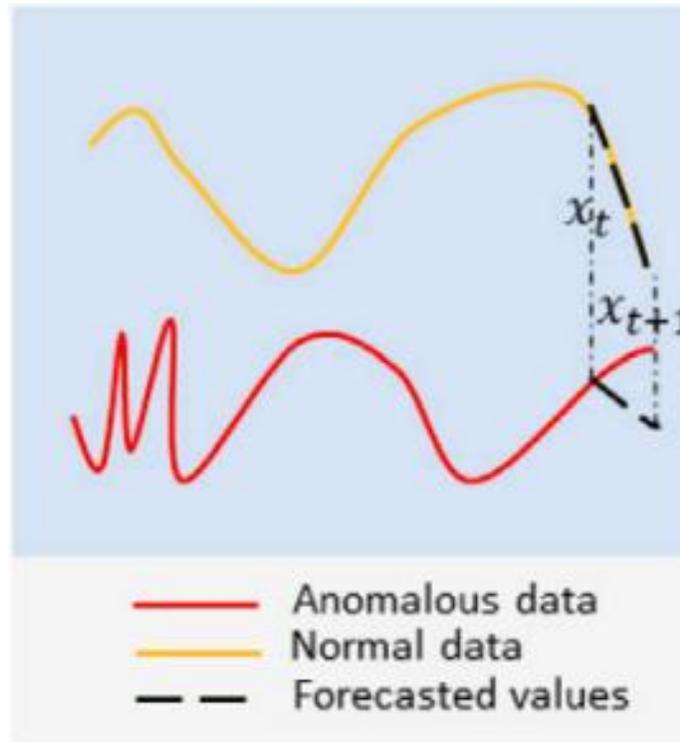
Wie können wir das testen?

Mit praktisch allen KI-Klassifikationsalgorithmen. Typisch sind hier z.B. ...

- Support Vector Machine (SVM)
- **Entscheidungsbäume**

Überwachtes Lernen in der Anomalieerkennung

Überwachtes Lernen stellt Anomalieerkennung als **Klassifikationsproblem** dar. Aus historischen Daten wird ein Model von normalen und anomalen Beobachtungen gelernt, für neue Punkte wird eine Klassen vorhergesagt.



Source: Mejri et al. (2024)
doi: 10.1016/j.eswa.2024.124922

Wie können wir das testen?

Mit praktisch allen KI-Klassifikationsalgorithmen. Typisch sind hier z.B. ...

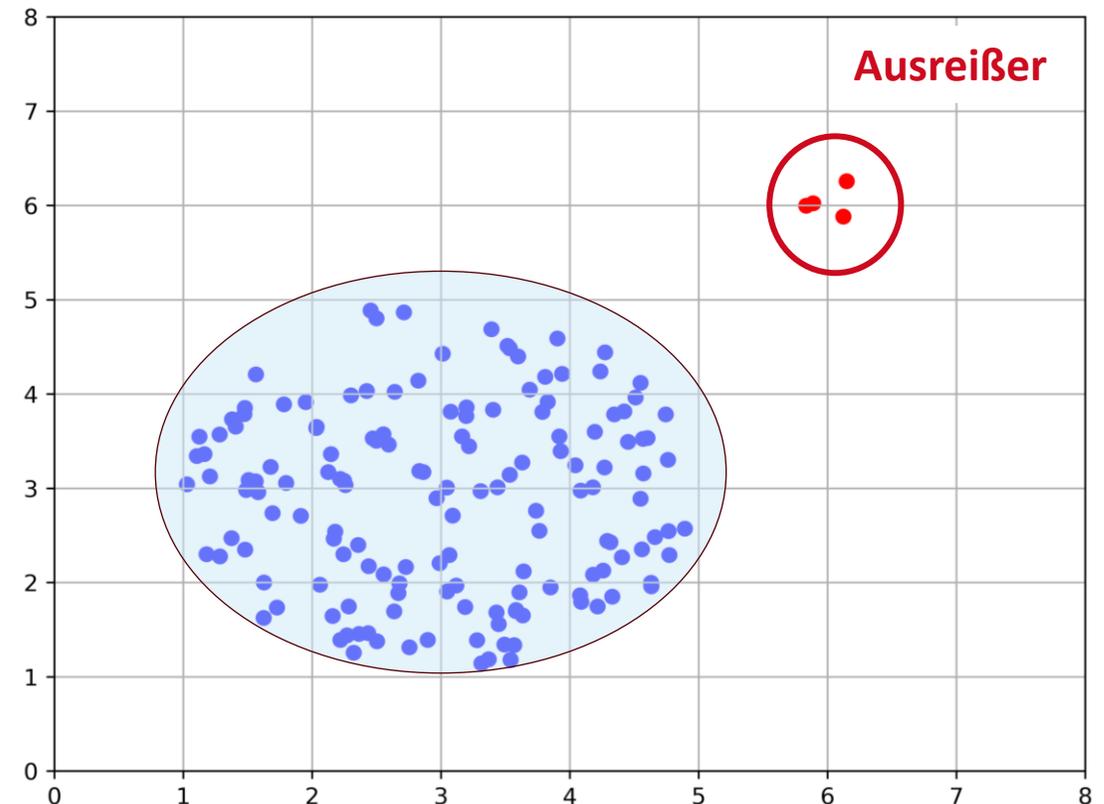
- Support Vector Machine (SVM)
- Entscheidungsbäume
- **Vorhersage nächster Werte**
 - **(S)ARIMA(X)** – autoregressive Vorhersage mit integrierter Differenzierung und laufenden Durchschnitten
 - **Neuronale Netze**, z.B. RNN, LSTM, Transformer

