

Einführungsveranstaltung SoSe 2023

Quelle: Discovery

Energiedatenanalyse - Datamining

Lehrender



WICHTIG!

Vorlesungszeiten

Di 14:30 – 17:45

Prof. Dr. Ralf Schemm

- Professor für Energiewirtschaft, insb. elektrische Energiesysteme
- Ausbildung
 - Wirtschaftsingenieur TU-Berlin
 - Stadtwerk-> Energiehandel
 - 10 Jahre Unternehmensberatung Energiewirtschaft
 - Seit Febr. 2020 FH-Aachen:
Schwerpunktfächer im Master Energiewirtschaft & Informatik
- Kontakt:
 - schemm@fh-aachen.de
 - Raum 01B32
 - Sprechstunde: Bitte Anmeldung per Mail

1. Einleitung

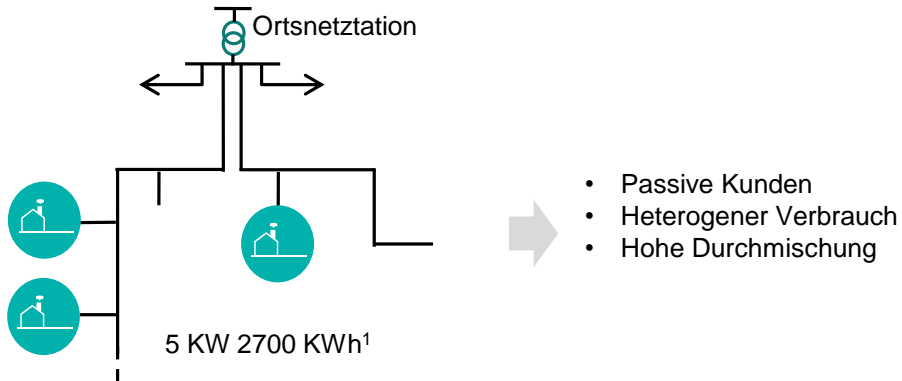
2. Aufbau

3. Prüfung



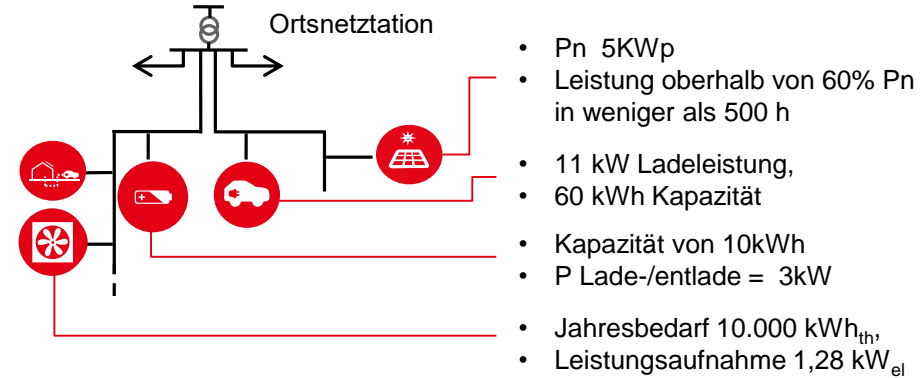
Aktuelle Herausforderungen und Notwendigkeit der Betrachtung von Flexibilität im Verteilnetz: Neuartige dezentrale Verbraucher/Erzeuger müssen in das Verteilnetz integriert werden

Früher



- Wenig flexible Verbraucher und Erzeuger
- Geringe Anzahl betrieblicher Freiheitsgrade, insbesondere im NS-Netz
- Unidirektionale Leistungsflüsse
- Aus Sicht des Vertriebs durch Standardlastprofile abgebildet

Heute/Zukünftig



- Vielzahl neuer, volatile und flexibler Verbraucher: Elektromobile, Wärmepumpen, Energiespeicher
- Neue betriebliche Freiheitsgrade
- Bidirektionale Leistungsflüsse
- Neuartige Geschäftsmodelle für Vertrieb

- **Marktteilnehmer benötigen Zugriff auf Flexibilität zur Etablierung neuer Geschäftsmodelle**
- **Netzbetreiber benötigen Erkenntnisse über Netzzustand und Novellierung der Eingriffsrechte**

¹ Ohne Durchlauferhitzer für einen Drei-Personenhaushalt

Tarifanwendungsfälle sollen dem Vertrieb ermöglichen, neuartige Produkte und Dienstleistungen mit Hilfe zur Verfügung

Auswahl aktueller Vertriebsprodukte

1

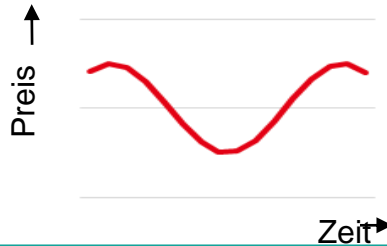
Zeitzonentarif
(statische
Zeitfenster)



envia^M
Meinstrom Digital
Happy Hour

2

Dynamischer Tarif,
abhängig vom
Marktpreis



NEXT
KRAFTWERKE
Best of 96:
dynamische
Stromtarif

3

Zugriff auf
Flexibilität



sonnen
sonnenFlat
home

Abbildung durch intelligente Messsysteme

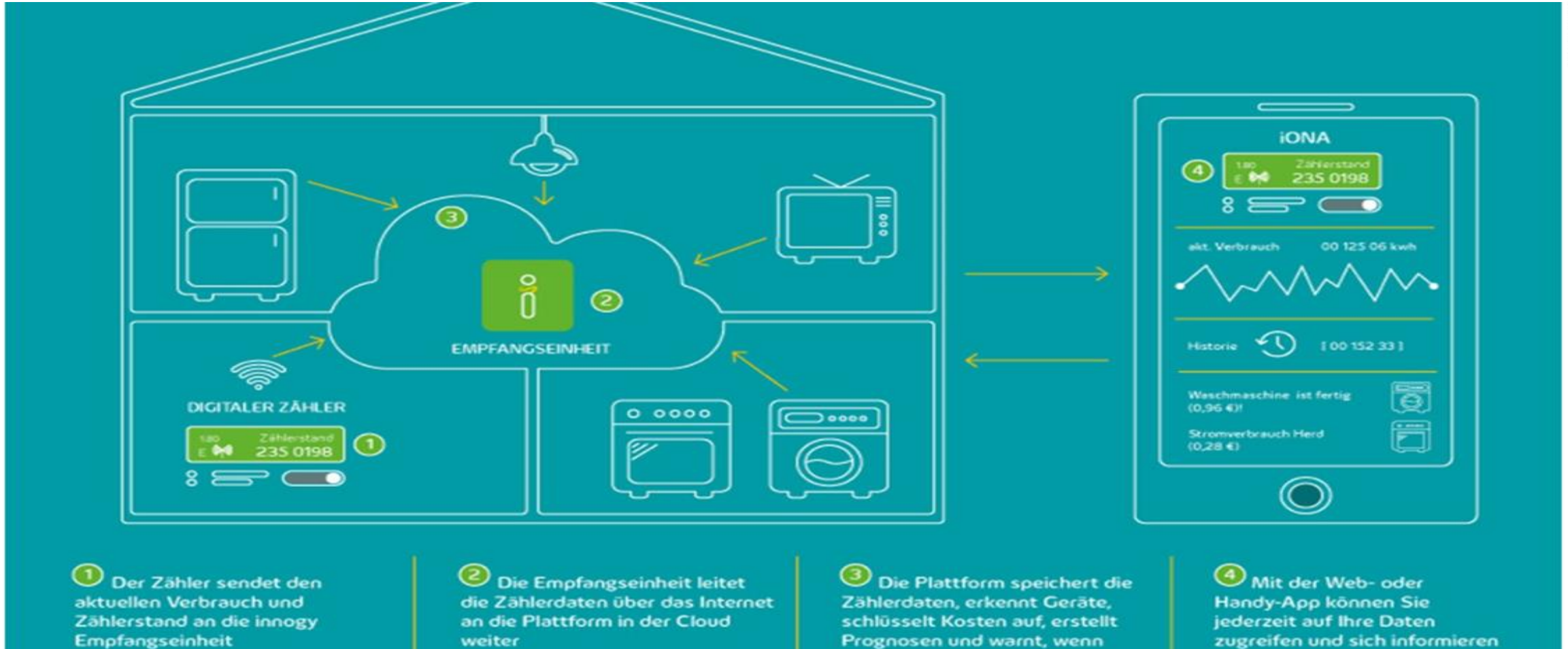
Mögliche Vertriebsprodukte müssen durch die im Smart Meter Gateway (SMG) hinterlegten zertifizierten Tarifanwendung (TAF) abbildbar sein

Mögliche Tarifanwendungen der 1. Generation nach TRI-03901-1

TAF	Auslesung
TAF 1	Jahresverbrauchsmenge
TAF 2	Verbrauchsmenge je Zeitstufe
TAF 7	15-Minuten-Lastgang

➤ Bislang wird die Nutzungsmöglichkeit/ der Zugriff auf Flexibilität durch die zertifizierten TAF nicht unterstützt

Energieversorger können mit Alternativen zum Einbau eines zertifizierten intelligenten Messsystems deutlich mehr Daten erhalten



Quelle: <https://www.iona-energy.com/iona/ablauf>

Anwendungsfall: Verbrauchserkennung

Ab März bietet enviaM das Stromprodukt „Mein Strom digital“ in Zusammenarbeit mit einem Dienstleister und proprietärer Messtechnik (kein iMSys, kein mME) an

Zusatzprodukte Smart Meter

ZEK+ | SMART ENERGY

12.03.2019

Großes Kundeninteresse an Envia-M-Produkt "Mein Strom digital"

Schon über 1000 Vorbestellungen kann der ostdeutsche Energiedienstleister für sein Digitalisierungspaket verzeichnen – darunter auch zahlreiche Neukunden. Dazu gehört ein Zähler mit Echtzeitdatenauflösung und Verbrauchsaufschlüsselung einzelner Haushaltsgeräte. Das BSI steht dem Gerät derzeit noch skeptisch gegenüber.