Unüberwachtes Lernen in der Anomalieerkennung

Mit Techniken des unüberwachten Lernens wird die **Mehrheitsverteilung** der Datenpunkte modelliert. Neue Punkte werden als Anomalien erkannt, wenn sie weit außerhalb dieser gelernten Verteilung liegen.

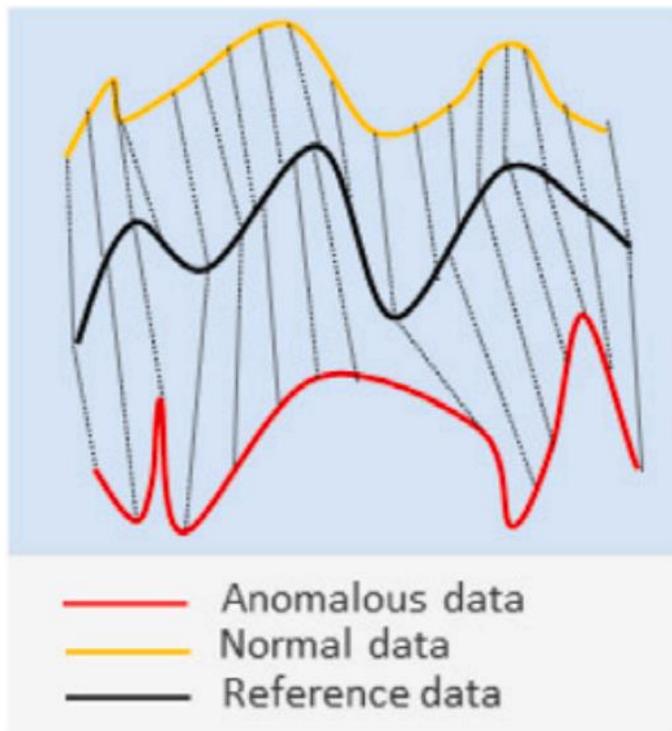
Was ist "normal"?

- Im Vorhinein nicht bekannt, Verteilung der Daten wird im Training gelernt
- Kein automatisches Feedback zu Genauigkeit
- Abgrenzung anormaler Beobachtungen händisch oder statistisch realisierbar
- Zeitaufwand für Interpretation der Ergebnisse



Unüberwachtes Lernen in der Anomalieerkennung

Mit Techniken des unüberwachten Lernens wird die **Mehrheitsverteilung** der Datenpunkte modelliert. Neue Punkte werden als Anomalien erkannt, wenn sie weit außerhalb dieser gelernten Verteilung liegen.



Source: Mejri et al. (2024)
doi: 10.1016/j.eswa.2024.124922

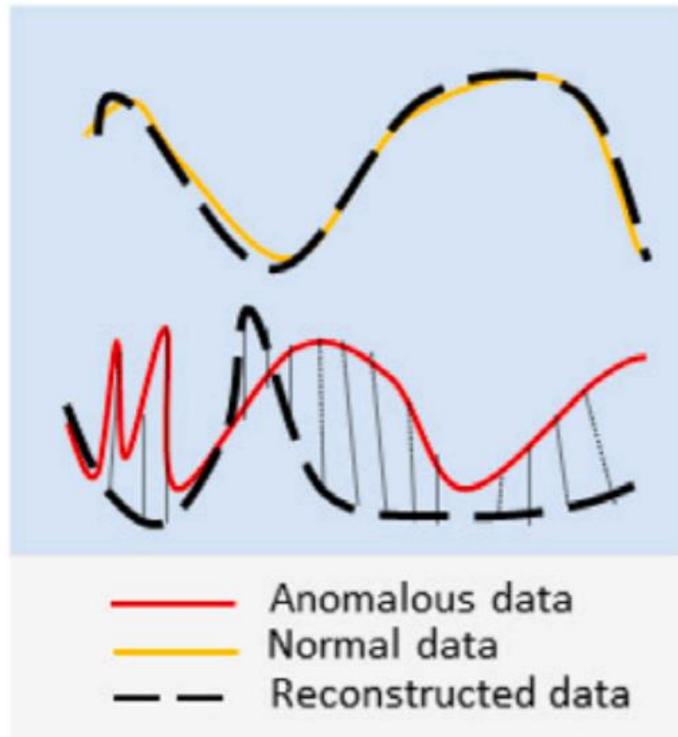
Wie können wir das testen?

▪ Distanz-basierte Methoden

- Vergleich neuer Zeitreihen gegenüber bekannten Datensätzen
- Anomalieerkennung bei hoher Abweichung zu (Teilmenge von) besten Matches
- z.B. **Dynamic-Time-Warping, DAMP**

Unüberwachtes Lernen in der Anomalieerkennung

Mit Techniken des unüberwachten Lernens wird die **Mehrheitsverteilung** der Datenpunkte modelliert. Neue Punkte werden als Anomalien erkannt, wenn sie weit außerhalb dieser gelernten Verteilung liegen.



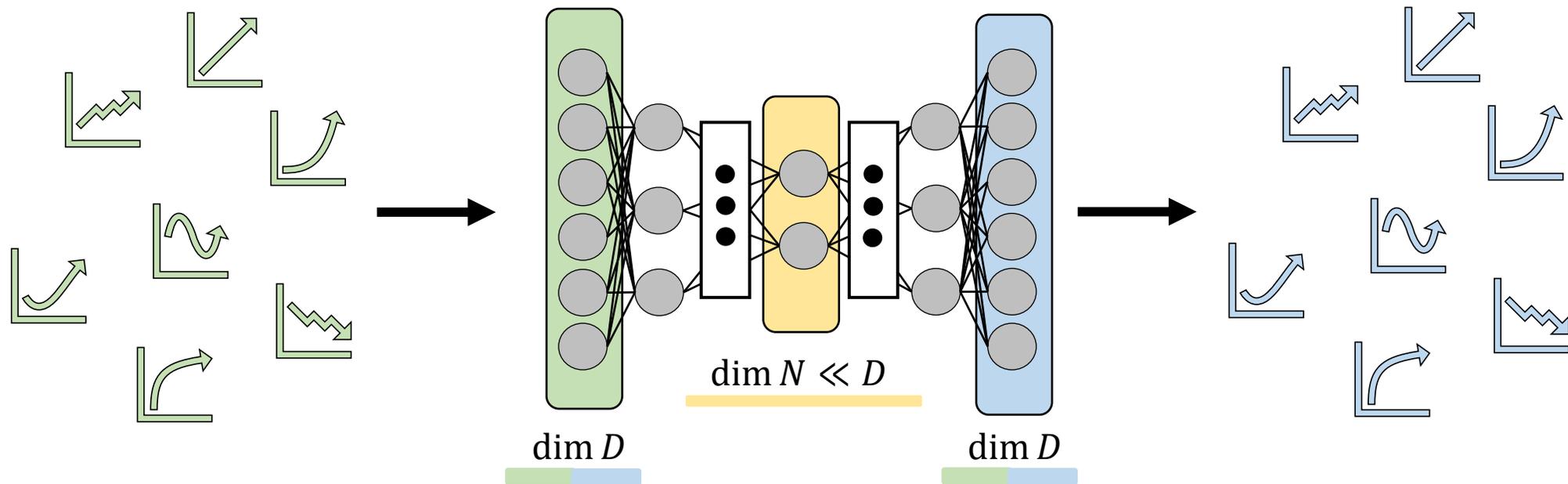
Source: Mejri et al. (2024)
doi: 10.1016/j.eswa.2024.124922

Wie können wir das testen?