Rahmenbedingungen

- Zielkunden: Haushalte < 6.000 kWh/a
- Proprietäre Zählertechnik und Empfangsbox der Firma lemonbeat (Innogy)
- Datenübertragung per Funk und WLAN inhouse auf die App
- Alternativ: Datenübertragung über Kunden-Router zur innogy-cloud und ins Portal

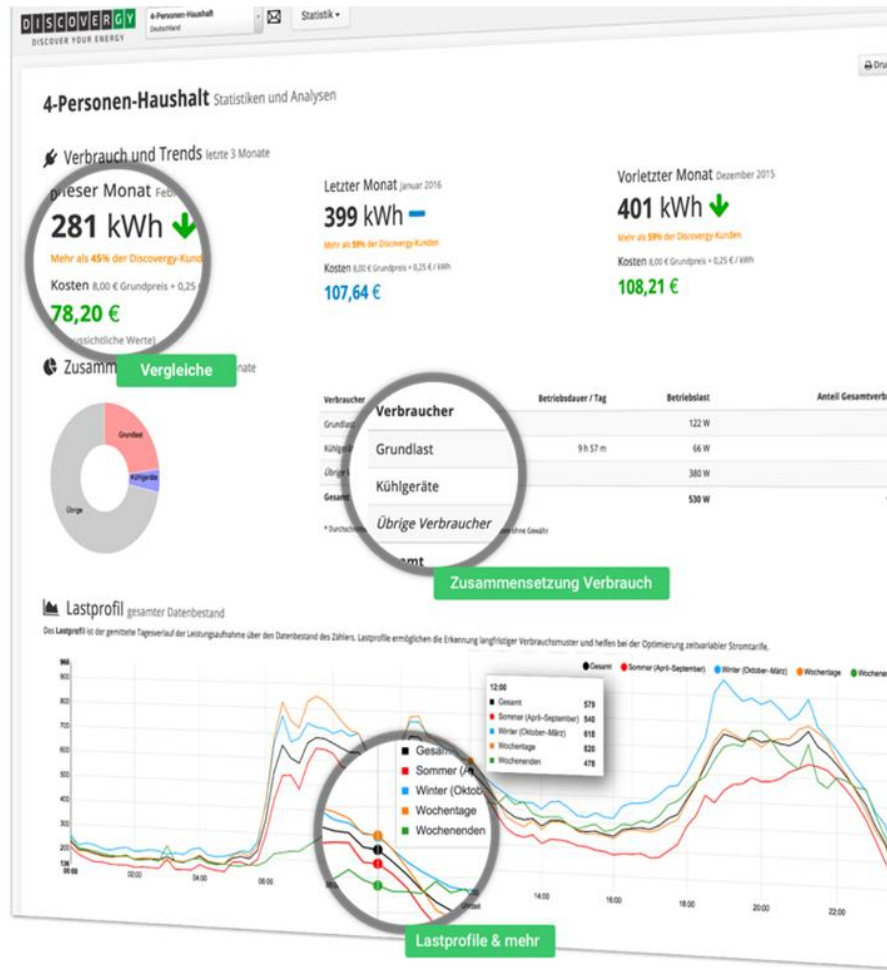
Leistungsumfang

- Stromverbrauch in Echtzeit, Verbrauchsstatistik und Analysen
- Geräteverbrauch und Kosten*
- Happy-Hour: täglich 19 – 20 Uhr eine Stunde Strom zum halben Preis



? Warum macht der Kunde / enviaM dies?

Anwendungsbeispiel: Geräteerkennung mit Hilfe von Smart Metern



Quelle: Discovergy

Anwendungsfall: Verbrauchserkennung

Anwendungsbeispiel: Geräteerkennung mit Hilfe von Smart Metern

Geräteerkennung

Discovery setzt Meilensteine in der Forschung und Entwicklung von Algorithmen, die Verbrauchsmuster einzelner Geräte und Anlagen erkennen. Mit dieser Technologie erhalten Sie nie dagewesene Möglichkeiten für intelligentes Energiemanagement und eine zielgerichtete Steuerung Ihres Verbrauchs.



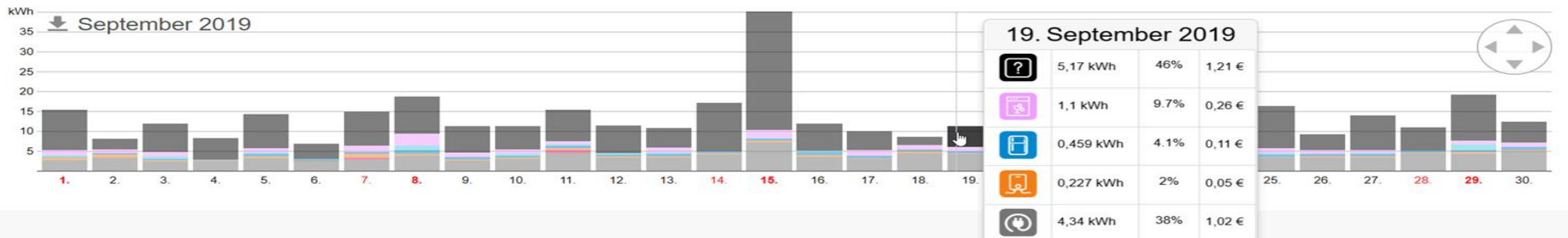
Erkennen des Verbrauchs

Mit detaillierten Verbrauchsanalysen haben Sie die Verbrauchsmuster Ihrer Geräte immer im Blick. Bei ungewöhnlichen Ereignissen können Sie Benachrichtigungen aktivieren und vermeiden unnötigen Verbrauch oder Ausfälle.



Predictive Maintainance

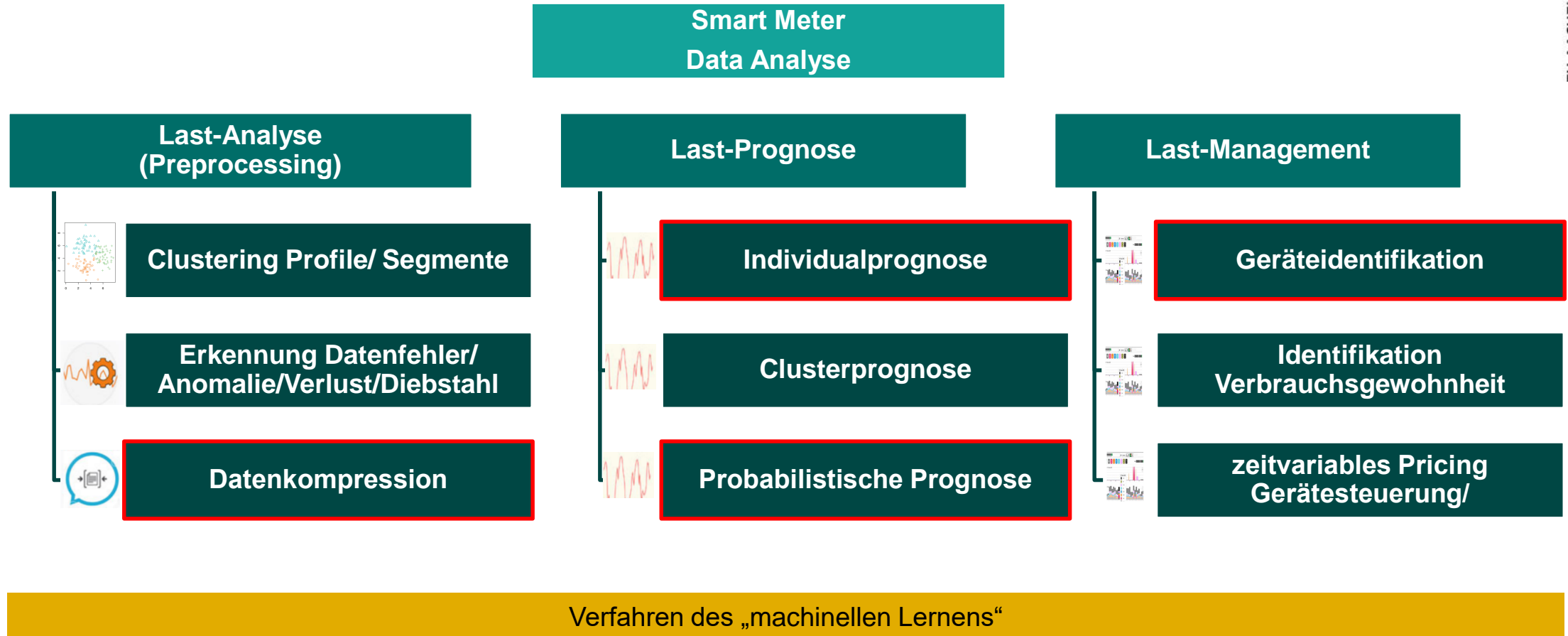
Durch unsere genauen Analyse-Tools können Sie jederzeit erkennen, ob Geräte und Anlagen einwandfrei funktionieren. So lassen sich Einstellungsfehler und drohende Geräteausfälle frühzeitig erkennen und vermeiden.



Quelle: Discovery

Anwendungsfall: Verbrauchserkennung

Systematisierung der Fragestellungen und Bezug zu den relevanten Machine-Learning Technologien



Anwendungsbeispiel: Predictive Maintenance für Windkraftanlagen

Windkraftanlagen enthalten Sensoren, welche kontinuierlich Daten aufnehmen



- Die von den Sensoren aufgezeichneten Zeitserien können mit Hilfe von Mustererkennungsverfahren bearbeitet werden.
- Von Predictive Maintenance spricht man, wenn anhand der Mustererkennung Rückschlüsse auf den Bauteilzustand bzw. den Verschleiß getroffen werden und eine Anpassung der Instandhaltungsstrategie erfolgt.
- Das heißt, Bauteile werden ausgetauscht, bevor sichtbare Defekte auftreten, beziehungsweise zum optimalen Zeitpunkt dem Überholungskreislauf wieder zugeführt.

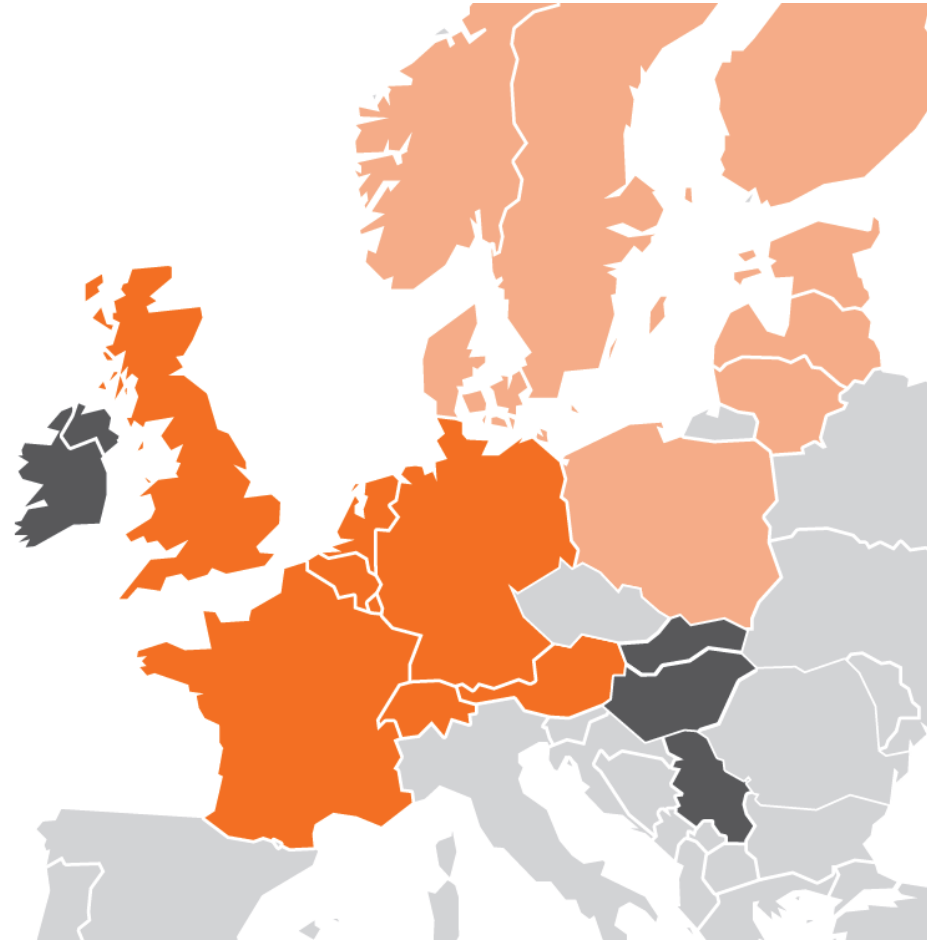
Durch Predictive Maintenance sollen die Lebenszyklen von Windkraftanlagen verlängert und effiziente Wartungskreisläufe etabliert werden.