- Distanz-basierte Methoden
 - Vergleich neuer Zeitreihen gegenüber bekannten Datensätzen
 - Anomalieerkennung bei hoher Abweichung zu (Teilmenge von) besten Matches
 - z.B. Dynamic-Time-Warping, DAMP
- **Rekonstruktionsbasierte Methoden**
 - Die gesamte Zeitreihe soll rekonstruiert werden
 - Höhere Abweichungen ggüb. Original als Anzeichen für Anomalien
 - **z.B. Autoencoder, GANs**

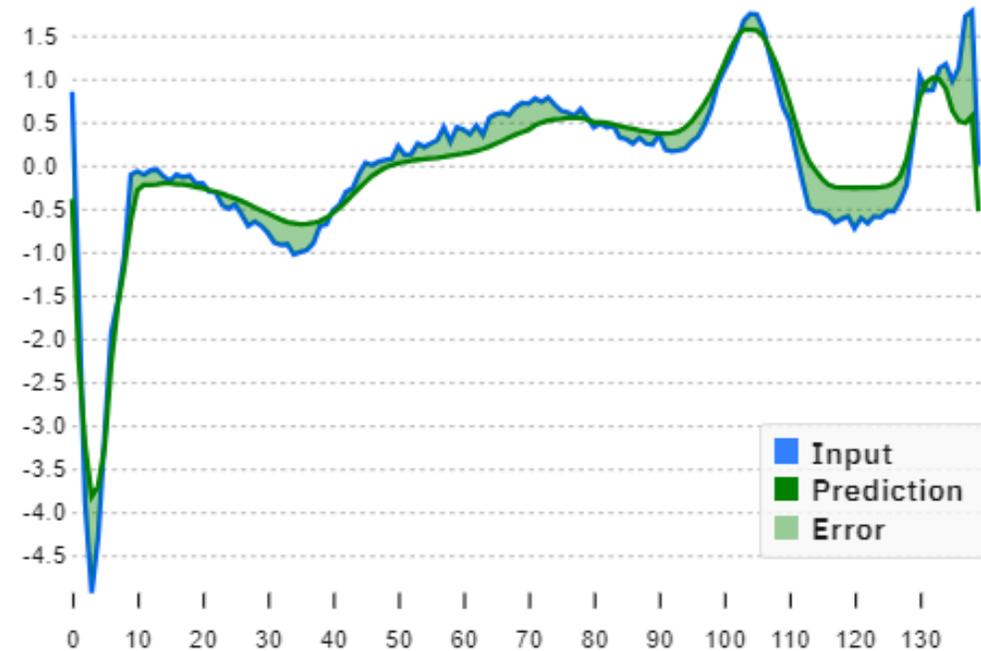
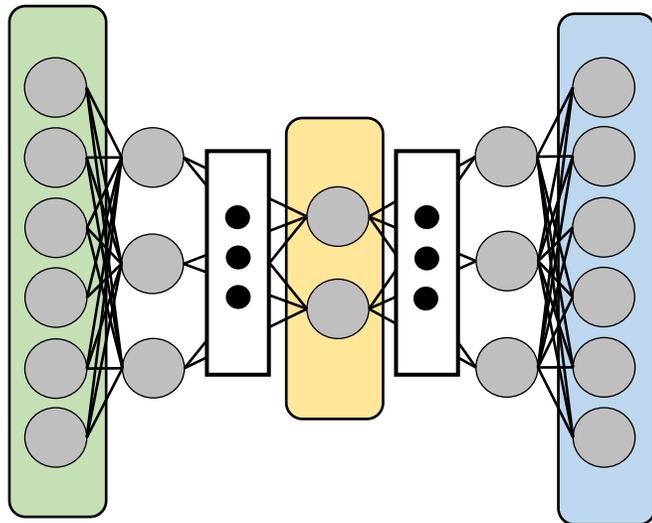
KI-Anwendungsbeispiel - Autoencoder

Ein **neuronales Netz** lernt **Daten zu rekonstruieren**, die dem Normalverhalten des Systems entsprechen. Anomalien stellen unbekanntes Verhalten dar, weshalb bei deren Rekonstruktion größere Fehler auftreten.



KI-Anwendungsbeispiel - Autoencoder

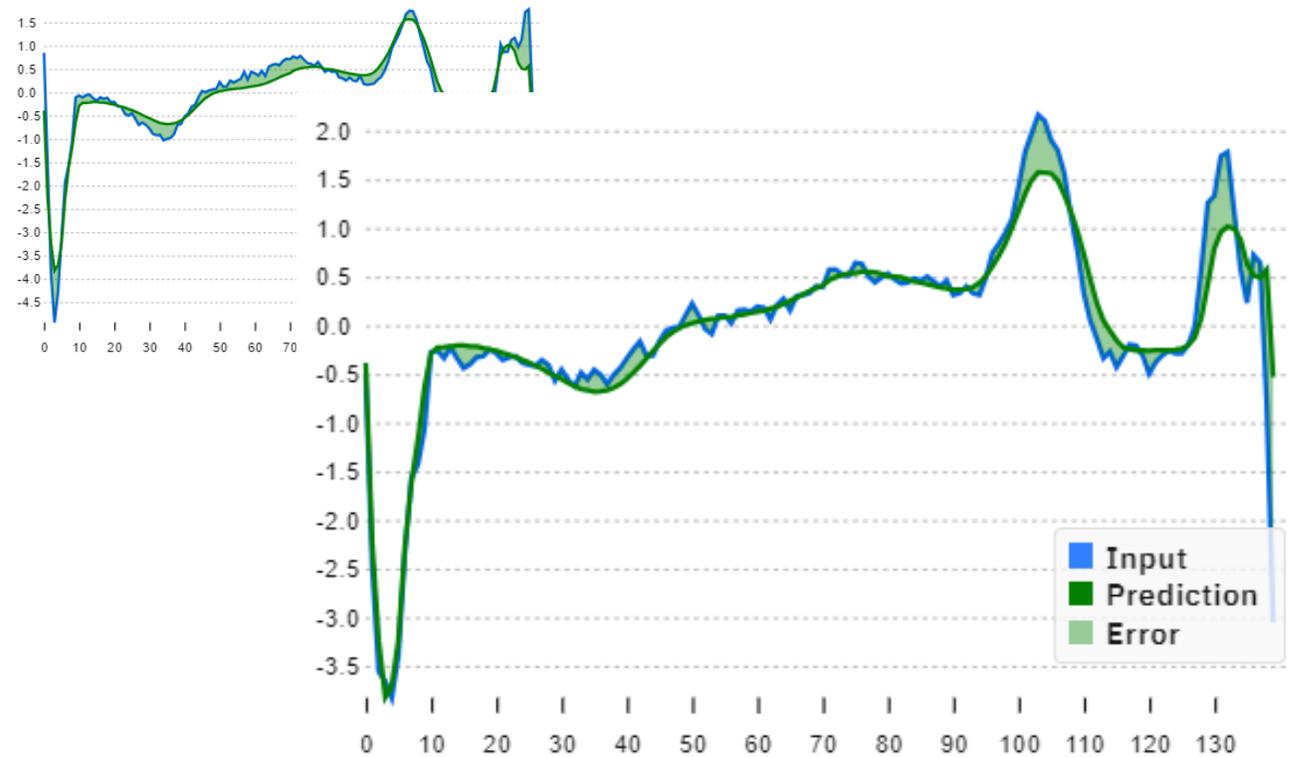
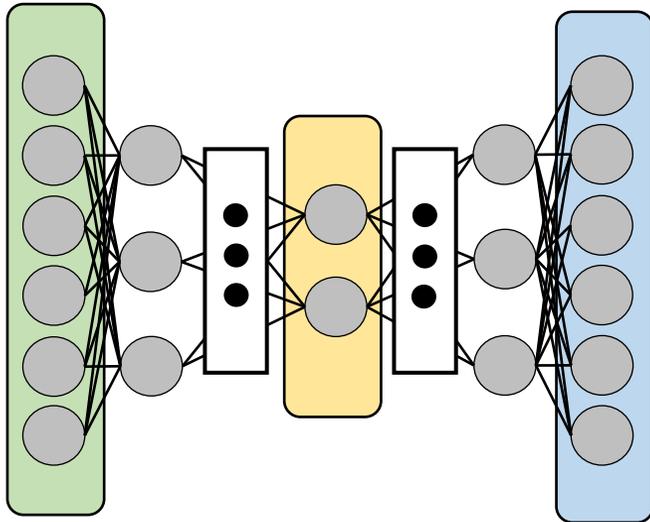
Ein neuronales Netz lernt Daten zu rekonstruieren, die dem Normalverhalten des Systems entsprechen. **Anomalien** stellen unbekanntes Verhalten dar, weshalb bei deren Rekonstruktion **größere Fehler auftreten**.