Anwendungsbeispiel: Einsatz von Regressionsverfahren zur Preisprognose I

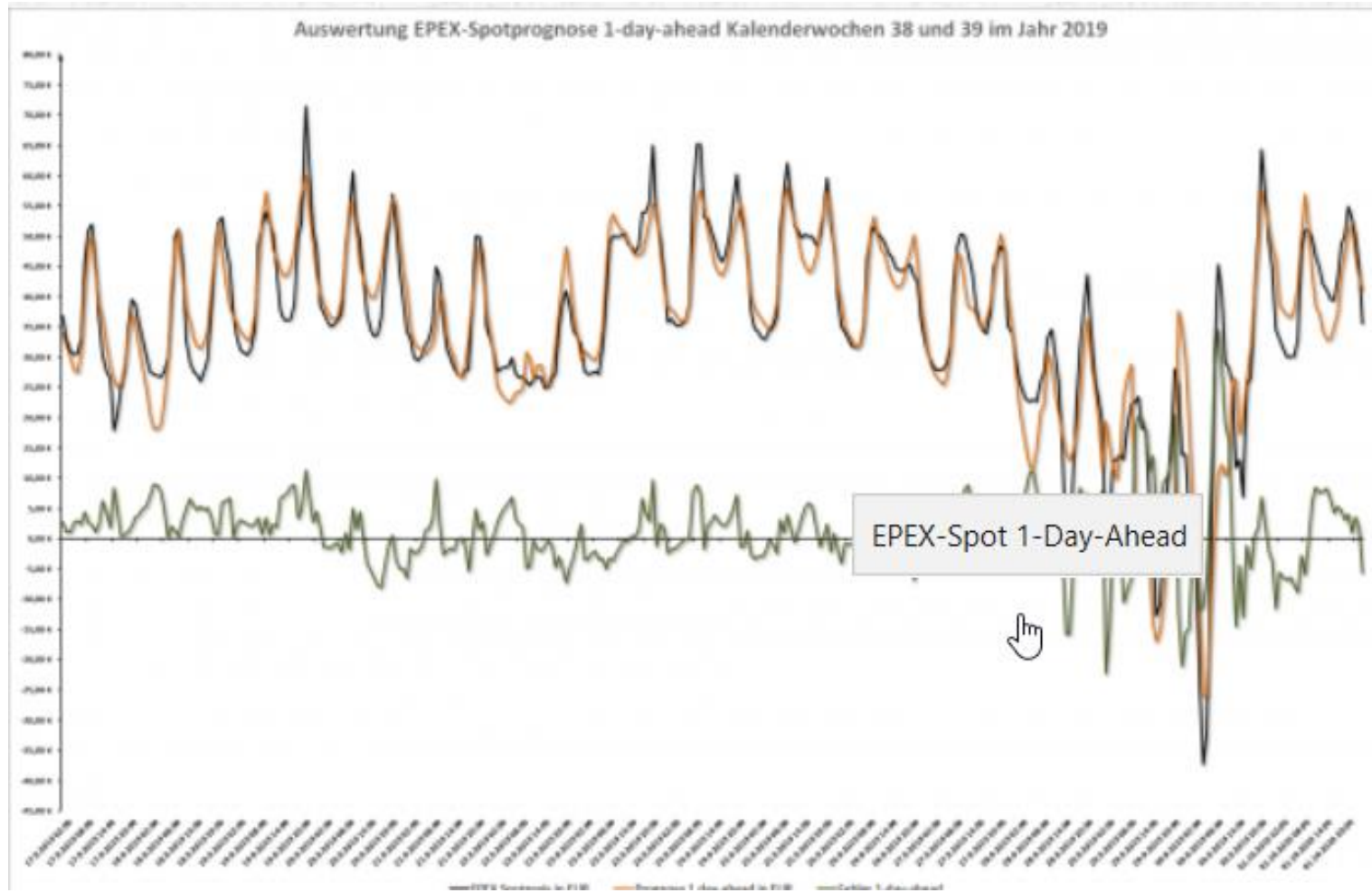
Markets

The European Power Exchange EPEX SPOT SE operates the short-term electricity markets for Germany, France, the United Kingdom, the Netherlands, Belgium, Austria, Switzerland and Luxembourg; markets representing 50% of European electricity consumption. In 2018, its 289 members traded 567 TWh. The creation of a pan-European power market is what defines EPEX SPOT's mission.

-  EPEX SPOT markets
-  Coming soon
-  Serviced Power Exchanges



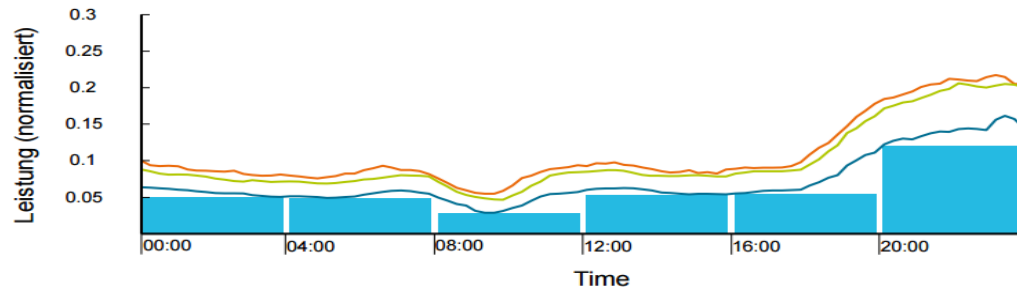
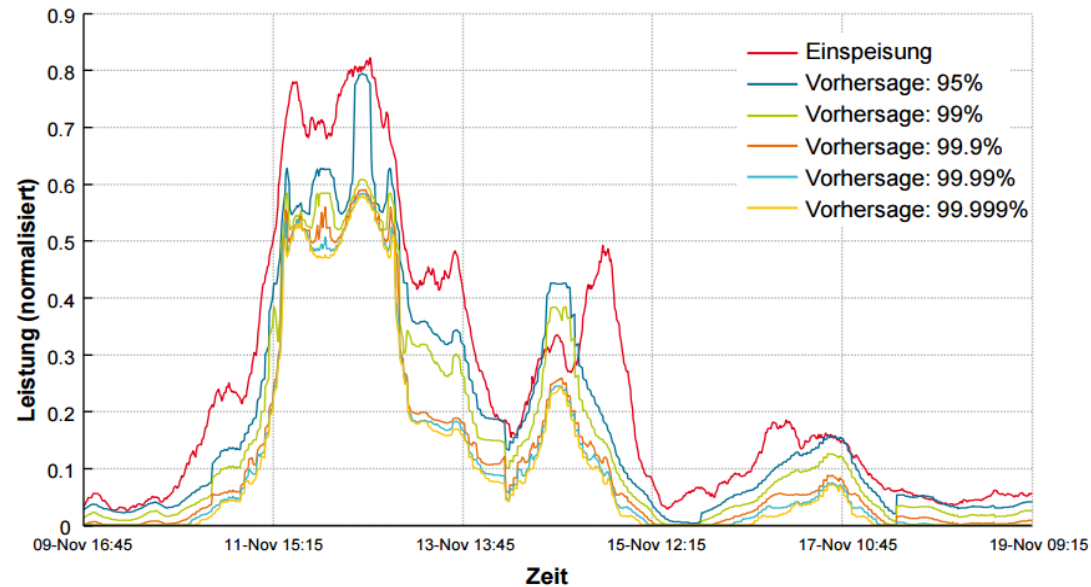
Anwendungsbeispiel: Einsatz von Regressionsverfahren zur Preisprognose II



Quelle: Priceit

Anwendungsfall: Preis-Prognose

Probabilistische Prognosen eignen sich besonders für die Regelenenergievermarktung



Regelenenergievermarktung

- Grundlage: Einsatz probabilistischer Einspeiseprognosen
- Berücksichtigung einer Vorhersagegenauigkeit zu einem definierten Quantil
- Ableitung einer Mindesteinspeisung als Grundlage der RE-Vermarktung

- Sichere Vortagsprognose
- Sichere Untertagsprognose
- Einspeisung
- Angebot Regelleistung