Normal

KI-Anwendungsbeispiel - Autoencoder

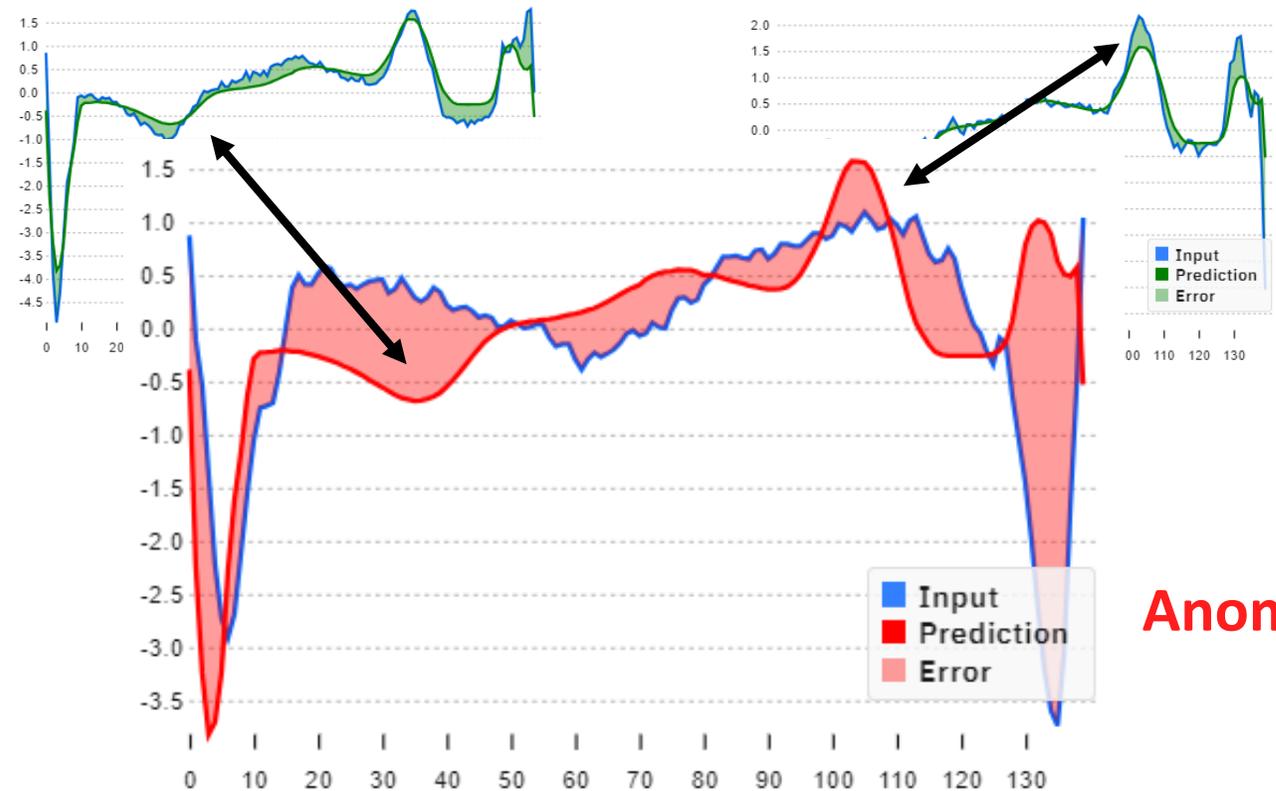
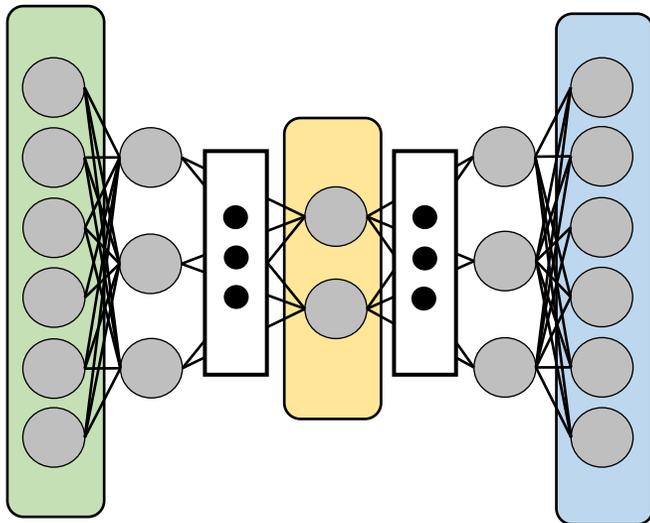
Ein neuronales Netz lernt Daten zu rekonstruieren, die dem Normalverhalten des Systems entsprechen. **Anomalien** stellen unbekanntes Verhalten dar, weshalb bei deren Rekonstruktion **größere Fehler** auftreten.



Normal

KI-Anwendungsbeispiel - Autoencoder

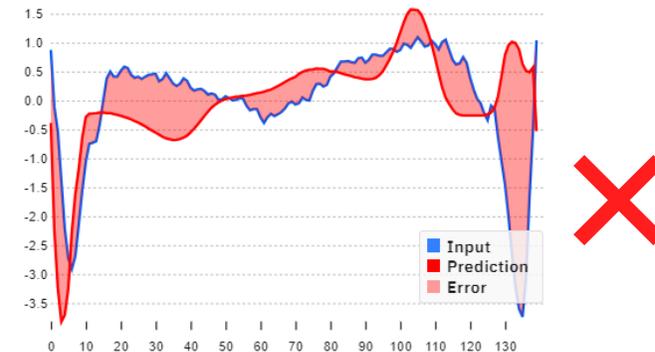
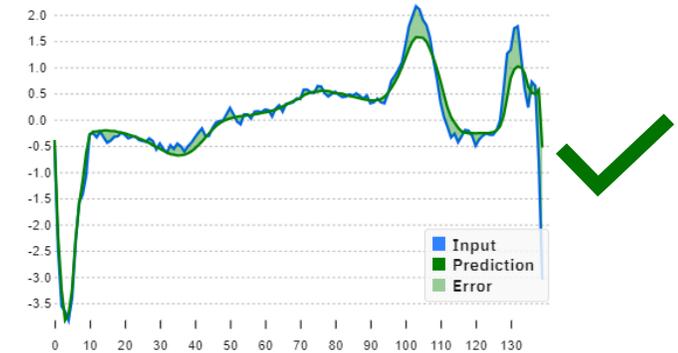
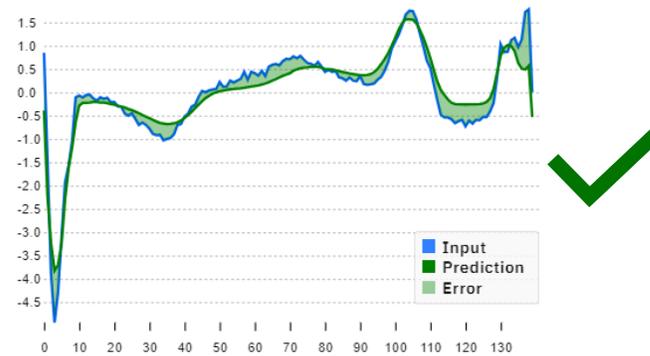
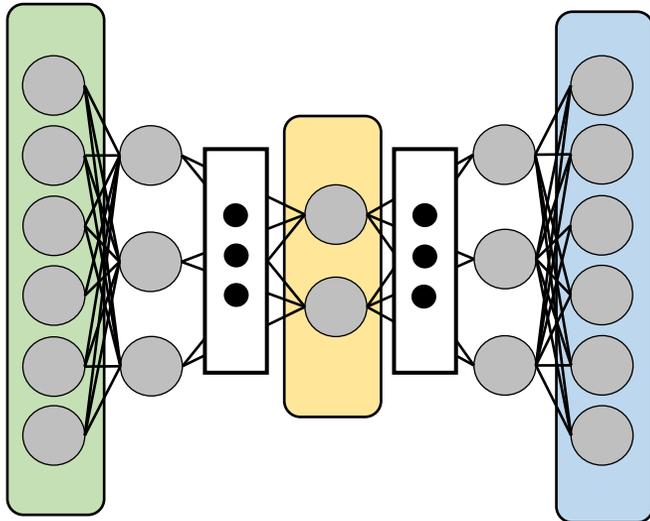
Ein neuronales Netz lernt Daten zu rekonstruieren, die dem Normalverhalten des Systems entsprechen. **Anomalien** stellen unbekanntes Verhalten dar, weshalb bei deren Rekonstruktion **größere Fehler** auftreten.