1. Einleitung

1. Anwendungsfälle

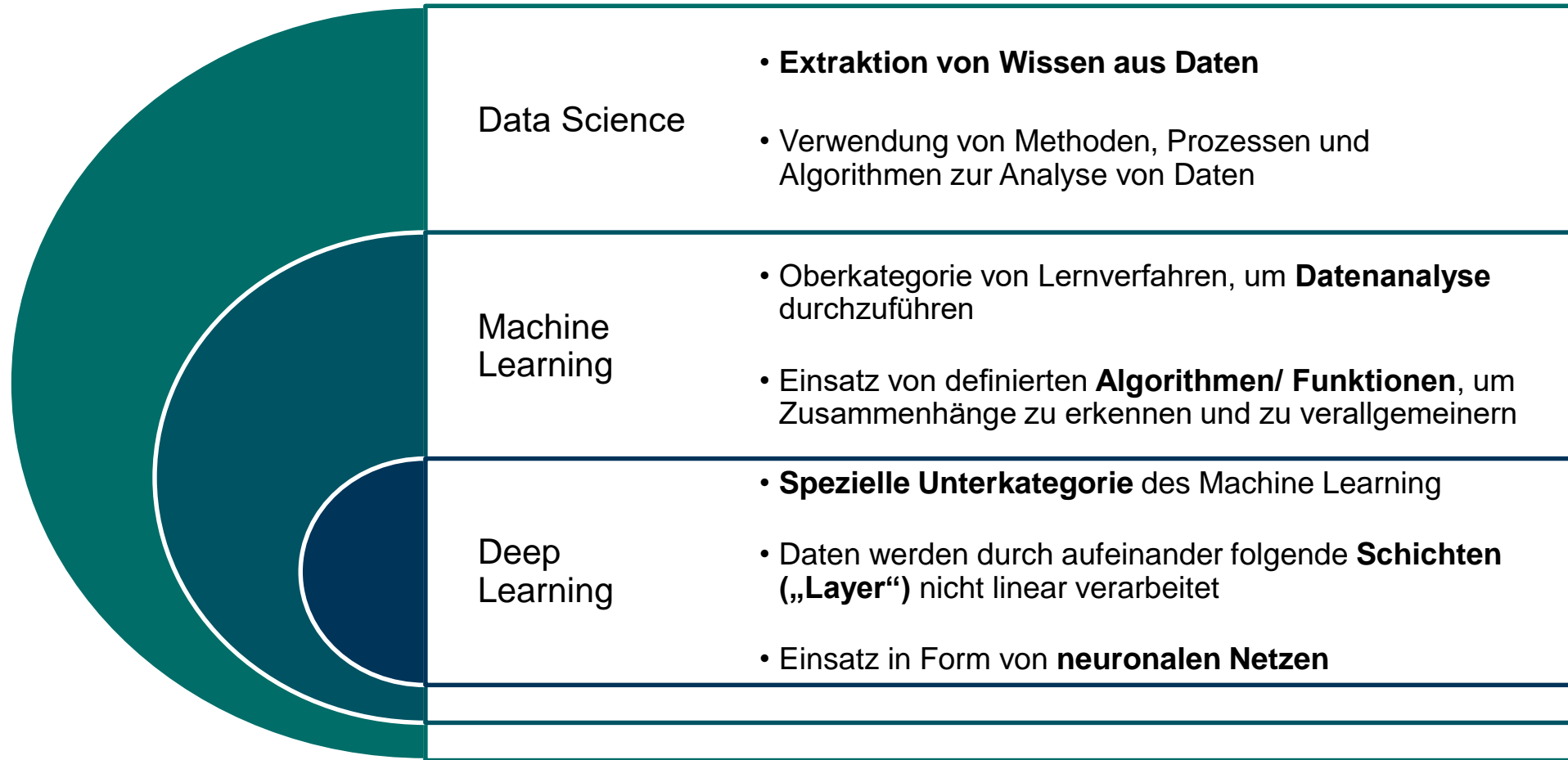
2. Überblick Algorithmen

2. Aufbau

3. Prüfung



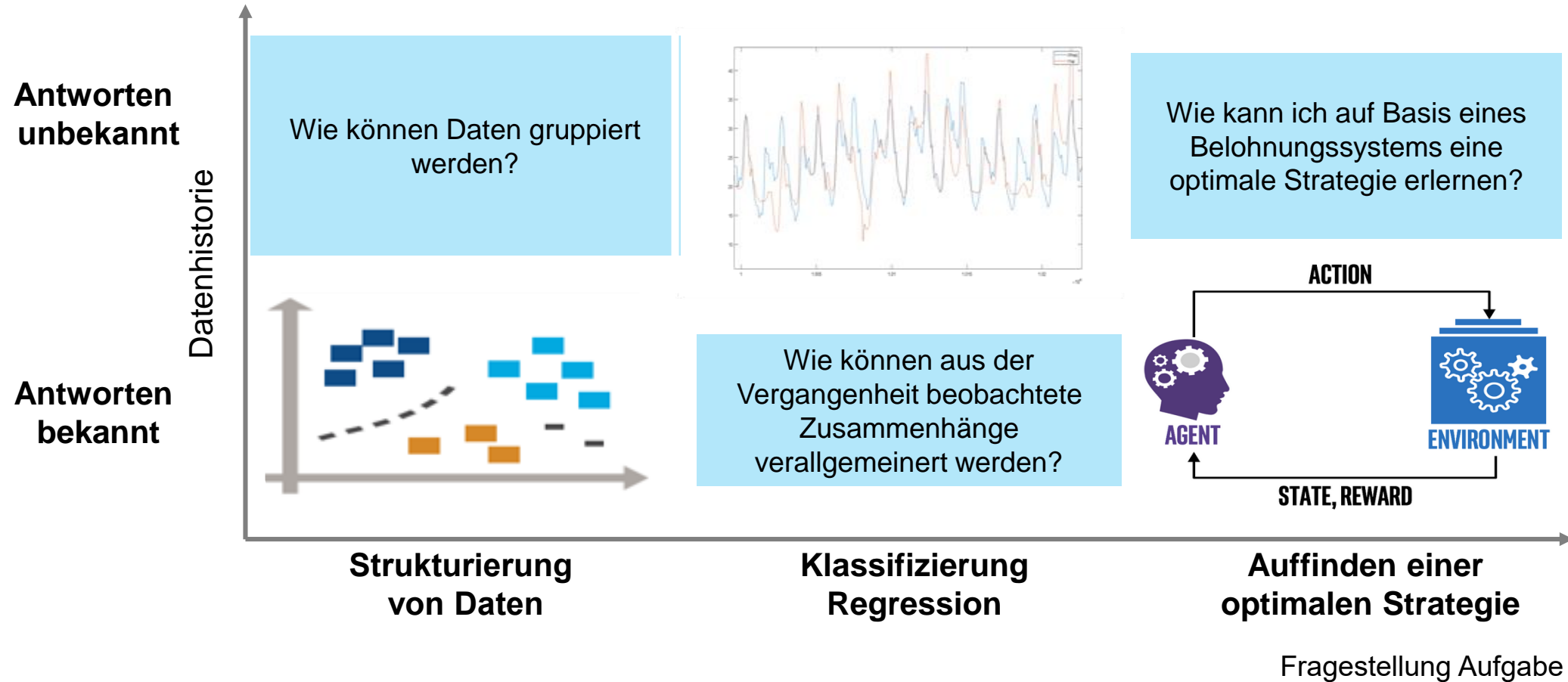
Zuerst müssen ein paar Begriffe voneinander abgegrenzt werden



Bereiche der Datenanalyse: eine nicht überlappungsfreie Zuordnung

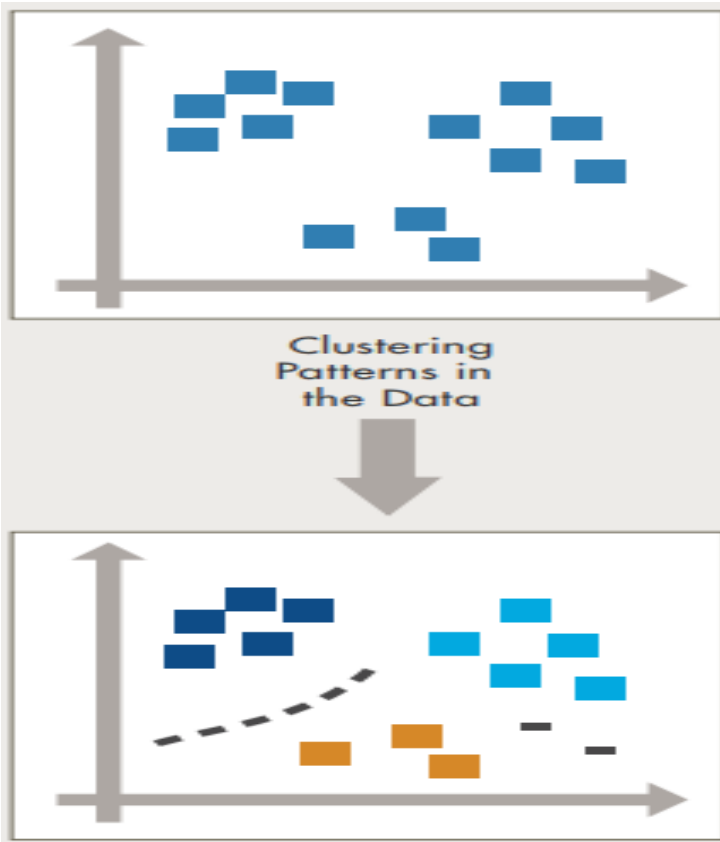


Fragestellungen/ Aufgaben werden in Abhängigkeit der Struktur der Daten und der Aufgabe systematisiert



Unsupervised Learning versucht durch Gruppierung Datenstrukturen und Muster aufzuzeigen

Clustering



- Unsupervised Learning versucht versteckte Muster und Strukturen in Daten aufzudecken.
- Es wird explizit nicht der Zusammenhang zwischen Input- und Outputgröße betrachtet.

Betrachtete Methoden:

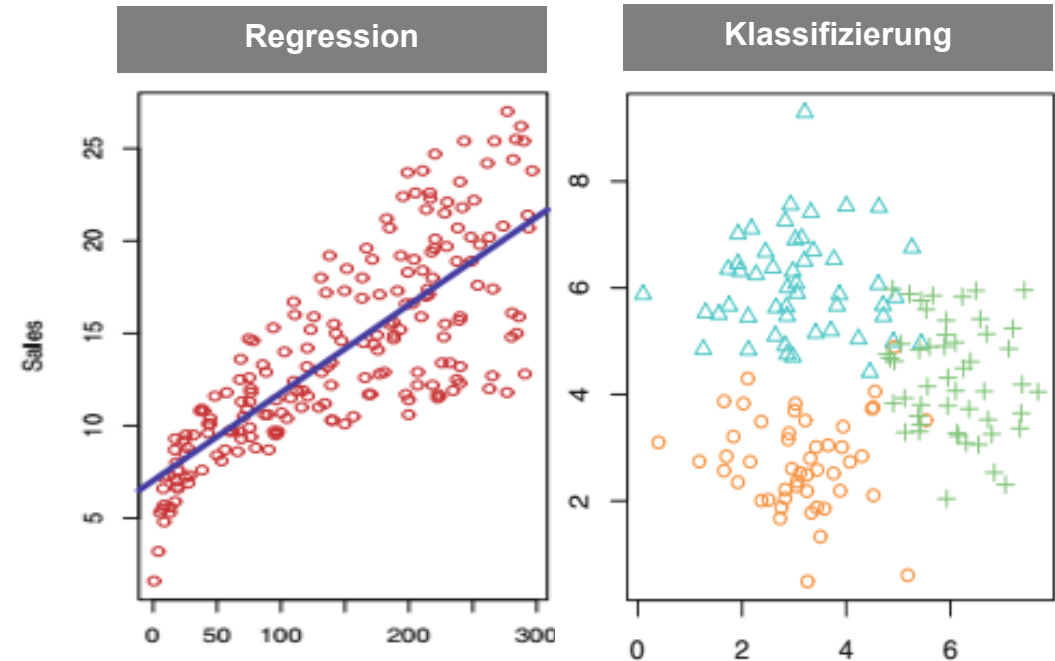
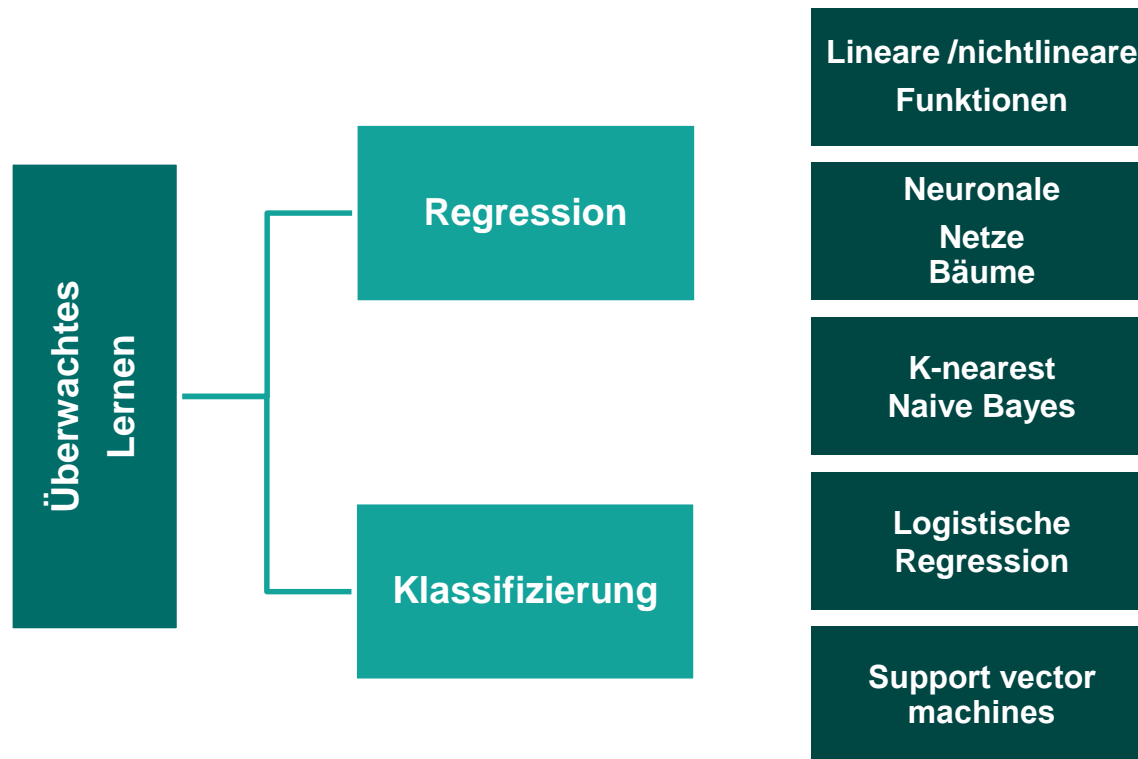
Principle Component Analysis

K-Means

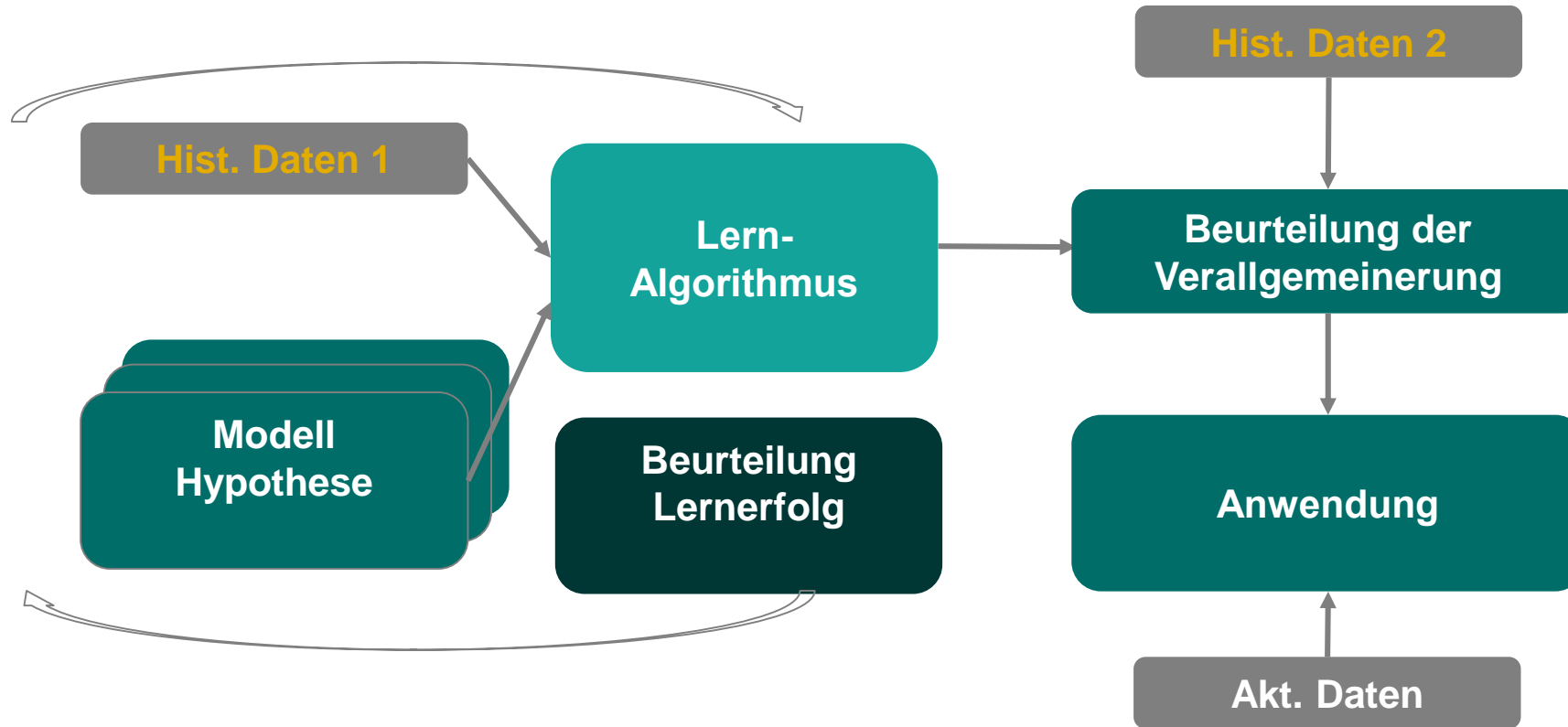
Hierachical Clustering

Supervised Learning versucht Abhängigkeiten zwischen Input- und Outputdaten zu analysieren

- Supervised Learning modelliert die Abhängigkeit einer Ergebnisgröße Y und Eingangsgrößen X
- Es wird **ein funktionaler Zusammenhang** mit Hilfe eines definierten Modells abgebildet



Eine wesentliche Aufgabe des Anwenders ist somit die richtige Modellauswahl

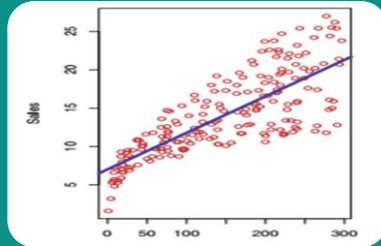


Zusammenfassung: Die drei Arten des Machine Learnings leiten sich aus den damit verbundenen Zielsetzungen und Voraussetzungen ab



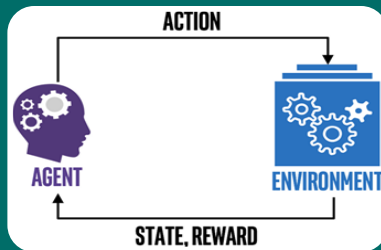
Unsupervised Learning (Strukturanalyse)

- Keine übergeordnete Zielgröße
- Kein Feedback
- Anwendungszweck: Strukturen sichtbar machen



Supervised Learning

- Vorliegen gekennzeichneteter Daten (Input/ Output)
- Direktes Feedback zur Vorhersage
- Anwendungszweck: Prognose

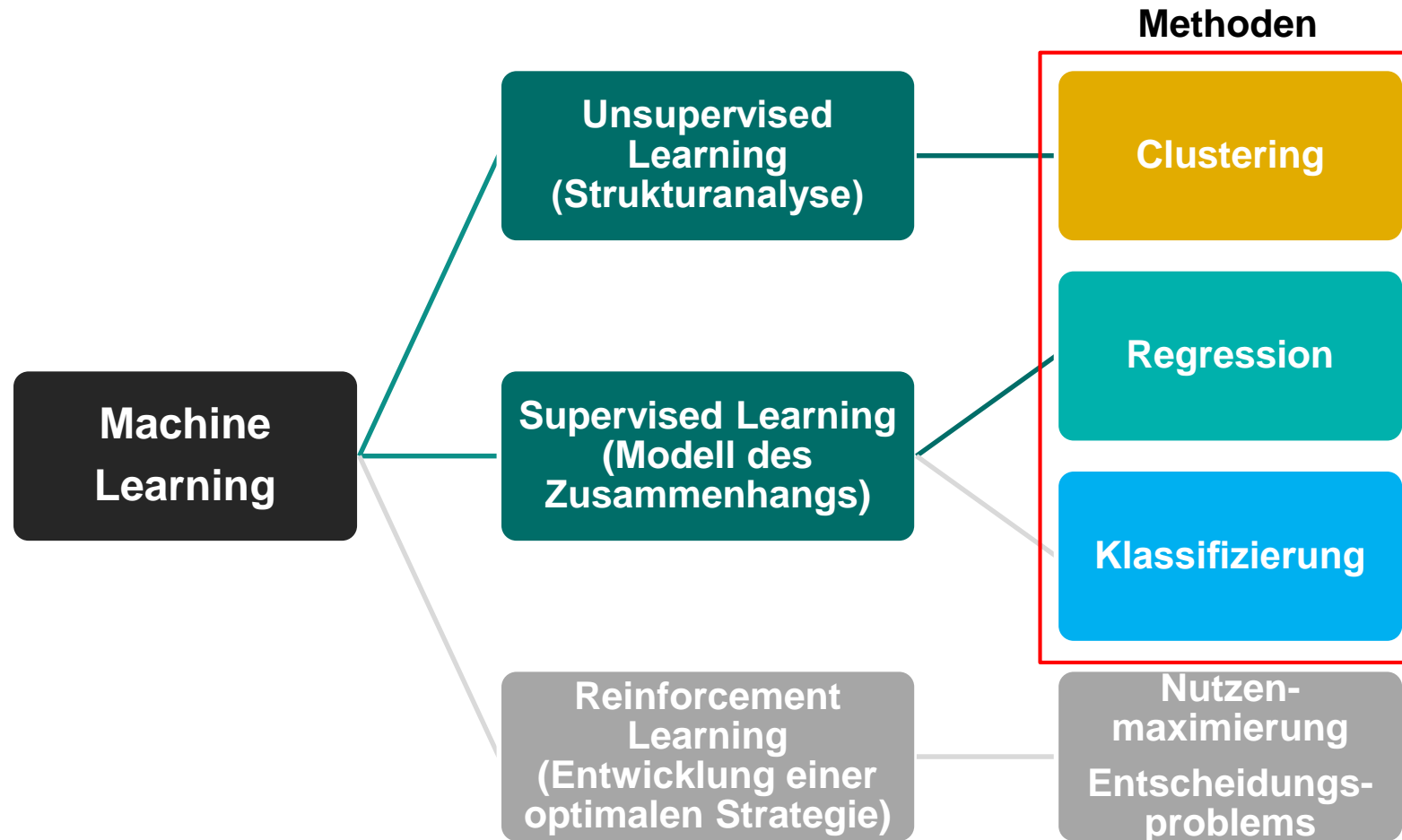


Reinforcement Learning

- Entscheidungsvorgang
- Erlernung anhand Belohnungssystem
- Anwendungszweck: Einsatzentscheidung

Die betrachteten quantitative Methoden fallen unter die Rubrik Machine Learning

Ziel: Aufzeigen Beziehungen in Datenstrukturen mit Hilfe von Algorithmen



1. Einleitung

2. Aufbau

3. Prüfung

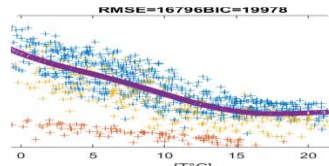


Die Voraussetzungen für den Kurs sollten bei allen Teilnehmern vorliegen, um den Inhalten folgen zu können

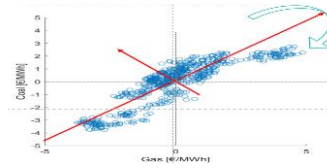


Was haben wir vor ...

Lineare und nichtlineare Regression



Feature engineering und Datenreduktion (PCA)

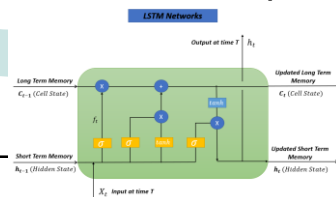


Technik

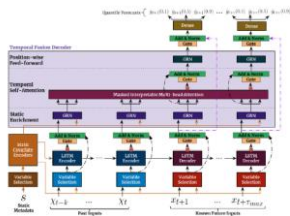
Einführung in Tensorflow und Keras



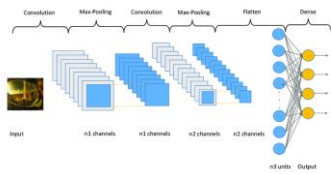
Rekurrente Neuronale Netze (LSTM)



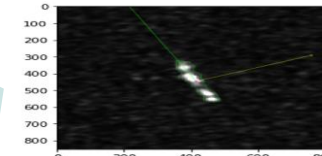
Temporal Fusion Transformer



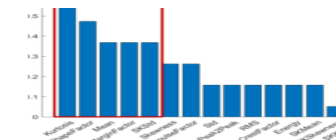
Convolutionale Neuronale Netze (CNN)



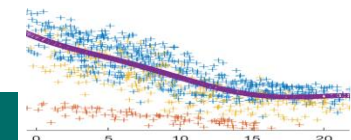
Ausrichtungsanalyse Windkraftanlage



Predictive Maintenance

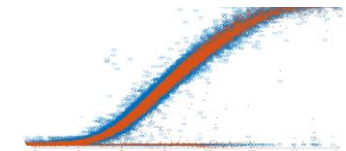


Multivariate Gasnachfrage

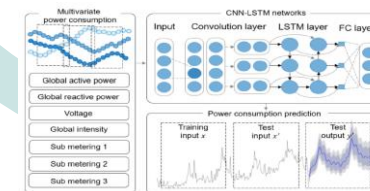


Anwendung

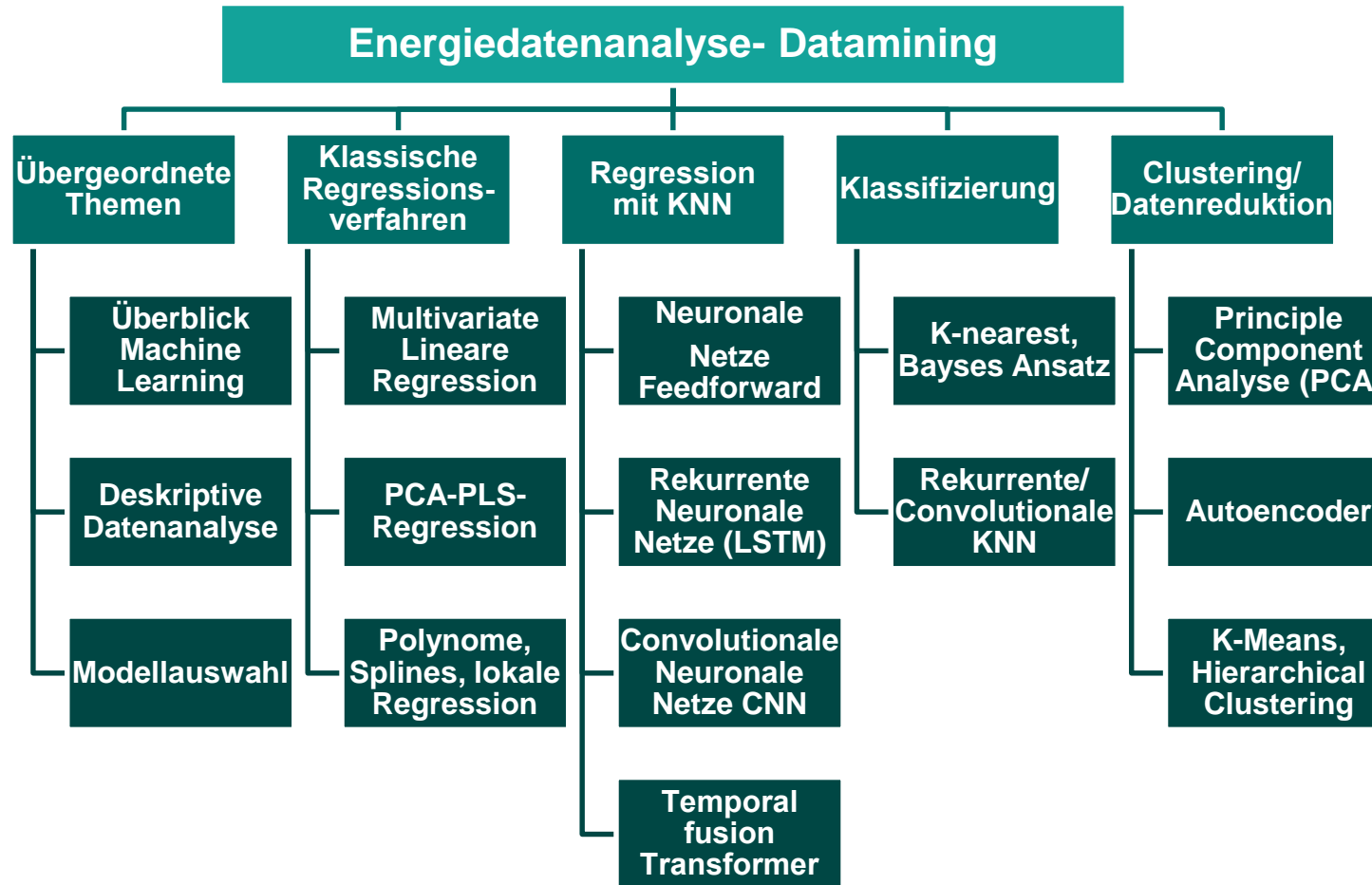
Einspeise-Prognose Windkraftanlage



Last-/Preisprognose



Die Themengebiete der Veranstaltung verknüpfen Modelle des „Machine Learning“ mit energiewirtschaftlichen Fragestellungen



WICHTIG: Rückfragen zur Case Study werden nach Vorlesungsschluss erst wieder am 15.08.2023 beantwortet!

VL	Termin	Inhalt
28.03.2023	1	Einführung Lineare Regression
04.04.2023	2	Lineare Regression und und Modellauswahl
11.04.2023	3	Nichtlineare Regression (Spline und Kernel Regression)
18.04.2023	4	Datenreduktionstechniken (PCA)
25.04.2023	5	Einführung in Keras, Tensorflow
02.05.2023	6	Fed Forward-Netze (Modellierung Windeinspeisung)
09.05.2023	7	Hyperparameter: Einführung in Keras Tuner und Tensorboard
16.05.2023	8	Rekurrente Netze: KNN LSTM Regression
23.05.2023	9	Rekurrente Netze: KNN LSTM CNN Regression
30.05.2023		Arbeiten an Case Study Benchmarkmodell
06.06.2023	10	Temporal Fusion Transformer
13.06.2023	11	Temporal Fusion Transformer
20.06.2023	12	Anwendung: Non Intrusive Load Management
27.06.2023	13	Anwendung: Predictive Maintanance
15.08.2023	14	Rückfragen Case Study Teil B
10.09.2023		Abgabe Case Study

Benchmark-
modell



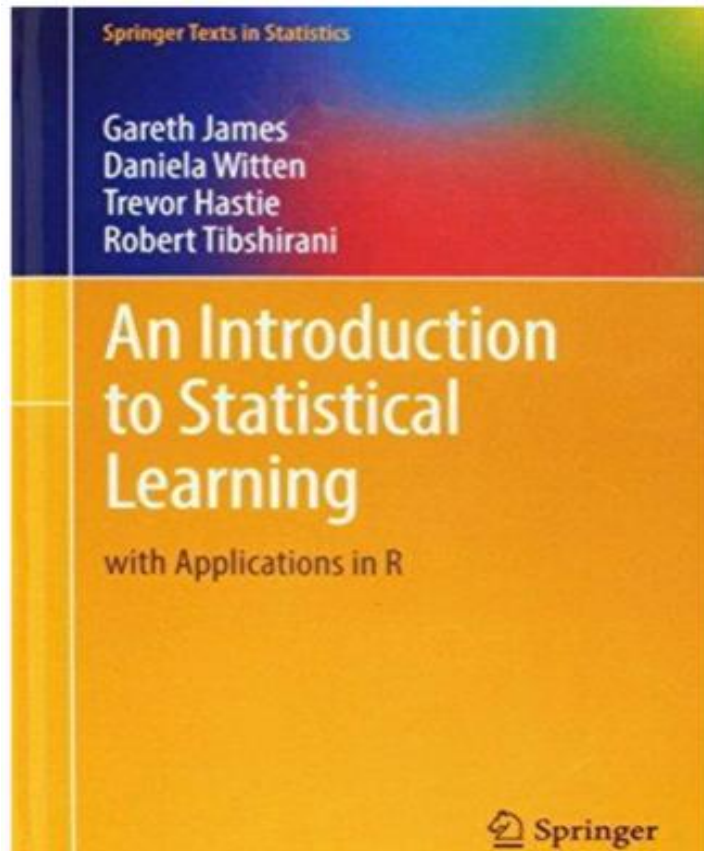
Zeitslot der
Synchronveranstaltung
14:30 – 17:45

Vertiefendes
Modell



Der Stoff wird sehr kondensiert vermittelt, so dass eine Literaturbegleitung zwingend erforderlich ist

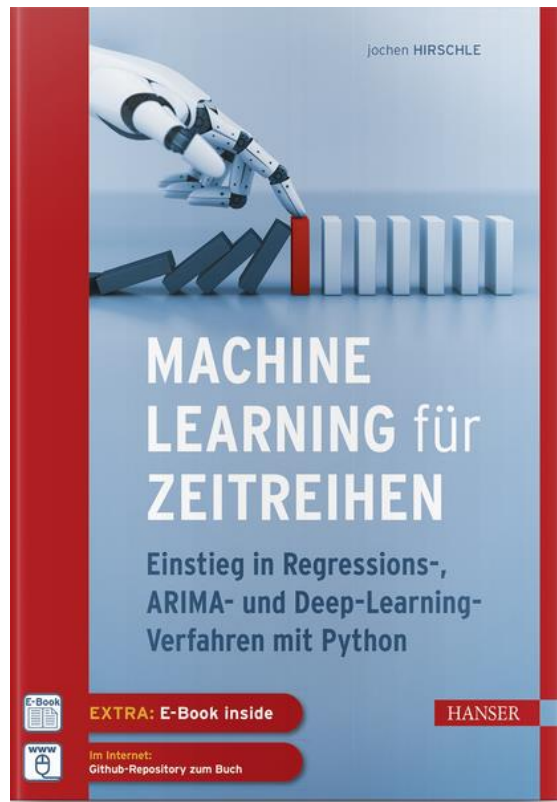
Literatur Methoden



- Viele der betrachteten Methoden sind im Begleitbuch ausführlich beschrieben.
- Algorithmen mit Verwendung neuronaler Netze werden mit eigenen Literaturquellen angekündigt
- In Bibliothek verfügbar
- Online: <https://www.statlearning.com/>
- Themen
 - Modellauswahl
 - Regression, Lasso
 - PCA
 - Splines, lokale Regression
 - K-Means- Hierachical Clustering

Der Stoff wird sehr kondensiert vermittelt, so dass eine Literaturbegleitung zwingend erforderlich ist

Literatur Umsetzung

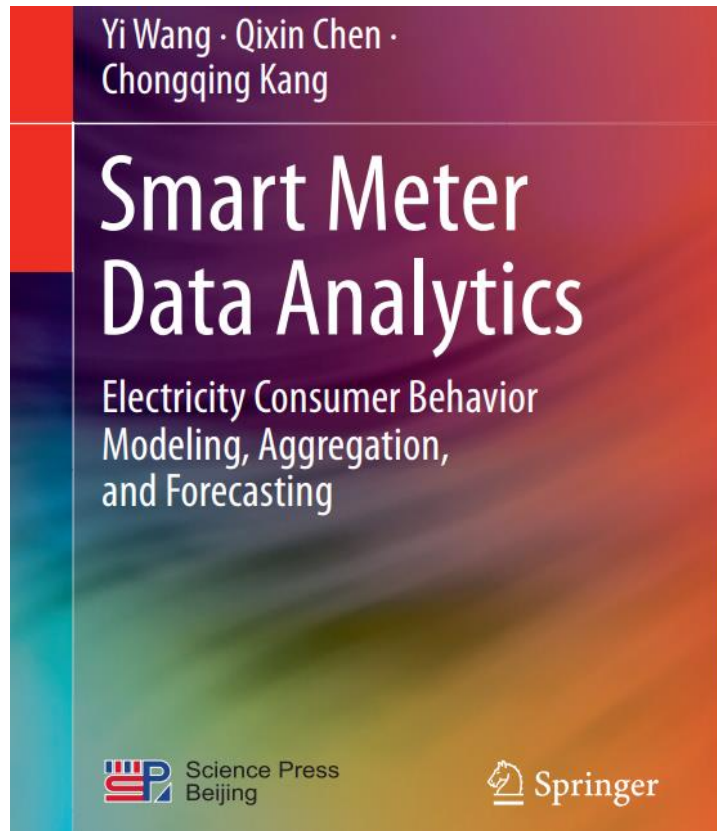


Machine Learning für Zeitreihen

- Einstieg in Regressions-, ARIMA- und Deep Learning-Verfahren mit Python. Inkl. E-Book
- [Jochen Hirschle](#) (Autor)
- Buch | Hardcover
- 277 Seiten
- 2020 | 1. Auflage Hanser, Carl (Verlag) 978-3-446-46726-2 (ISBN)
- Implementierung neuronaler Netze in Python für Zeitreihenanalyse
- Neuronale Netze

Der Stoff wird sehr kondensiert vermittelt, so dass eine Literaturbegleitung zwingend erforderlich ist

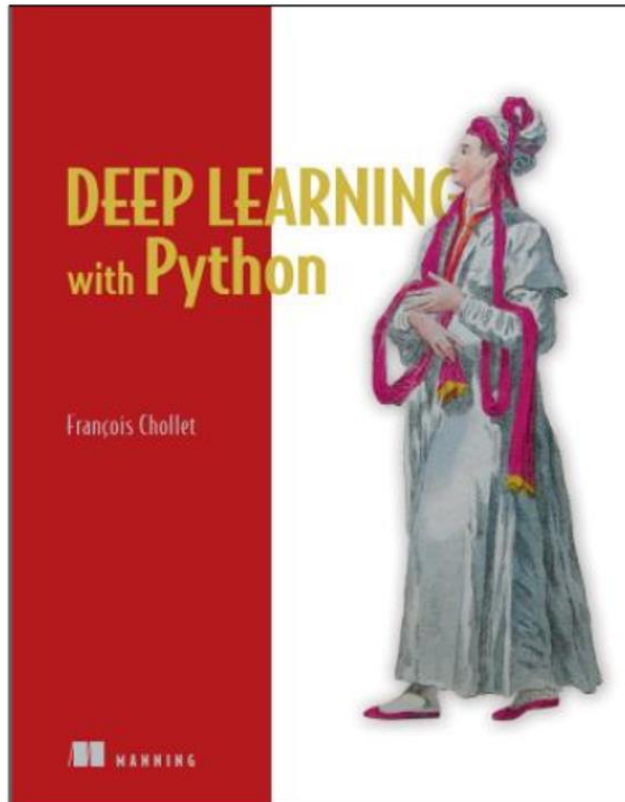
Literatur Anwendung



- ISBN 978-981-15-2623-7 ISBN 978-981-15-2624-4 (eBook)
- <https://doi.org/10.1007/978-981-15-2624-4>
- In Bibliothek als Onlineexemplar frei verfügbar
- Viele verschiedene Machine Learning Algorithmen auf Fragestellungen rund um Smart Meter Data Analytics umgesetzt

Der Stoff wird sehr kondensiert vermittelt, so dass eine Literaturbegleitung zwingend erforderlich ist

Literatur Modellierung mit tensorflow



- Umfangreiche anwendungsorientierte Abhandlung mit Keras/Tensorflow-Umgebung
- Leider Kapitel zur Regression ggü. Klassifikation unterrepräsentiert
- Aktuell: 2.te Auflage

Bezug: Bibliothek

<https://www.manning.com/books/deep-learning-with-python>

1. Einleitung

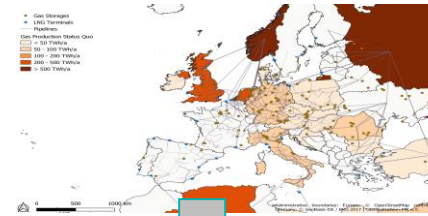
2. Aufbau

3. Prüfung



Der Gaspreis wird von einer Vielzahl von Einflussfaktoren bestimmt.

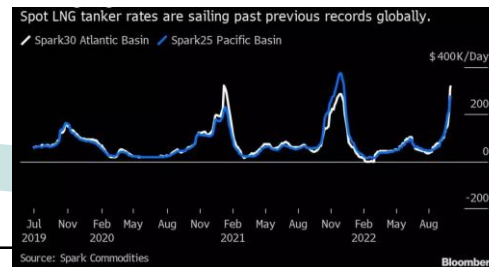
Infrastruktur



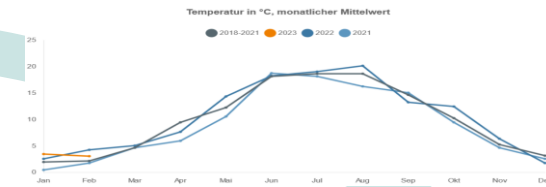
Gas-Preise



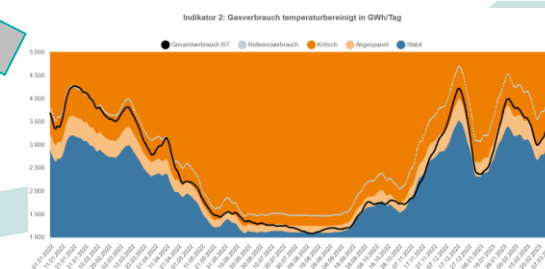
LNG-Frachtraten



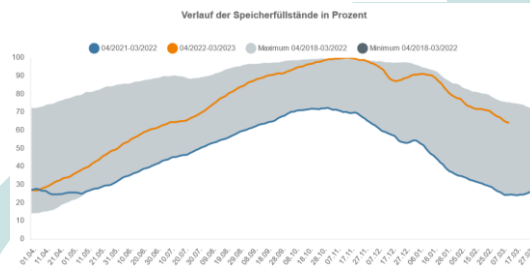
Temperatur



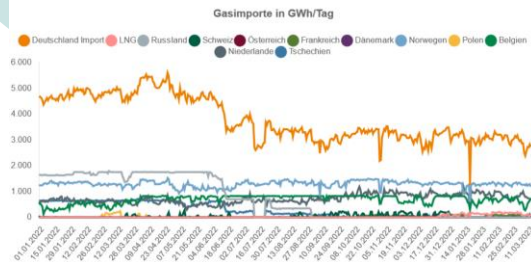
Nachfrage (Haushalte, Industrie, Kraftwerke)



Speicherfüllstände



Import/Export



Die Untersuchung erfolgt in zwei Abschnitten (Modellfamilie parametrische und KNN-Regressionsmodelle)

Energiedatenanalyse- Datamining

Datenvorbereitung

Deskriptive
Datenanalyse

Modellauswahl

Regressionsmodelle

Lineare Regression

Nichtlineare
Regression

PCA/PLS

Deep-Learning Modelle

Neuronale Netze

Hyperparameterfestlegung

Benchmarkmodell

Vertiefendes Modell

Für die Prüfungsleistung wird eine Seminararbeit angefertigt

■ Zielsetzung:

- Abbildung der Zusammenhänge in einem energiewirtschaftlichen Datensatz mit Hilfe der erlernten Methoden in Python
- Wissenschaftliche Dokumentation der Erkenntnisse im Rahmen einer Seminararbeit im jupyter-notebook

■ Vorgehensweise (sowohl in Block 1 als auch in Block 2):

1. Definition einer wissenschaftlichen Fragestellung
2. Auswahl der Untersuchungsobjekte
3. Deskriptive Analyse der Daten
4. Ableitung von Modell-Hypothesen zur Auswahl der Einflussfaktoren und Modelle
5. Modellparametrierung und Modellvalidierung
6. Modellauswahl
7. Energiewirtschaftliche Interpretation der Ergebnisse
8. Schlussfolgerung und Anwendungsnutzen

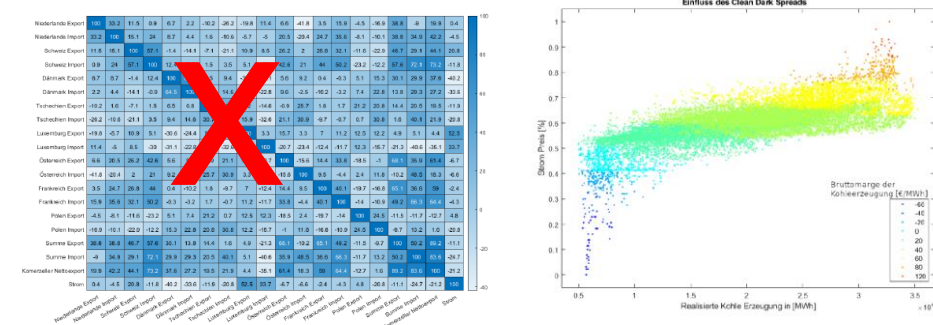
Struktur der Arbeit

- 1. Definition einer wissenschaftlichen Fragestellung
 - Kurzes Intro zum Day-Ahead-Markt und energiewirtschaftlicher Relevanz des Untersuchungsobjektes
 - Konkretisierung der Aufgabenstellung und Gliederung der Arbeit (Herangehensweise)
- 2. Auswahl der Untersuchungsobjekte und deskriptive Analyse
 - Ableitung möglicher relevanter Objekte aus technoökonomischen Zusammenhang
 - Auswahl der Datenquellen und deren Zeitreihen
 - Strukturierung der Daten (was wird in welcher Form dokumentiert (Granularität/Auflösung, Fristigkeit))
 - Ableitung der relevanten Zeitreihen

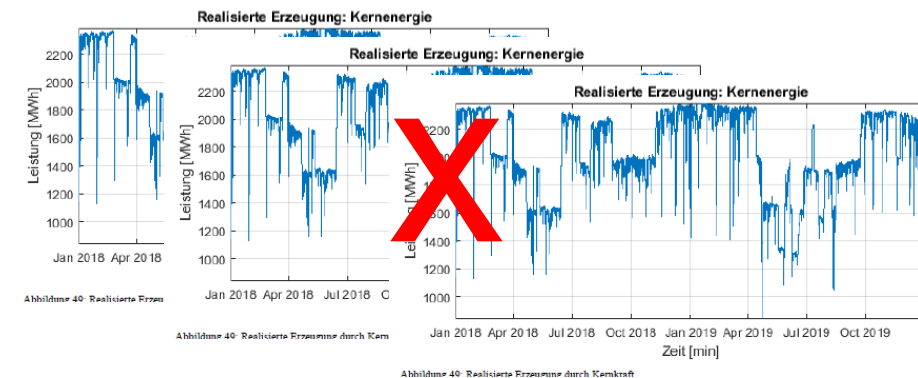
3. Deskriptive Analyse der Daten um Wirkungszusammenhang abzuleiten

- Darstellung visueller und tabellarischer Auswertungen (z.B. Lage und Streuungsparameter), um einen Überblick der Daten zu erhalten
- Analyse im Zeitablauf (inbs. visuelle Information in unterschiedlicher Auflösung (Jahr, Monat, Tag, Stunde))
- "Multivariate Darstellung (Korrelation, Scatter,...) der Ausgangszeitreihen, um Zusammenhänge zu analysieren (bspw. Trennung nach WT und WE oder Hoch und Niedrigpreisphasen oder Im und Exportsituationen oder ...)"
- Feature Engineering: ggf. weitergehende Transformierung/Aggregation der Daten, um detailliertere Analysen des Wirkungszusammenhangs vorzunehmen (Spreads, PCA-Analyse,...)
- Ableitung erster Hypothesen zur Relevanz der einzelnen Zeitreihen und mögliche Aggregationen in der Modellbildung

Bitte hier intelligente und aussagekräftige Analysen



Klasse statt Masse

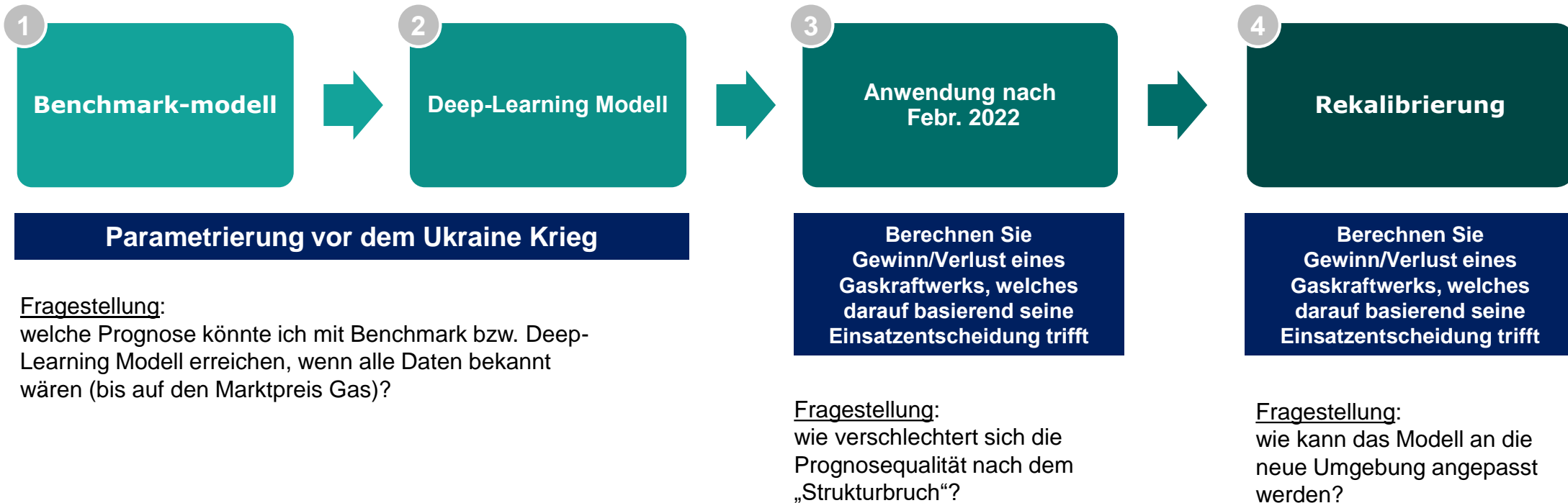


Kein Bild ohne Interpretation!

4. Ableitung von Modell-Hypothesen zur Auswahl der Einflussfaktoren und Modelle

■ Zielsetzung:

- Das Modell soll zum Einsatz der Day-Ahead-Planung eingesetzt werden



Benchmarkmodell

- Das Modell soll möglichst kompakt und in seiner Wirkungsweise noch nachvollziehbar bleiben
 - Einsatz der erlernten Techniken (Vorlesung 1 - 5) um ein Grundmodell anzuwenden
- Benchmarkmodell aufgrund einer plausiblen Modellhypothese mit Hilfe der multiplen linearen Regression
 - Feature Engineering
 - Modellaufbau
 - Modellauswertung
- Zu dem untersuchten Modell gehört:
 - Modellparametrierung: Diskussion der Parameterschätzer (Standardabweichung; Intervallschätzer; p-Wert)
 - Modellvalidierung: Analyse des Fits, resultierender Abhängigkeiten und der Residuen
 - Energiewirtschaftliche Einordnung: Darstellung der Schwächen und Motivation der weiteren Faktoren

Deep Learning Modell

- 2 Deep Learning Modelle aufgrund plausibler Modellhypothesen
 - Einsatz der erlernten Techniken (Vorlesung 6 - 10) um zwei alternative „vertiefende“ Modelle anzuwenden
- Zum untersuchtem Modell gehört:
 - Sinnvolles Feature-Engineering (bitte nicht einfach sinnlos aus Daten „reinschmeissen“)!
 - Hyperparameteroptimierung!
 - Modellvalidierung: Analyse des Fits, resultierender Abhängigkeiten und der Residuen
 - Modelldarstellung in Tensorboard
- Energiewirtschaftliche Einordnung:
 - Darstellung der Schwächen und Motivation der weiteren Faktoren
 - Vergleich zum Benchmark-Modell auf Basis von Testdaten

Anforderungen sowie Form und Aufbau der Arbeit

- Struktur und Aufbau
 - Siehe Struktur Vorgehensweise
- Abgabeform:
 - Dokumentiertes Jupyter-Notebook
- Abgabe auf ILIAS einstellen.
- Gruppen: 2-3 Personen pro Gruppe
- Gruppeneinteilung: t.b.d.

**Frist Gesamtdokument:
10.09.2023 22:00**

Bewertungskriterien (vorbehaltlicher Änderungen)

Bewertungskriterien	Relevanz
Form, Aufbau, Vollständigkeit, Abbildungen, Verzeichnisse, Stil, Sprache, Rechtschreibung	15%
Inhalt, Substanz, roter Faden, energiewirtschaftliche Einordnung	20%
Modellverständnis: Modellkomplexität, Kreativität, Modellauswahl, Anwendung, Umsetzung in Code	35%
Einsatz der Technik: Parameterschätzung, Hyperparametersetzung, Validierung und Samplingtechnik, „technisches Verständnis“ der Ergebnisse, Umsetzung in Code	30%