KI-Anwendungsbeispiel - Autoencoder

Ein neuronales Netz lernt Daten zu rekonstruieren, die dem Normalverhalten des Systems entsprechen.

Anomalien stellen unbekanntes Verhalten dar, weshalb bei deren Rekonstruktion **größere Fehler auftreten**.



Was haben wir **gelernt**?

- Was eine Anomalie ist und warum wir diese untersuchen möchten
- Beispiele für **statistische Methoden**, um Anomalien per Schwellwert zu finden
 - Z-Werte, IQR
 - STL-Dekomposition
- Vor- und Nachteile von **KI-basierter Anomalieerkennung**
- Überblick zu Anomalieerkennung mit KI
 - **Klassifikation** von gelabelten Daten
 - **Vergleich** oder **Rekonstruktion** bei ungelabelten Daten
 - **Vorhersage** einzelner Werte als Zwischenschritt
- **Anwendungsbeispiel** zur Anomalieerkennung mit **Autoencodern**

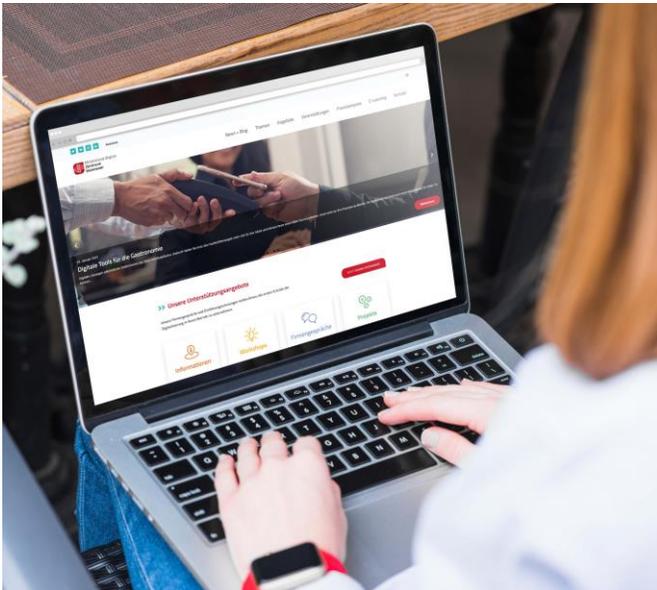


Source: pixabay.com/kinggodarts



Wie geht es **weiter?**

Lassen Sie uns in Kontakt bleiben und gemeinsam auf Ihre Erfolge aufbauen!



Workshop/Webinar



Firmengespräch



Digitalisierungsprojekt

Bleiben Sie **informiert!**

- **Newsletter** des Mittelstand-Digital Zentrums Hannover
- regelmäßige Infos über unsere kostenfreien Angebote, kommende Termine, aktuelle Projekte und Projektideen
- **Anmeldung** über QR-Code rechts oder online: www.digitalzentrum-hannover.de/newsletter



Kommen Sie **bei Fragen gerne** auf mich zu!



Uwe Hadler

Mittelstand-Digital Zentrum Hannover

KI-Trainer, Projektgenieur

+49 151 5022 9301

hadler@mitunsdigital.de

www.digitalzentrum-hannover.de



Mittelstand-
Digital 

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages