Machine Learning Applications

Energy Forecasting – Electric load forecasting by the example of a machine tool

Prof. Dr.-Ing. M. Weigold, Bastian Dietrich, M.Sc.

Darmstadt | 31.01.2020

ETA FABRIK

ENERGIEEFFIZIENZ, TECHNOLOGIE UND ANWENDUNGSZENTRUM

Agenda

Agenda

The ETA-factory

https://www.youtube.com/watch?v=agmGWmSSLG4

Research Group ETA

Energy Technologies and Applications in Production

TECHNISCHE UNIVERSITÄT DARMSTADT → FACHBEREICH MASCHINENBAU → PTW TU DARMSTADT → ETA FABRIK

Additive Manufacturing

Machine Tools and Industrial Robots

Machining Technology

Manufacturing Technology
Prof. Dr.-Ing. M. Weigold

Main Fields of Application
Mechanical Engineering | Automotive | Aerospace | Dental Technology

Production Organization
Prof. Dr.-Ing. J. Metternich

Energy Efficiency in Production
Prof. Dr.-Ing. E. Abele

CiP | Center for Industrial Productivity

Management of Industrial Production

ETA | Energy Technologies and Applications in Production

eta-fabrik.de

The Challenge

Holistic increase of the Energy Efficiency

Today: Isolated optimization of different sub-systems of a factory

Building

25%

Quelle: Prof. Dipl.-Ing. J. Eisele

Building services

20%

Machine

30%

Savings

< 30 %

Our vision: Holistic factory optimization including all sub-systems

Interaction of:

- **Machines**
- **Building services**
- **Buildings**

Synergies by energy controlling and recovery measures

Potential

~ 40 %

The interdisciplinary ETA research group

Energy Technologies and Applications in Production

ETA FABRIK

Research Fields: Energy Efficiency // Energy Flexibility // Resource Efficiency in production

WANTED: STUDENTS AND HIWIS

Software-oriented

Artificial Intelligence for Energy Systems

- Autonomous measuring systems
- Optimum control
- Optimum scheduling
- Intelligent Agents
- Energy Assistance Systems

Simulation of Energy Systems

- Object-oriented modeling
- System Identification
- Digital Twin
- Virtual Commissioning
- Digital Planning Tools

Energy management & monitoring

- Prediction Systems
- Energy Performance Indicators
- Condition & Quality Monitoring
- ICT-Infrastructure
- Data Acquisition Systems

Hardware-oriented

Energy-optimized production machine

Machining // Cleaning // Heat Treatment // Process Engineering

- Energy Recuperation
- New & existing project planning
- Cross-sectional Technology Benchmarks
- Energy Analyses

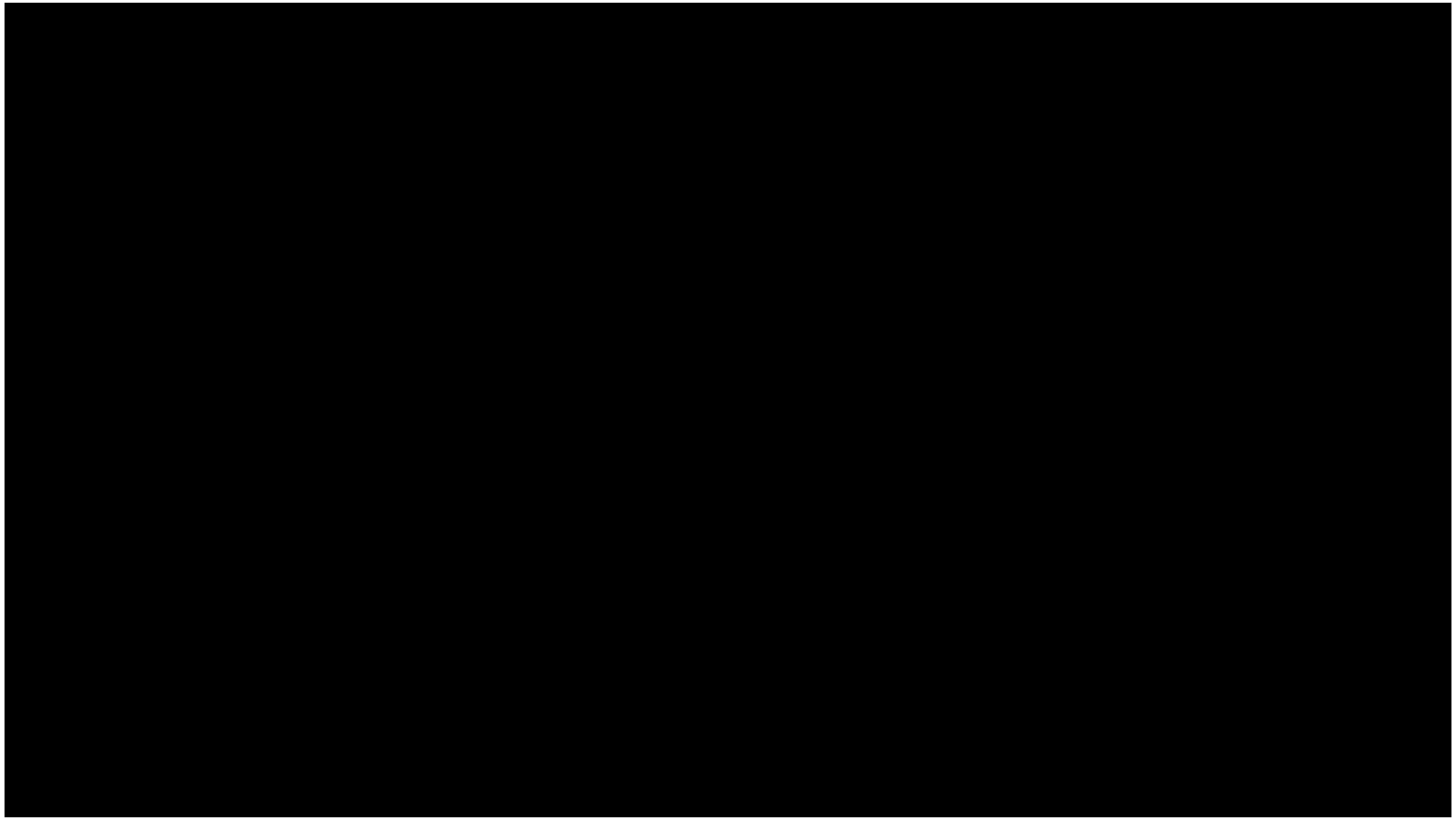
Energy-optimized supply systems

Cooling // Heating // Compressed Air // Air Conditioning // Ventilation

- Energy Recovery, Storage & Networks
- Topologies & Dimensioning
- Technology Benchmarks
- Energy system analyses

Transfer // Education

Projekt SynErgie



https://www.youtube.com/watch?v=vo8w4sOBv_4

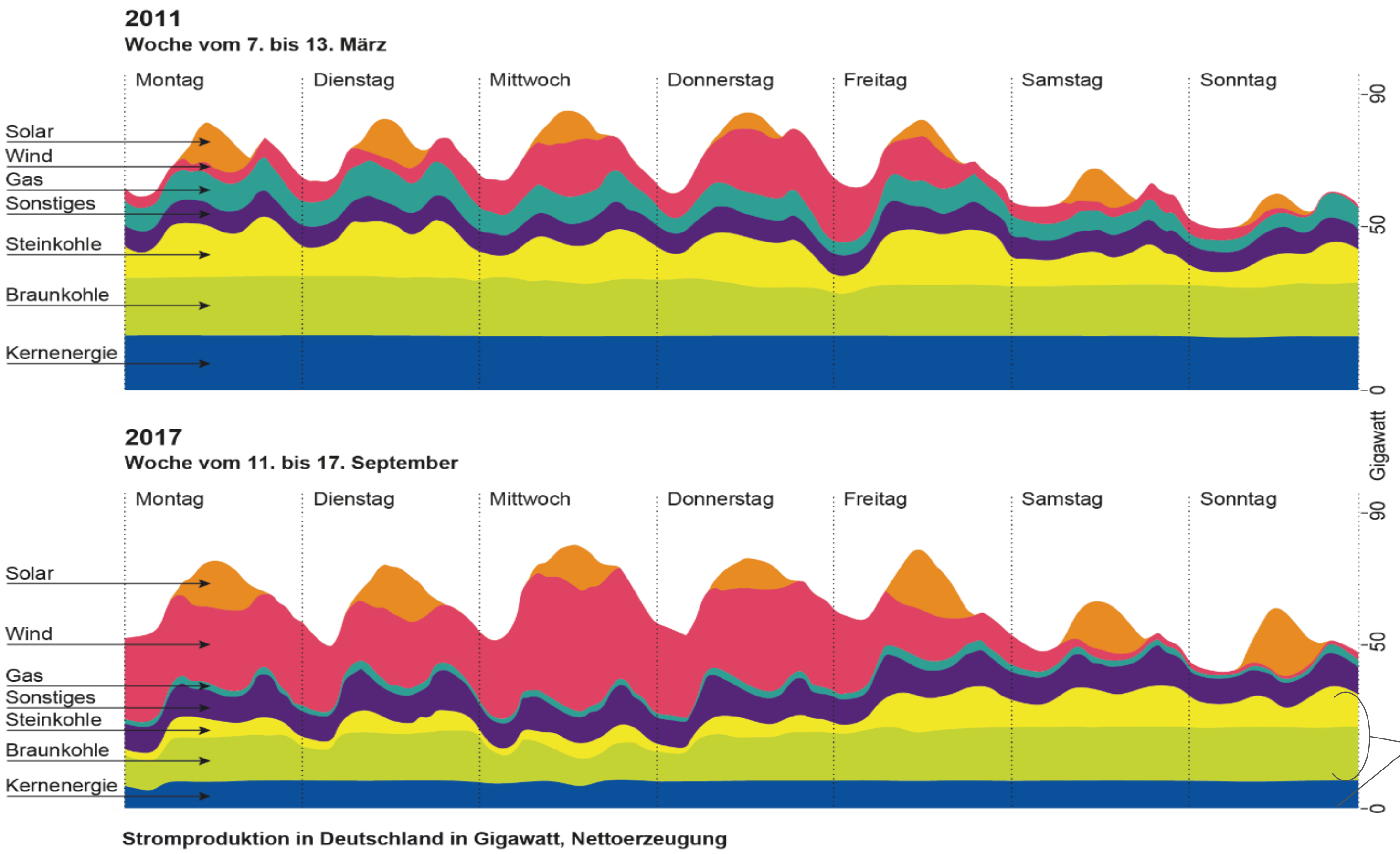
Agenda

Learning objectives

- General understanding: classification of forecasting in Machine Learning, delimitation of prediction
- Know the peculiarities of time series
- Know tools and algorithms for time-series-specific data preprocessing and forecasting
- Know metrics for the evaluation of the forecast quality and be able to name advantages and disadvantages in each case
- Be able to name forecasting applications in mechanical engineering and know typical problems and their solutions in implementation

Agenda

Progress of the Energy Transition since Fukushima



2019: 40,4% Renewable Energy (RE)
2011: 19 %

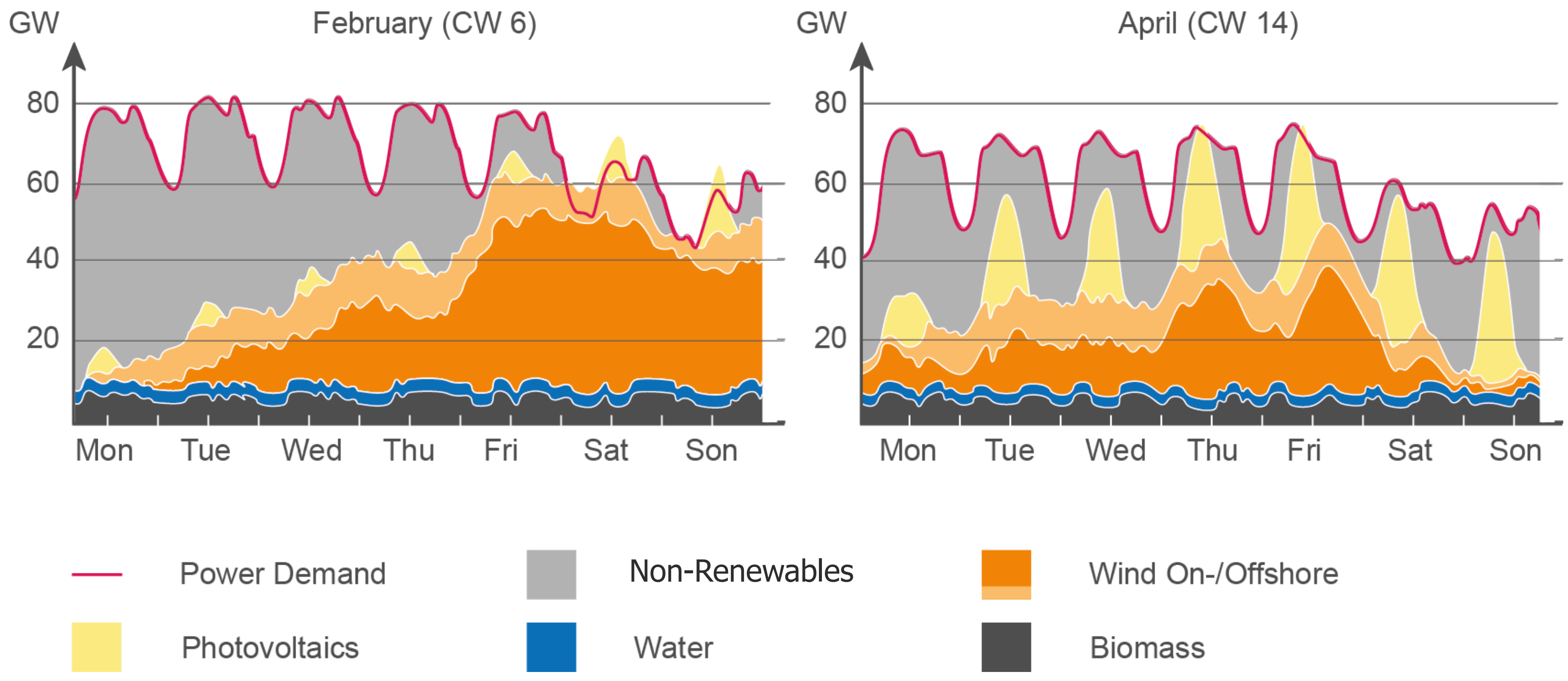
1 Hj.2019:
Contribution of RE to gross domestic electricity consumption: 290 billion kWh

2022: last nuclear power plant shut down

2038: last coal power plant shut down

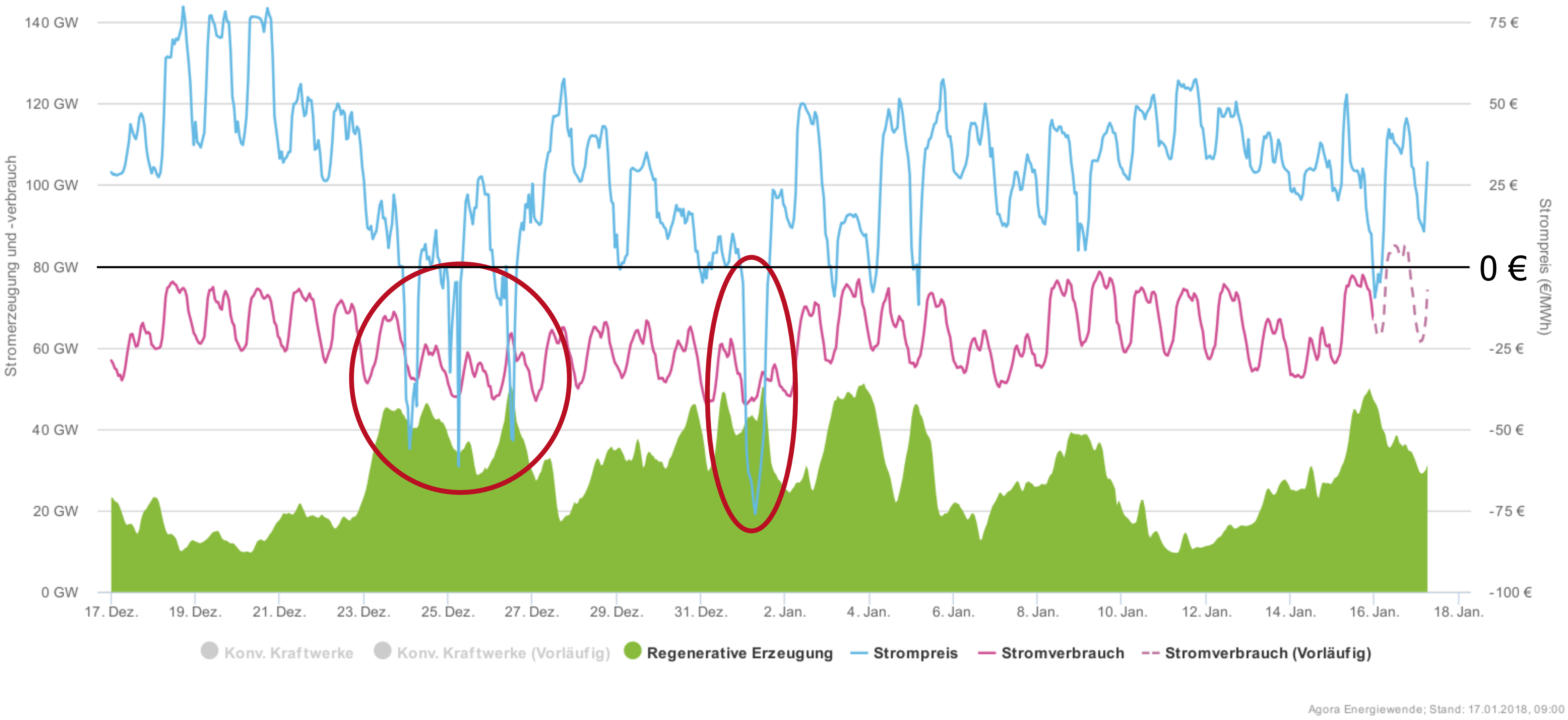
Challenges of the Energy Transition – Electric Power

Differences in Electric Power Demand and Electric Power Supply



Source: PTW, based on Agora Energiewende 2012

Electricity prices fluctuate due to demand and supply



- Usual fluctuation of electricity prices by >20€/MWh/d
- Negative electricity prices

Energy forecasting in industry – why?

- Pricing structures and price fluctuations at the electricity market **force the industry to adapt their electric load profile** to the electricity supply
- The load profile of the factory is strongly influenced by the load profiles of the production machines inside the factory.
- Knowing the future load profile of the production machines enables us to control the load profile of the factory.

Step 1: Create energy transparency

→ Step 2: Forecast the future load profile

→ Step 3: Put the knowledge to use

- Demand side management
- Peak load shaving

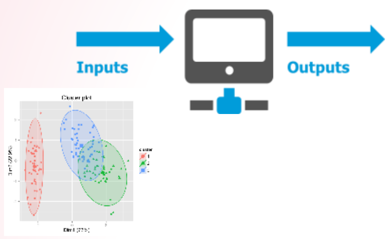
Why use Machine Learning for forecasting?

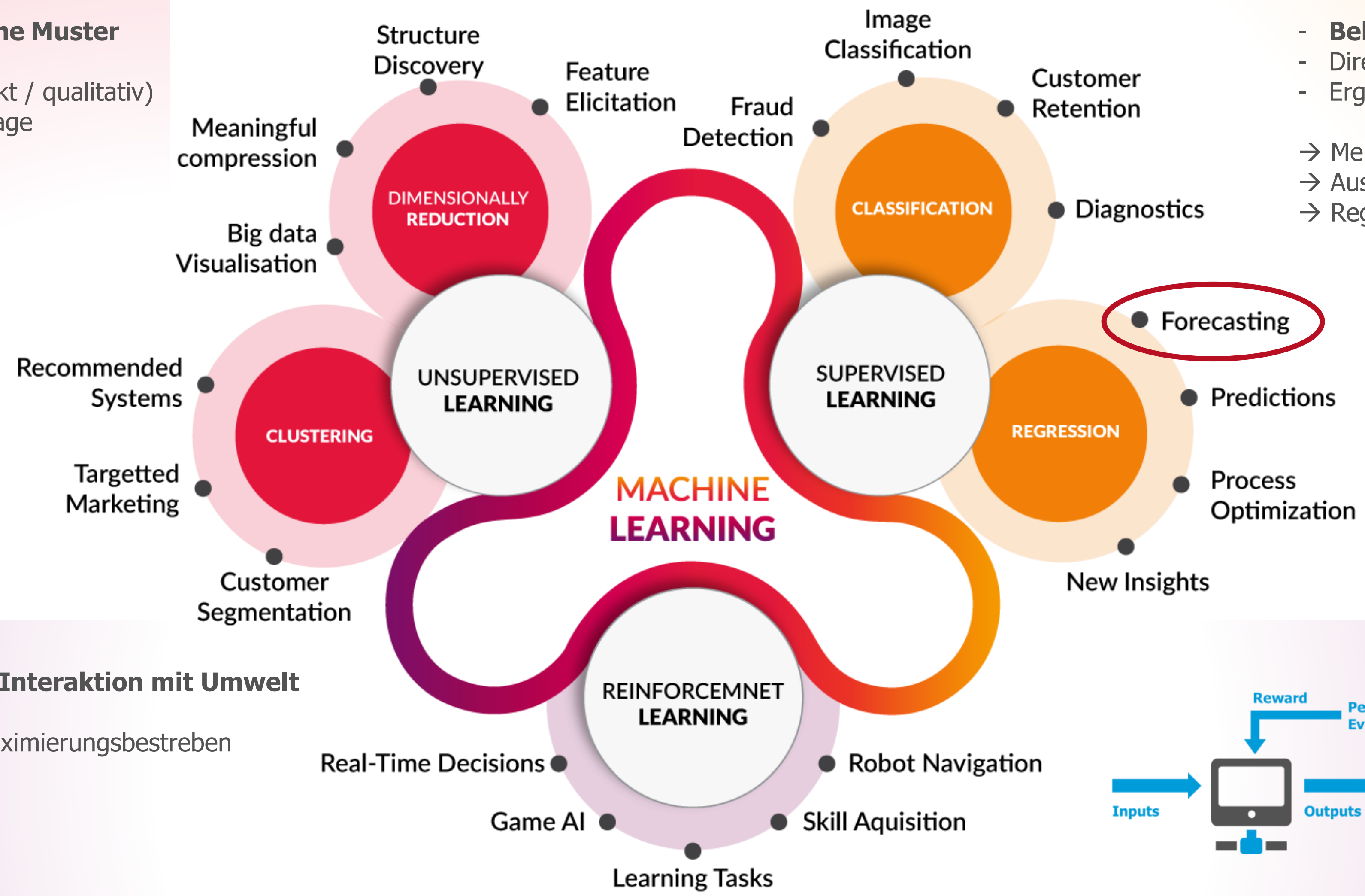
Advantages of Machine Learning	Possible drawbacks
Rising data availability through Industry 4.0 → Large data base	Data quality issues, installation of sensors necessary
More and better algorithms and rising computational power	Machine Learning expertise required (selection of algo's, parameters, data preparation)
Less domain expert knowledge required	Less model understanding
Forecasting based on real-time data (Forecasting horizon and sensitivity)	Continuous data acquisition and preparation necessary

Energy forecasting in the Machine Learning ecosystem

- **Unbekannte, verborgene Muster erlernen**
- Kein Feedback (nur indirekt / qualitativ)
- Keine spezifische Vorhersage

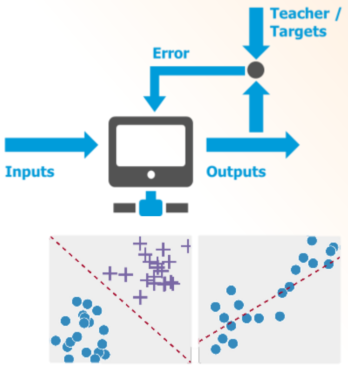
→ Gruppierung / Clustering





- **Bekannte Muster erlernen**
- Direktes Feedback
- Ergebnis/Zukunft vorhersagen

→ Merkmalsunterscheidung
→ Ausreißer-Erkennung
→ Regressionen



- **Muster erlernen durch Interaktion mit Umwelt**
- Entscheidungsprozesse
- Belohnungssystem mit Maximierungsbestreben
- „künstliche Intelligenz"

Quelle: Smartbasegroup.com

Forecasting vs. prediction

Definition of terms

Prediction:

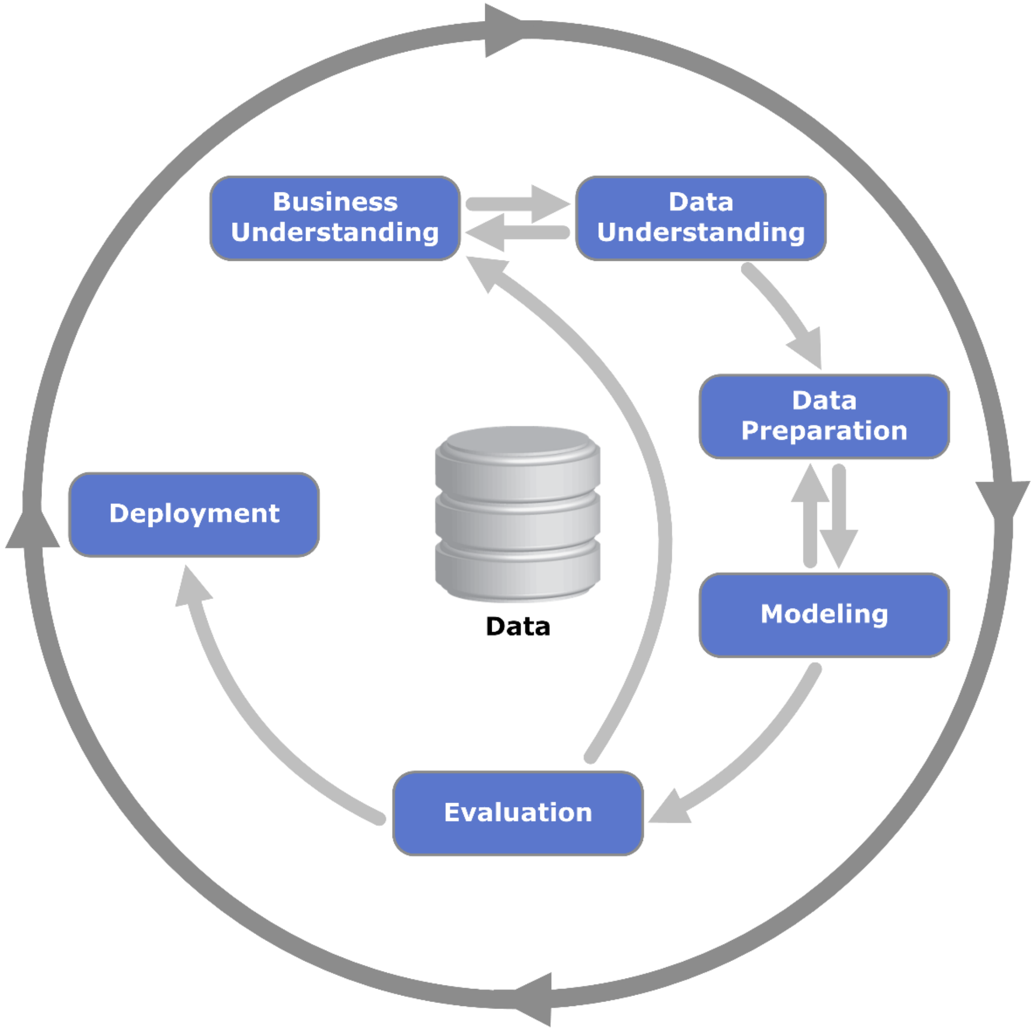
- General term → predict an unknown value from known inputs
- Example: Prediction of the net income of households from house location, house size, number of rooms, …

Forecasting:

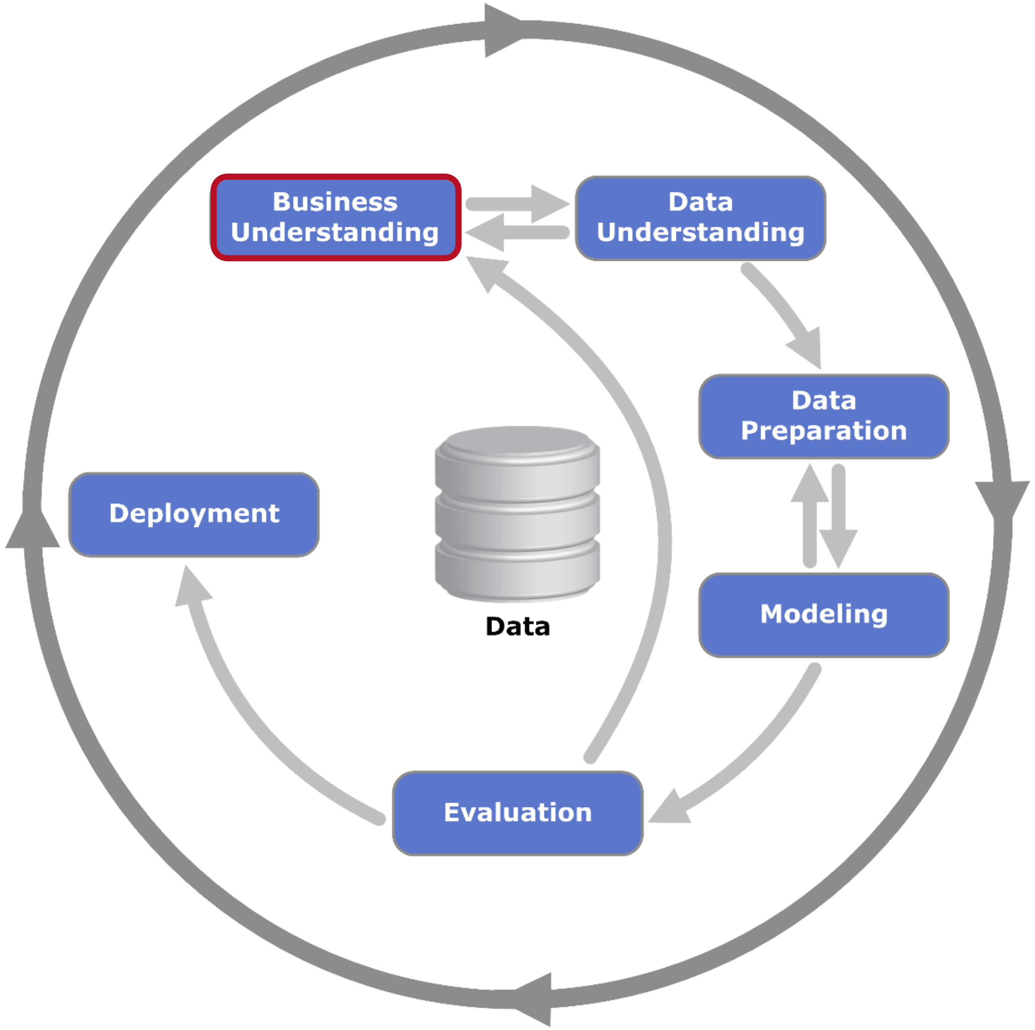
- Time related → forecast the future values of a time series
- Example: Weather forecast of tomorrow from current and past weather conditions, time of year, …
- Challenge of forecasting in Machine Learning:
  - Feature engineering gets a second dimension: Time
  1. Model the exogenous, non-temporal features (the feature model)
  2. Model the historical, temporal features (the temporal model)
  - Therefore, the feature set often becomes much larger than in regular Machine Learning problems, which can evoke the "curse of dimensionality"

Agenda

The CRISP-DM model



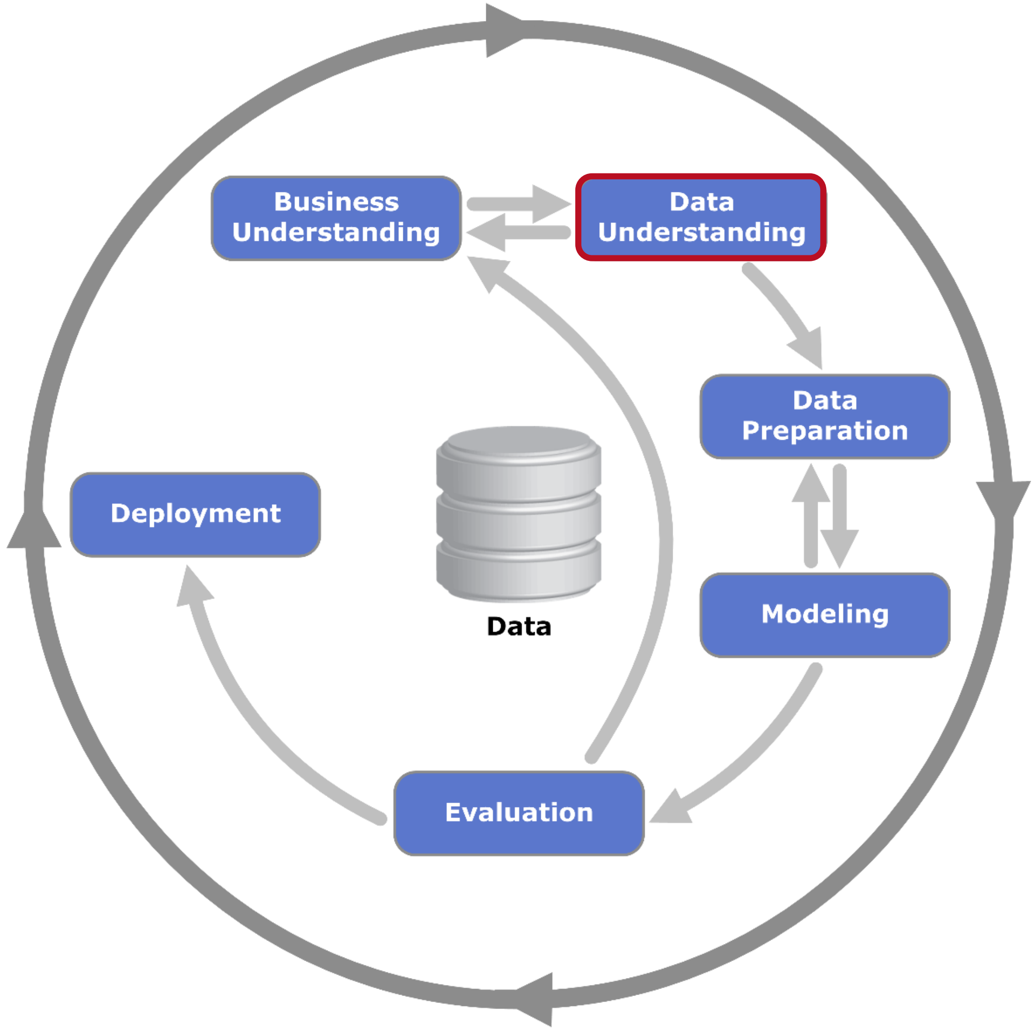
The CRISP-DM model – Business understanding



Use case: Load forecasting of a production machine

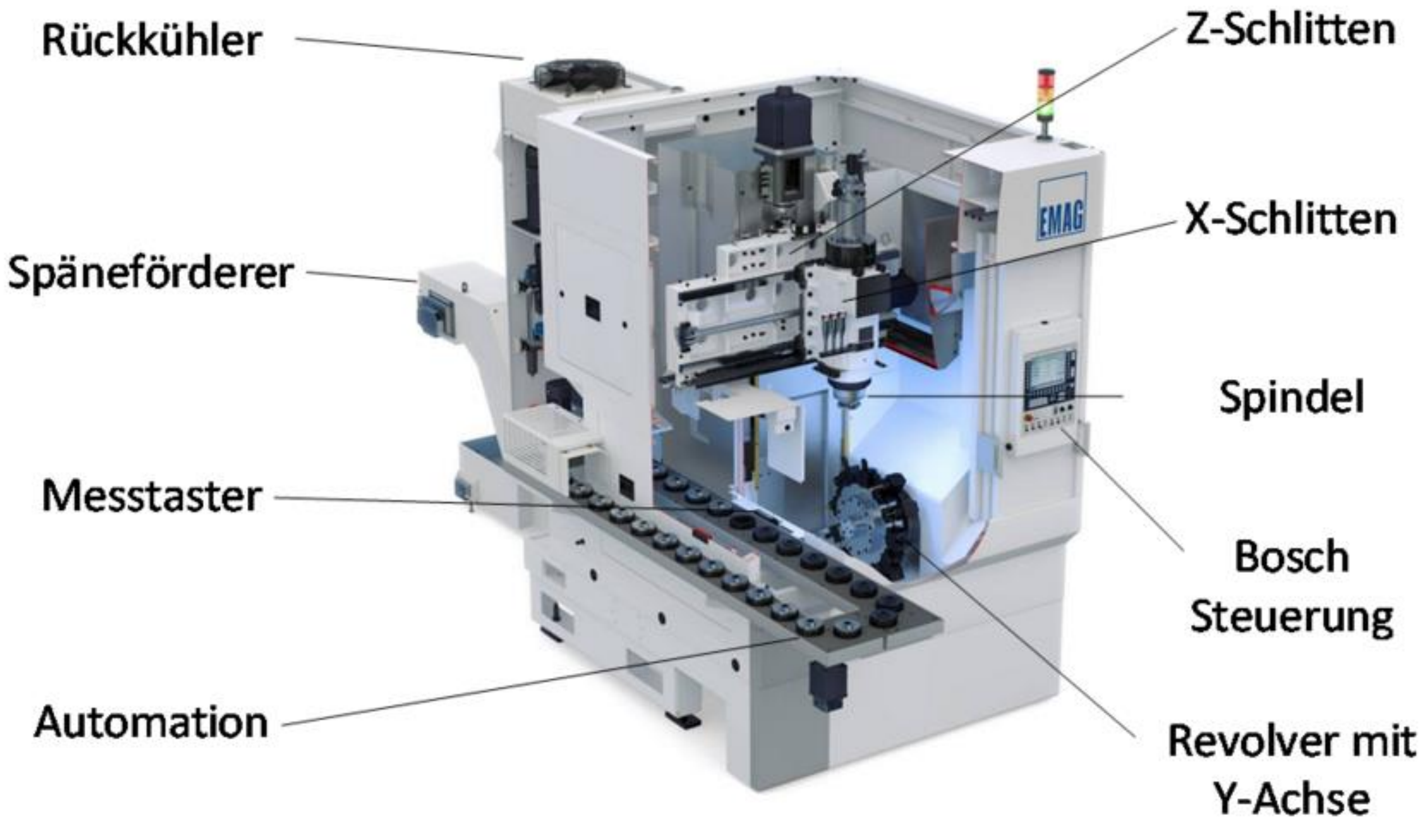
- **Business understanding:** Volatile energy market leads to need for better understanding of the demand side in industry to enable demand side management.
- **Business objective**: Forecast the active power of the machine tool EMAG VLC 100 GT
- **Model requirements**:
  - Forecasting horizon: 15 minutes
  - Time intervals: 1 second
  - The model requires minimal user input

The CRISP-DM model – Data understanding



The CRISP-DM model – Data understanding

The machine tool EMAG VLC 100 Y

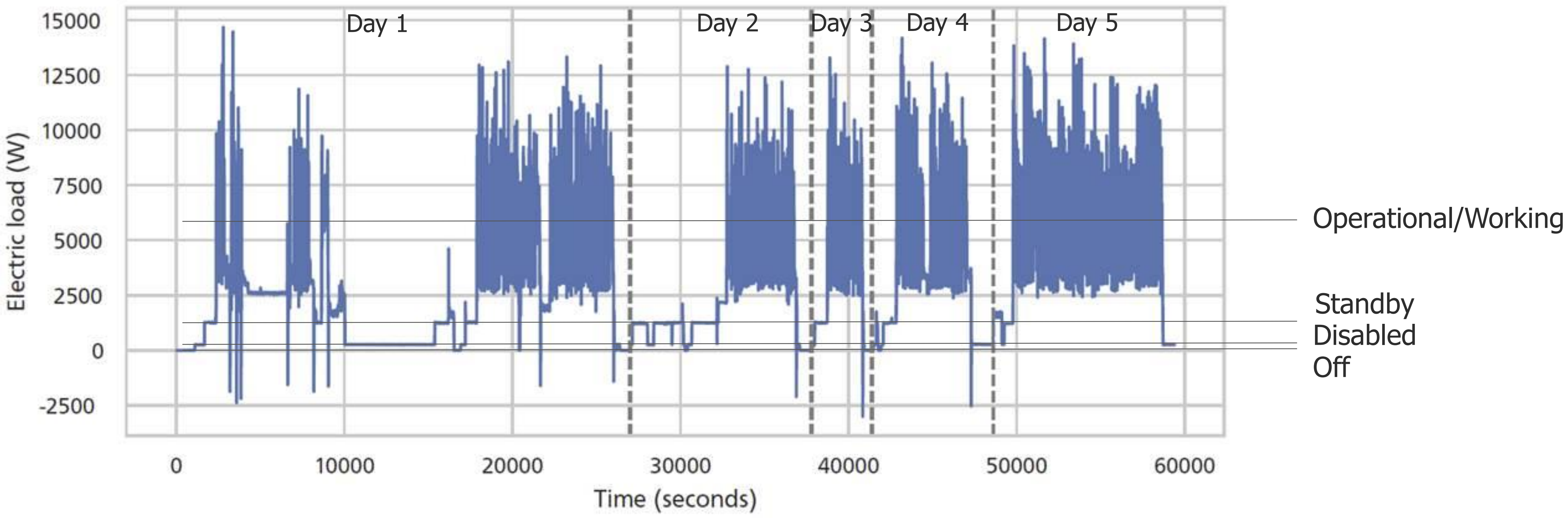


Distinguish the different kinds of power in machine tools:

- Installed power – Dimensioning of the power connection
- **Active power (P) – Real, time-dependent power input to the machine**
- Reactive power (Q)
- Apparent power (S)
- Mechanical power – power output to the part (power after losses)

The CRISP-DM model – Data understanding

Active power of the machine tool - Data set



The CRISP-DM model – Data understanding

Breakdown of machine tool power consumption

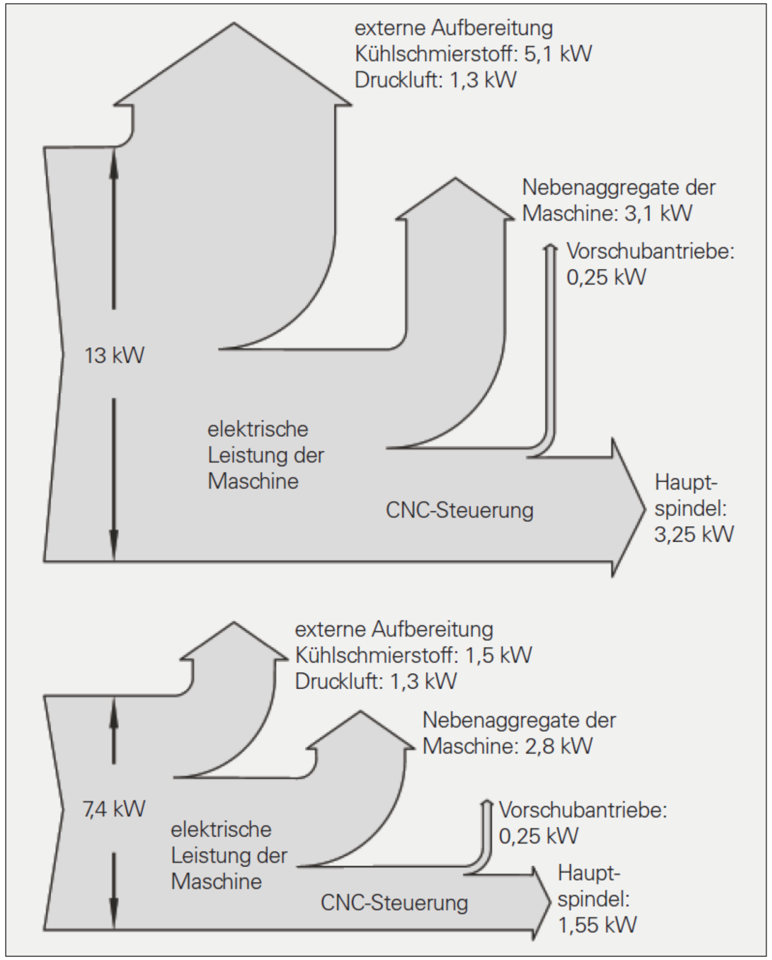
Leistungsbedarf

Der Energiebedarf verteilt sich auf die Verbrauchergruppen

- Kühlschmierstoffaufbereitung,
- Drucklufterzeugung,
- elektrisch gespeiste Nebenaggregate
- CNC-Steuerungspaket mit Hauptspindel

und Vorschubantrieben.

Gehäuse für die Betrachtung
des Leistungsbedarfs eines Fräsprozesses

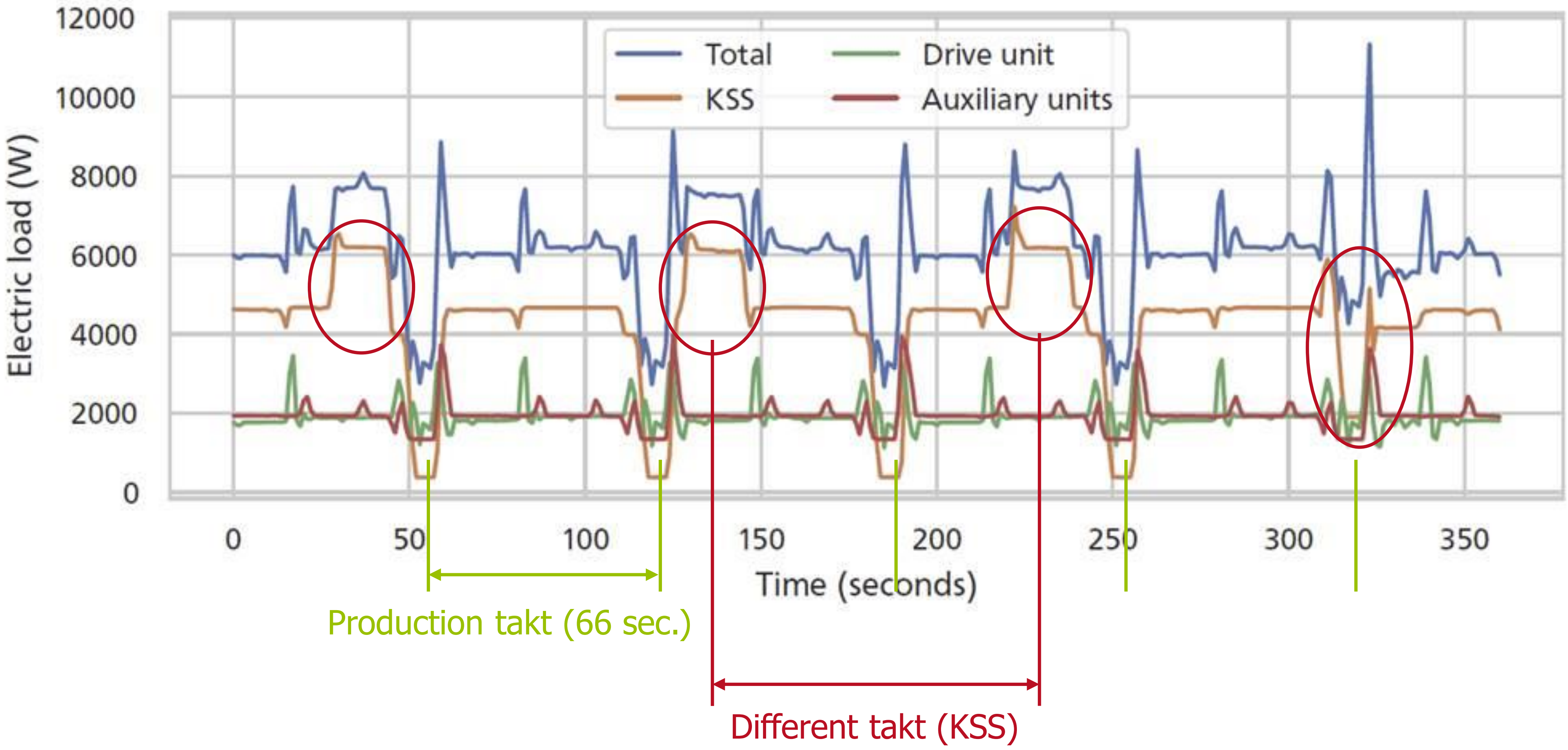


Mittlerer Leistungsbedarf für die Fertigung eines Gehäuseteils,
oben: Schruppen, unten: Schlichten

Source: Heidenhain

The CRISP-DM model – Data understanding

Breakdown of machine tool power consumption, time dimension



The CRISP-DM model – Data understanding
Zoom in zoom out

Drive unit

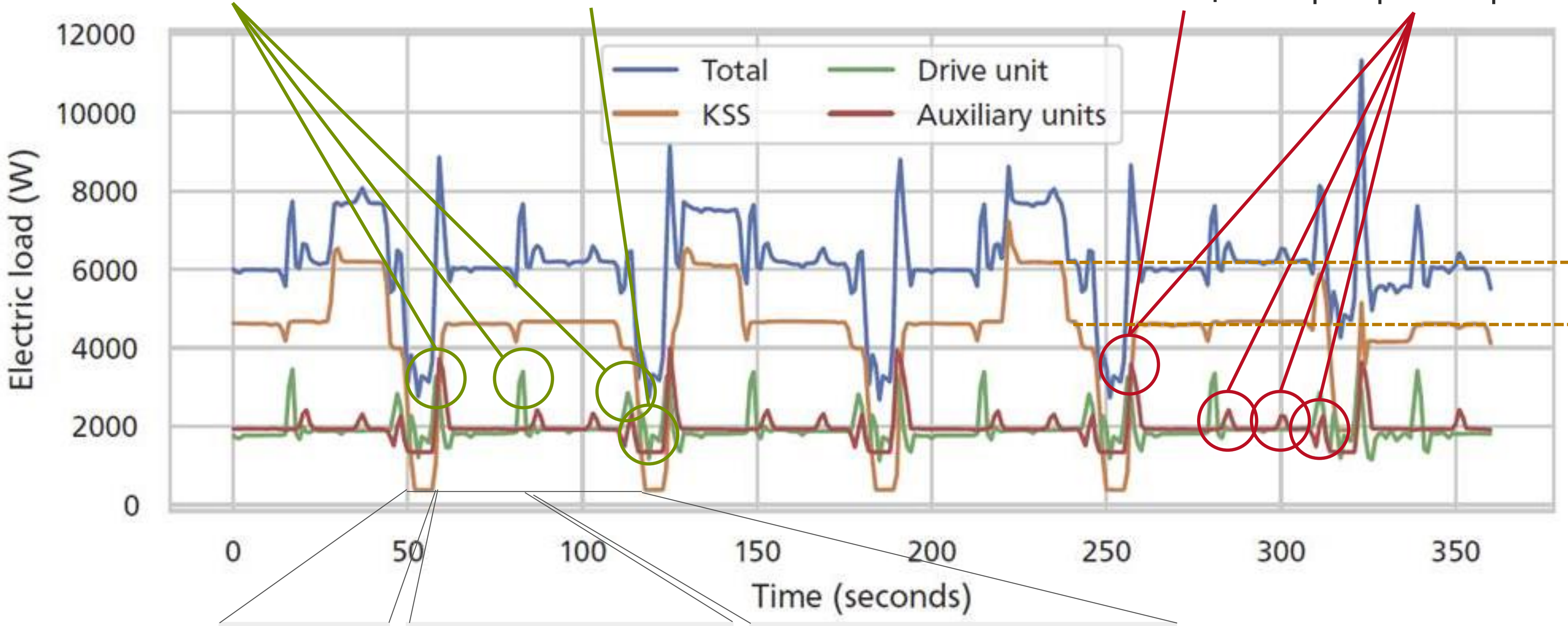
Peaks from rapid Z axis movement

Spindle 3 accelaration/decelaration

Auxiliary units

Suction startup

Hydraulic pump startup



KSS (Cooling lubricant)

Magnetic separator

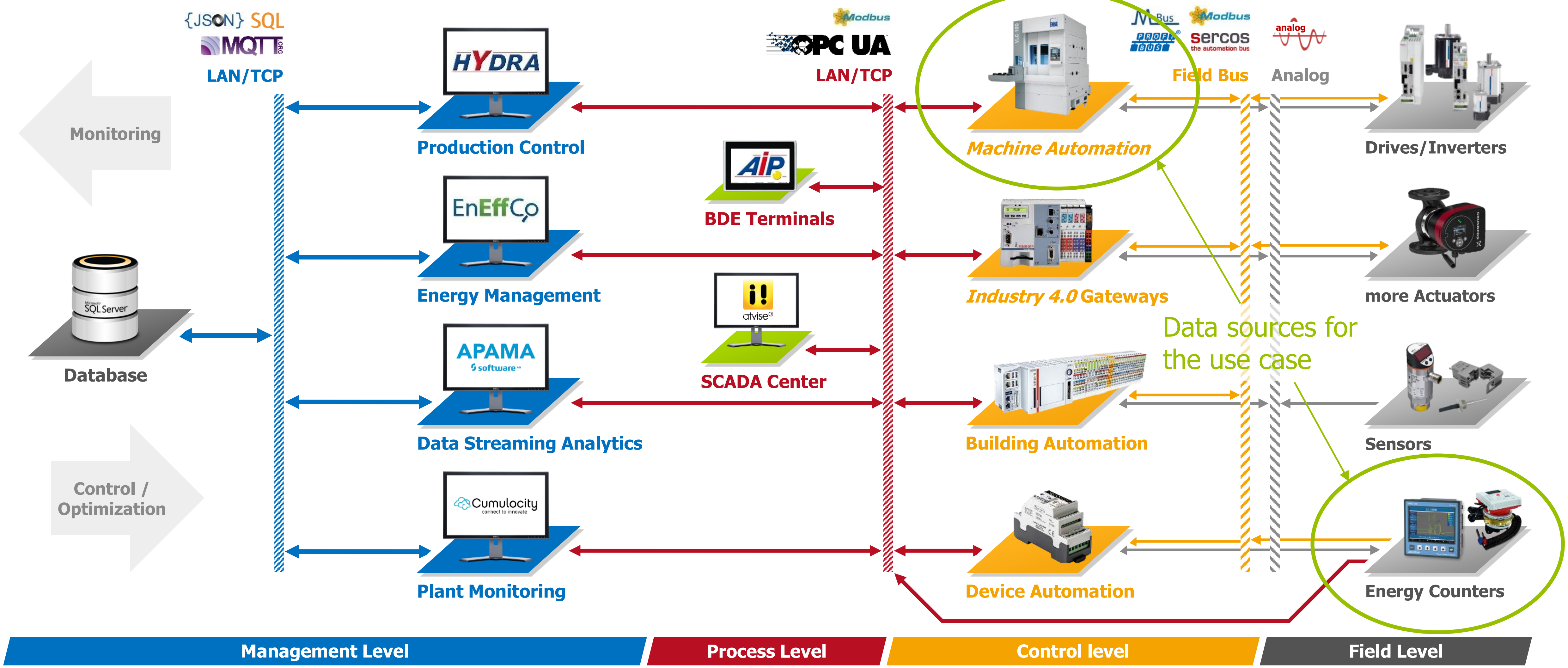
KSS pump

Production cycle

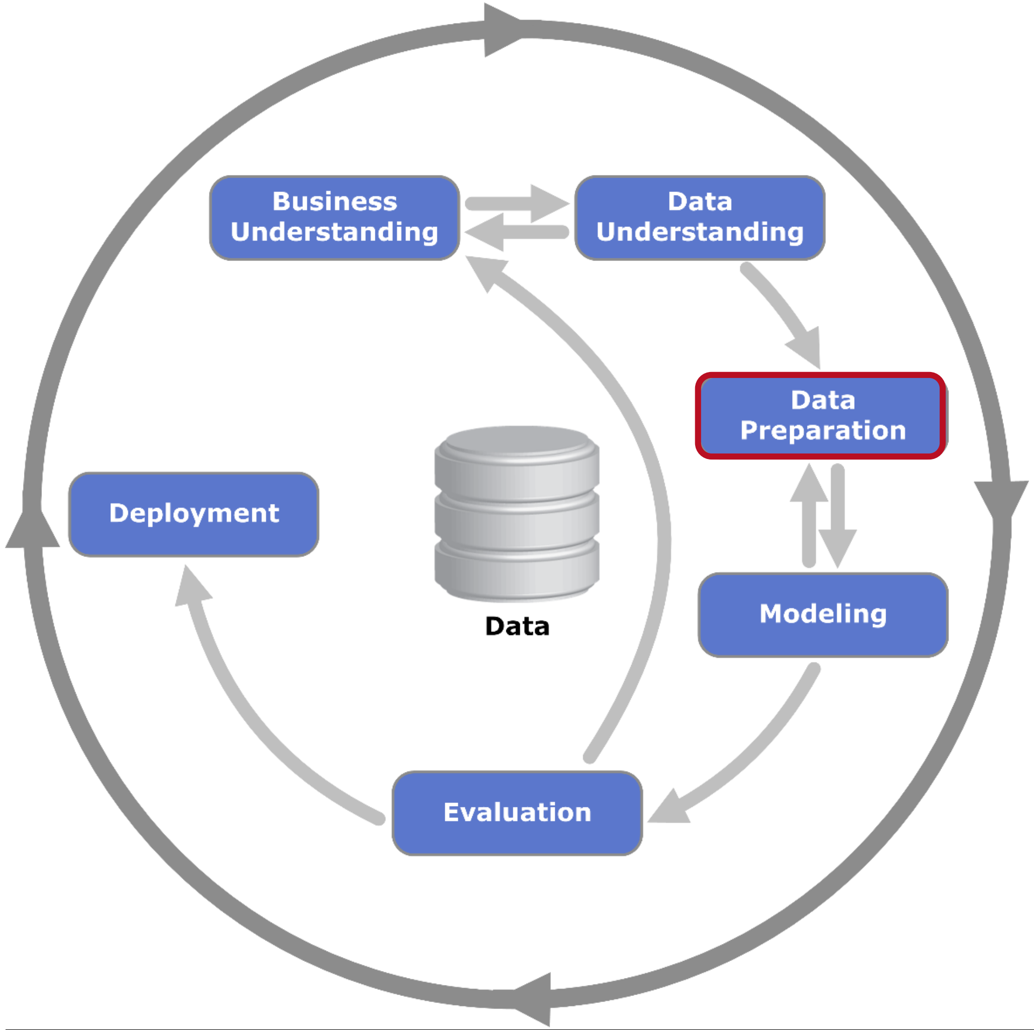
(Un-) Loading	Grinding step 1	Grinding step 2

The CRISP-DM model – Data understanding

Data Flows and Interfaces in the ETA-Factory



The CRISP-DM model – Data Preparation



The CRISP-DM model – Data understanding

Peculiarities of time series

- The values have a strong time dependency (autocorrelation)
  - Different splitting into training, validation and test data required
    - Test data must always be future values to ensure generalization capability of the model
    - No shuffling to prevent mixing future and past values
  - Preparation for supervised learning needed
    - Target must be time shifted so that the model learns the future behavior from the current/past inputs
  - Different feature engineering required
    1. Model the exogenous, non-temporal features (the feature model)
    2. Model the historical, temporal features (the temporal model)

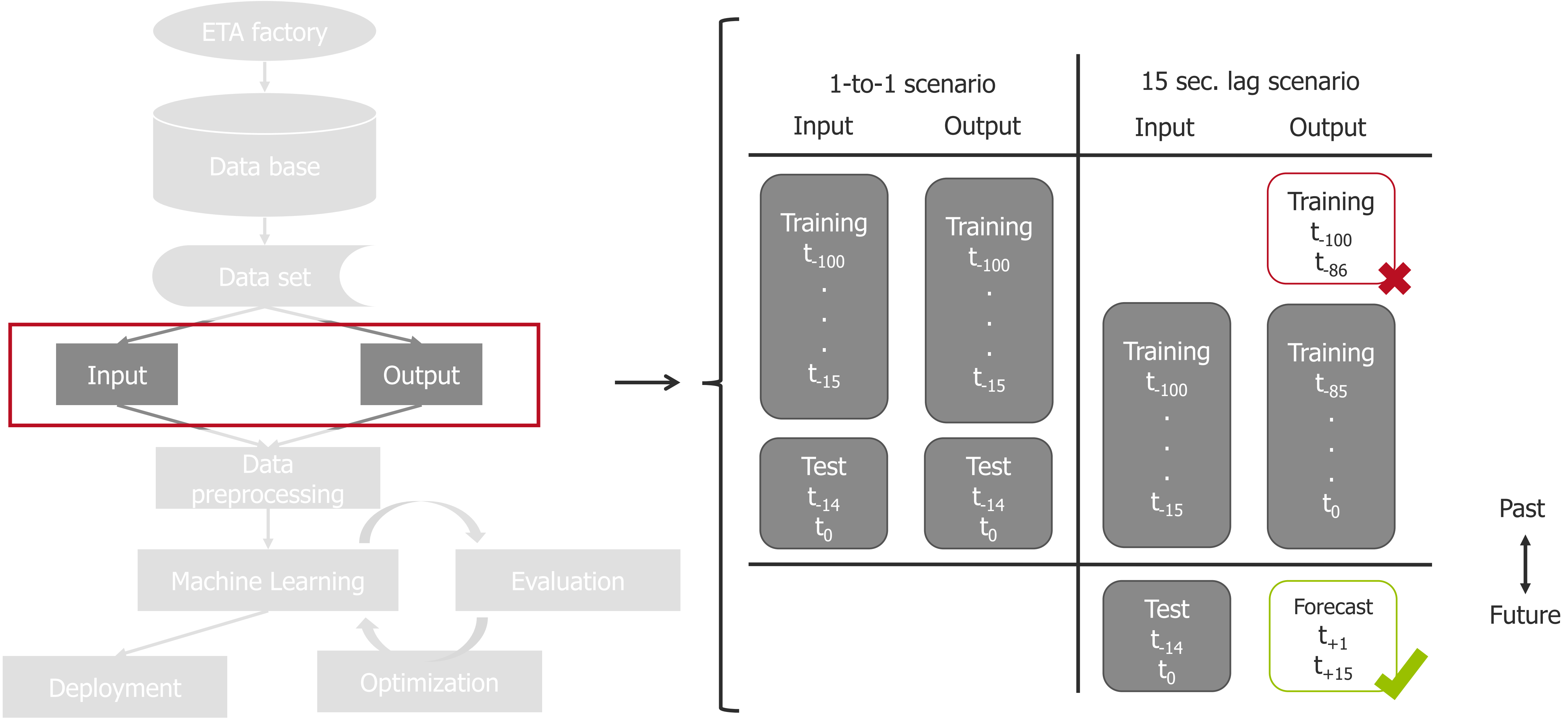
The CRISP-DM model – Data Preparation

Splitting into test, validation and training set

- Time series values are strongly dependent on values that are close in time
- Therefore, the test/validation set should always be a set of future values
- No shuffling to prevent mixing future and past values

The CRISP-DM model – Data Preparation

Time shift method for forecasting target preparation



The CRISP-DM model – Data preparation

Example of target time shift

Predict current value

	Time	Param 1	Param 2	Target (1-to-1)
now	t-0	$x_1^{(0)}$	$x_2^{(0)}$	$y^{(0)}$
	t-1	$x_1^{(1)}$	$x_2^{(1)}$	$y^{(1)}$
	t-2	$x_1^{(2)}$	$x_2^{(2)}$	$y^{(2)}$
past	t-3	$x_1^{(3)}$	$x_2^{(3)}$	$y^{(3)}$
	t-4	$x_1^{(4)}$	$x_2^{(4)}$	$y^{(4)}$
	…	…	…	…

Training example

Former target used as input feature

Forecast future value!

Time	Param 1	Param 2	Target (1-to-1)	Target (1 step forecast)
t-0	$x_1^{(0)}$	$x_2^{(0)}$	$y^{(0)}$	$y^{(+1)}$
t-1	$x_1^{(1)}$	$x_2^{(1)}$	$y^{(1)}$	$y^{(0)}$
t-2	$x_1^{(2)}$	$x_2^{(2)}$	$y^{(2)}$	$y^{(1)}$
t-3	$x_1^{(3)}$	$x_2^{(3)}$	$y^{(3)}$	$y^{(2)}$
t-4	$x_1^{(4)}$	$x_2^{(4)}$	$y^{(4)}$	$y^{(3)}$
…	…	…	…	$y^{(4)}$
…	…	…	…	…

The CRISP-DM model – Data Preparation

Feature engineering for the temporal feature set

Temporal features

	Time	Param 1	Engineered feature 1 (time lag 1 step)	Param 2	Engineered feature 2 (moving average)	Target
now	t-0	$x_1^{(0)}$	$x_{e1}^{(0)} = x_1^{(1)}$	$x_2^{(0)}$	$x_{e2}^{(0)} = \frac{1}{2}(x_2^{(0)} + x_2^{(1)})$	$y^{(0)}$
	t-1	$x_1^{(1)}$	$x_{e1}^{(1)} = x_1^{(2)}$	$x_2^{(1)}$	$x_{e2}^{(1)} = \frac{1}{2}(x_2^{(1)} + x_2^{(2)})$	$y^{(1)}$
	t-2	$x_1^{(2)}$	$x_{e1}^{(2)} = x_1^{(3)}$	$x_2^{(2)}$	$x_{e2}^{(2)} = \frac{1}{2}(x_2^{(2)} + x_2^{(3)})$	$y^{(2)}$
past	t-3	$x_1^{(3)}$	$x_{e1}^{(3)} = x_1^{(4)}$	$x_2^{(3)}$	$x_{e2}^{(3)} = \frac{1}{2}(x_2^{(3)} + x_2^{(4)})$	$y^{(3)}$
	t-4	$x_1^{(4)}$	$x_{e1}^{(4)} = x_1^{(5)}$	$x_2^{(4)}$	$x_{e2}^{(4)} = \frac{1}{2}(x_2^{(4)} + x_2^{(5)})$	$y^{(4)}$
	…	…	…	…	…	…
	t-n	$x_1^{(n)}$	X	$x_2^{(n)}$	X	

Training example

The CRISP-DM model – Data Preparation

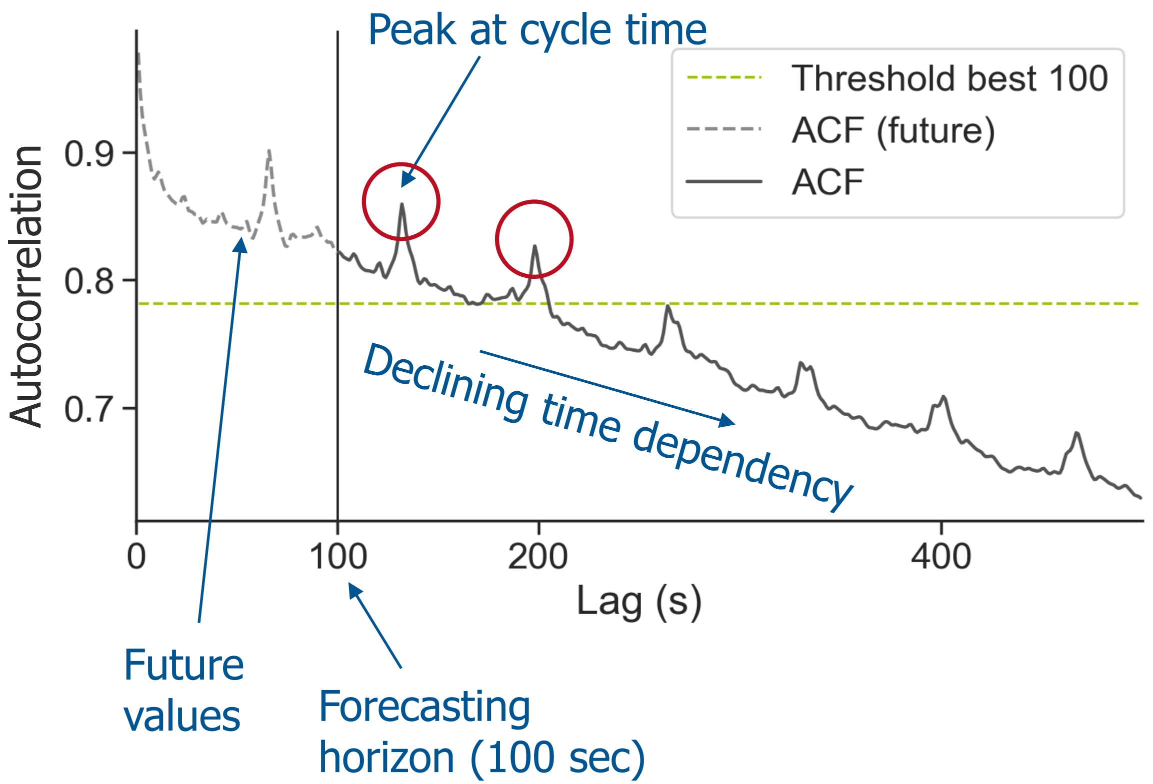
Feature engineering for the temporal feature set

Which time step(s) are significant?

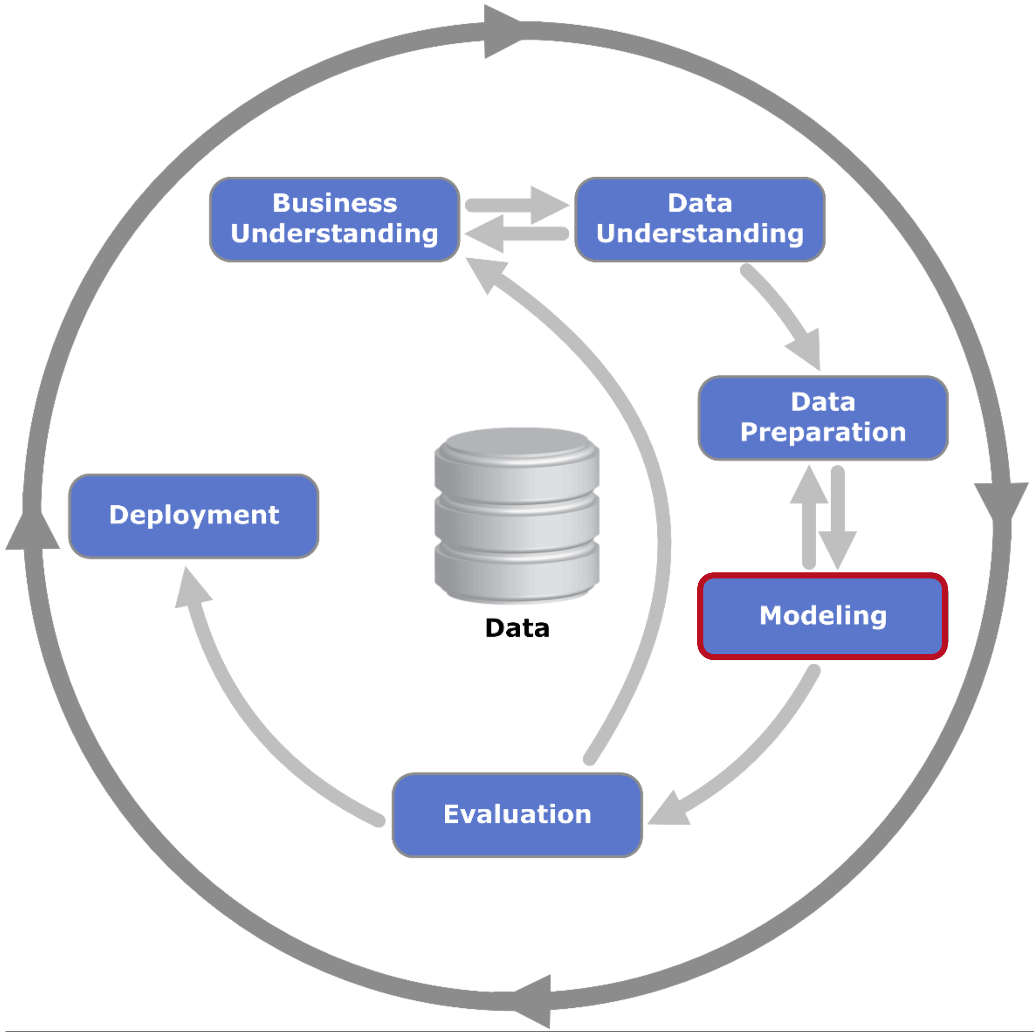
Time	**Param 1**	**Engineered feature 1 (time lag ?? steps)**	**...**
t-0	$x_1^{(0)}$	$x_{e1}^{(0)} = x_1^{(??)}$	
t-1	$x_1^{(1)}$	$x_{e1}^{(1)} = x_1^{(??)}$	
t-2	$x_1^{(2)}$	$x_{e1}^{(2)} = x_1^{(??)}$	
t-3	$x_1^{(3)}$	$x_{e1}^{(3)} = x_1^{(??)}$	
t-4	$x_1^{(4)}$	$x_{e1}^{(4)} = x_1^{(??)}$	
…	…	…	

The autocorrelation function (ACF)

Goal: Identify promising time lags for Feature Engineering



The CRISP-DM model – Modeling



The CRISP-DM model – Modeling

Suited algorithms for time series forecasting

Statistical models:

- Time series analysis (e.g. ARIMA),
- **Linear Regression**,
- …

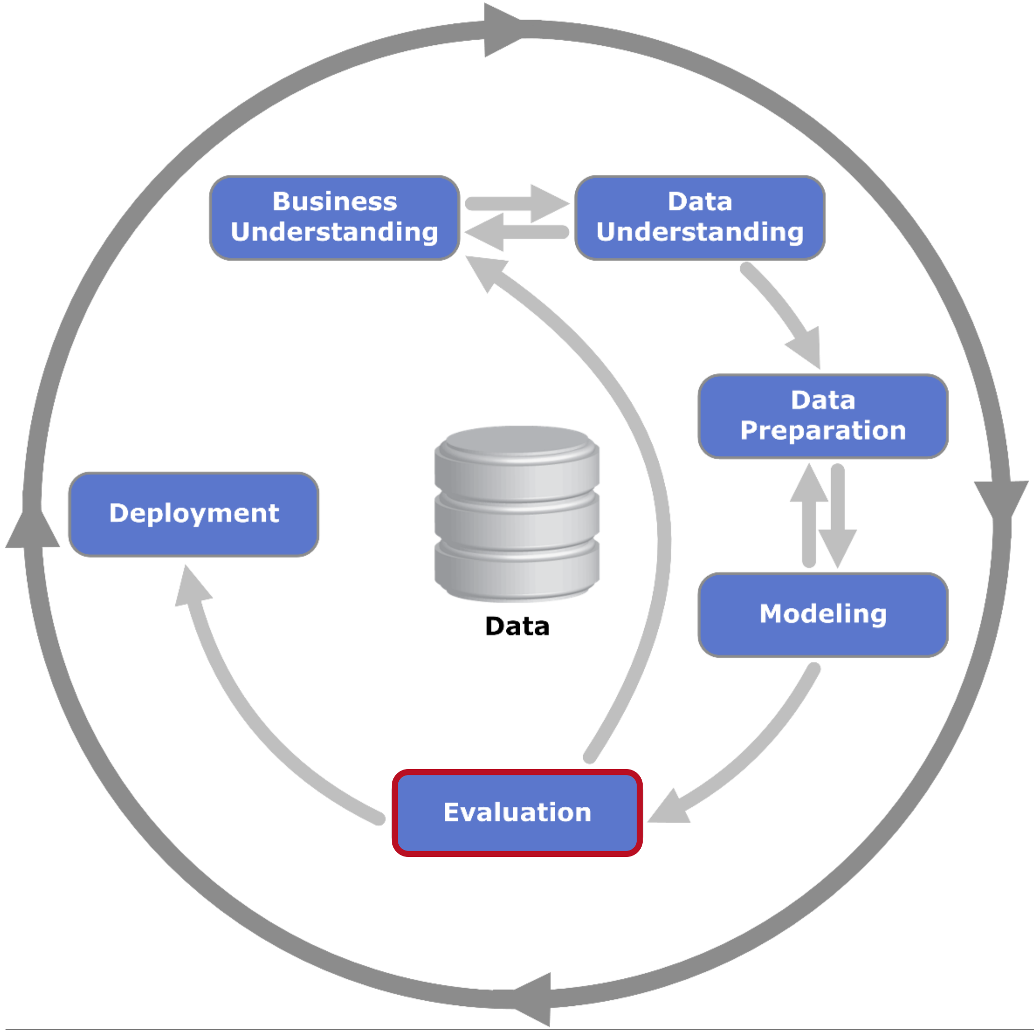
Machine Learning models for supervised learning:

- K Nearest Neighbors,
- Decision Tree,
- **Random Forest**,
- Support Vector Machine,
- **Artificial Neural Networks**,
- …

ARIMA = Auto Regressive Integrated Moving Average

→ Linear model of past values and moving average

The CRISP-DM model – Evaluation



The CRISP-DM model – Evaluation

Evaluate the forecasting accuracy

Goal: Is the achieved forecasting accurary sufficient for the business objective or is re-work required?

Guidelines to choose error metrics for the forecasting accuracy:

- Scale dependent: MAE, MdAE, RMSE
- Scale independant: nRMSE, R^2-score
- Always consider several different metrics in order to level out the weaknesses of the single metrics
- Decide how to measure the forecasting quality before training the model
- Decide what is an acceptable forecasting quality before training the model

The CRISP-DM model – Evaluation

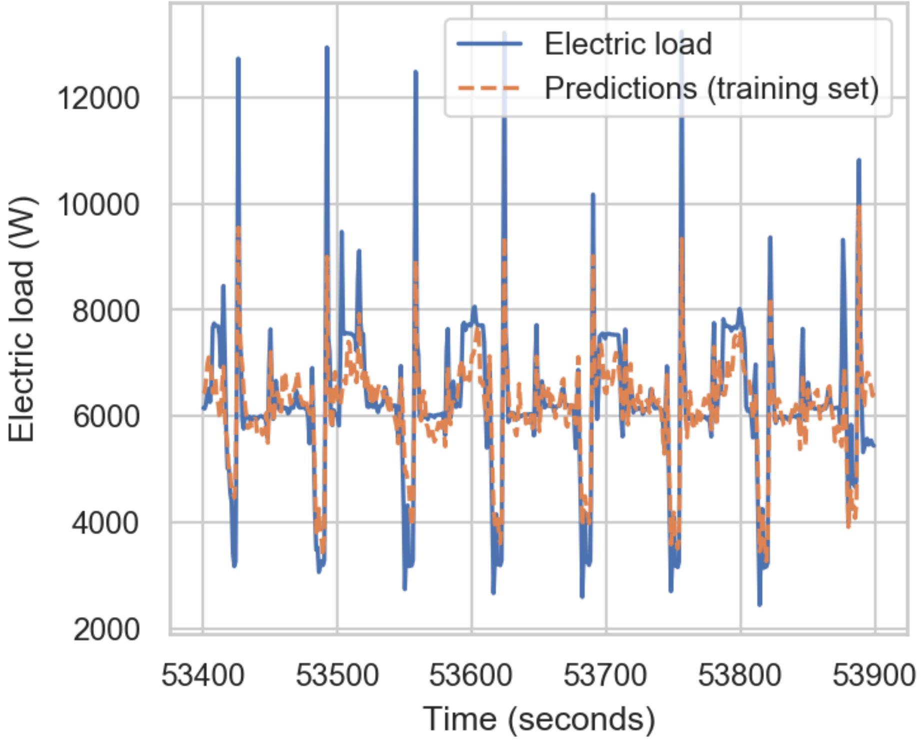
Error Metrics

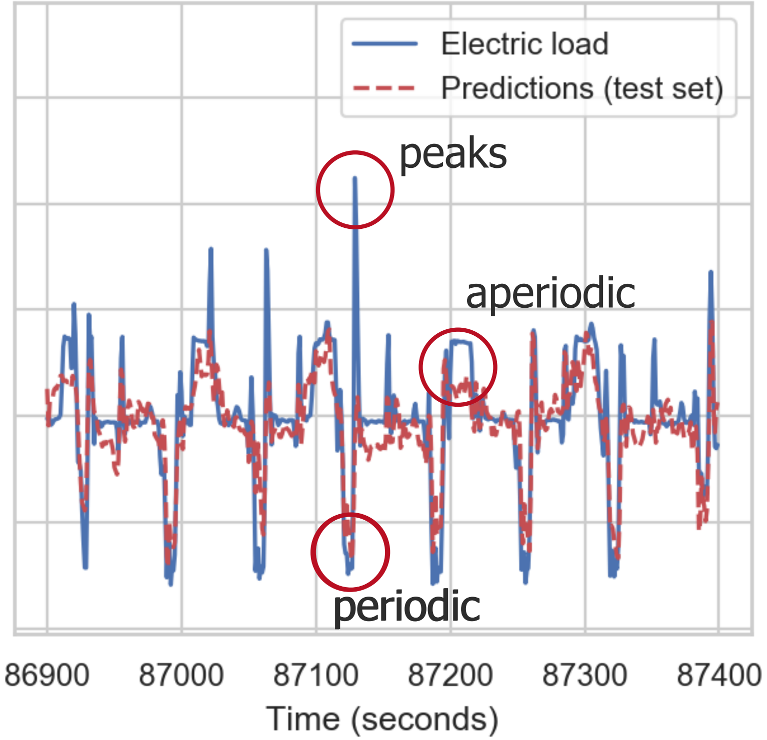
Group	Metric	Derivation	Advantage	Disadvantage
Scale-dependant	Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}$	+ Recommended for forecasting + High weight on large errors	- Sensitive to outliers
	Mean Absolute Error (MAE)	$MAE = \frac{1}{m}\sum_{i=1}^{m}\lvert y_i - \hat{y}_i\rvert$	+ Less sensitive to outliers than RMSE + Good to interpret	- Sensitive to outliers (less than RMSE)
	Median Absolute Error (MdAE)	$MdAE = \underset{i=1\ldots m}{median}(\lvert y_i - \hat{y}_i\rvert)$	+ Not very outlier-sensitive	- Harder to interpret than MAE and RMSE
Scale-independant	R²-score	$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \bar{y})^2}$	+ Standard metric in scikit-learn + Well-suited to estimate the generalization error + Normalized scale	- Sensitive to outliers
	Normalized RMSE (nRMSE)	$nRMSE = \frac{1}{n}RMSE$ with n = scaling factor	+ Normalized scale	- Sensitive to outliers - Scaling factor n has significant influence on the error metric

The CRISP-DM model – Evaluation

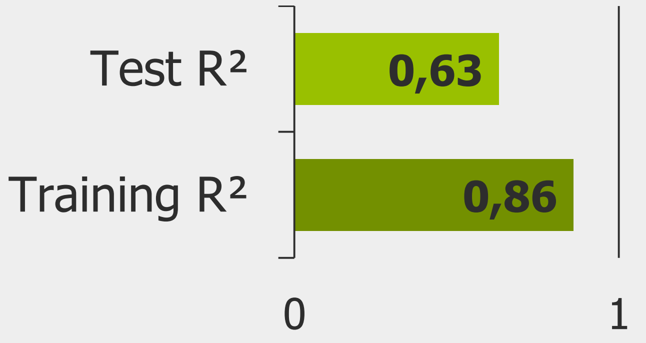
Results of the use case – Load forecasting of the machine tool EMAG VLC 100 GT

Target and predictions, zoom to seven production cycles:





Test-RMSE: 1207 W



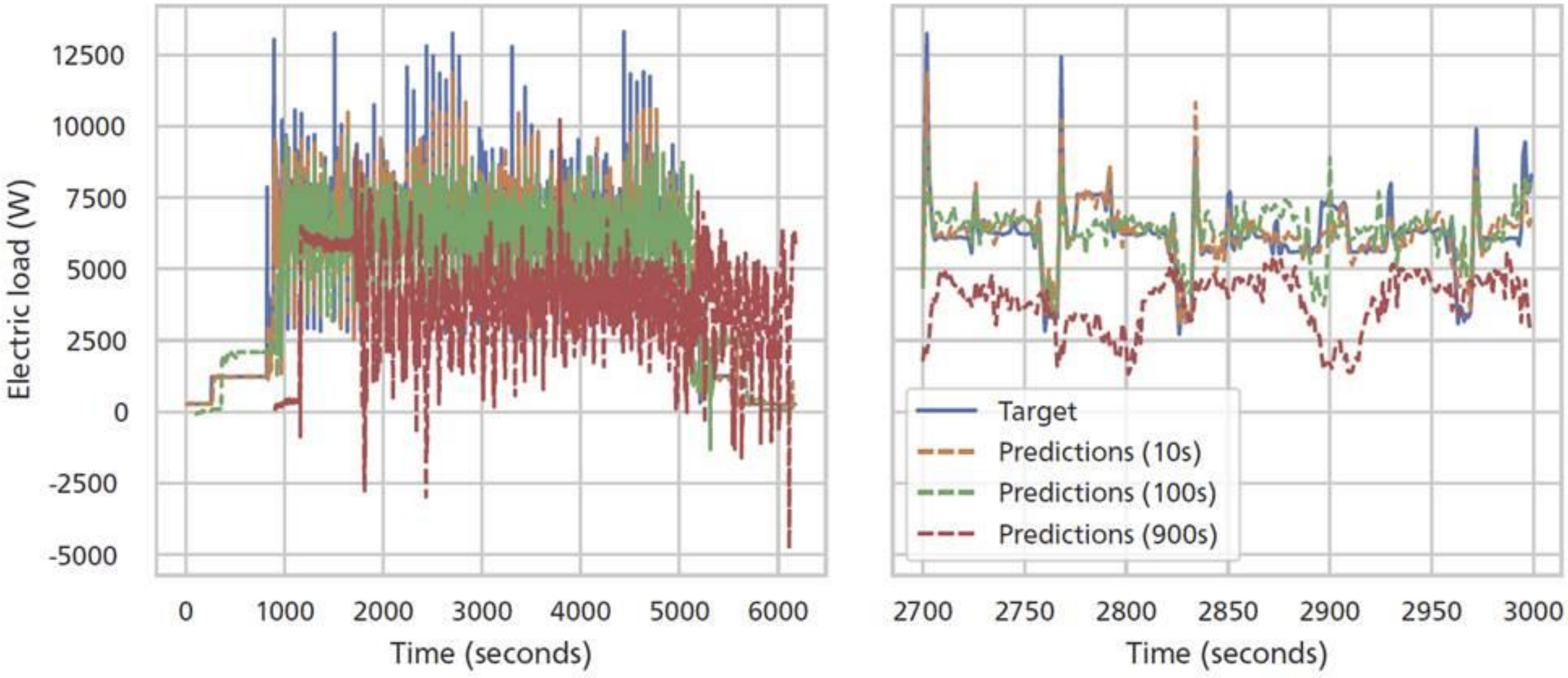
Result understanding:

- Periodic elements are captured well
- Aperiodic components (different cycle time) disturb the model → Improvement potential

Quelle: Walther, J.; Generic Machine Learning Approach for very short term Load Forecasting of production machines (2019)

The CRISP-DM model – Evaluation

Forecasting accurary of different forecasting horizons



The CRISP-DM model – Evaluation

Use case machine tool: Result of the hyperparameter optimization

Solution space, selected solutions marked in grey. Selection process was iterative with hyperparameter optimization in each step.

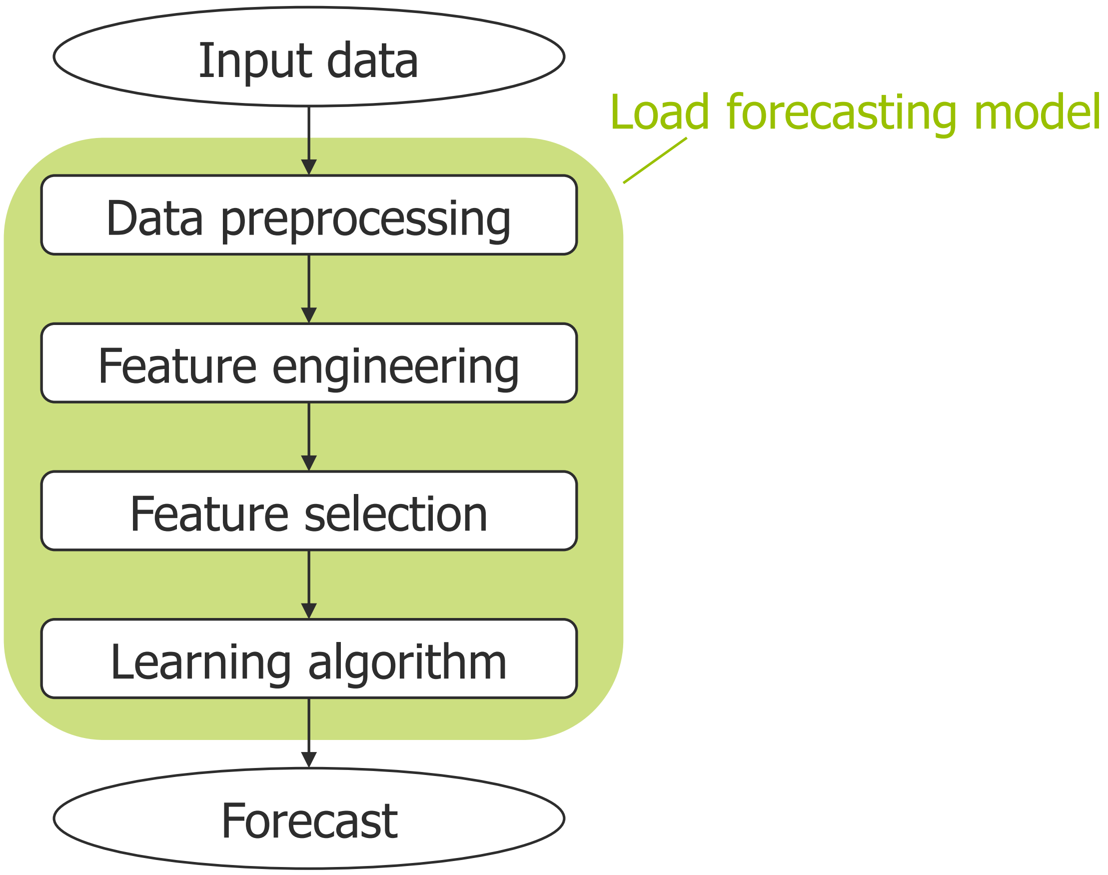
Problem	Solutions				
Forecasting horizon	10 seconds	**100 seconds**	900 seconds		
Learning algorithm	Linear Regression	KNN	Decision Tree	Random Forest	**ANN**
Imputation	Mean	**Median**			
Outlier treatment	Median and MAD	**None**			
Scaling	RobustScaler	**StandardScaler**	MinMaxScaler		
Feature engineering	**Univariate**	Multivariate	Combination		
Feature selection	By variance	By VIF	Recursive	**Combination**	

Conclusions

- Energy forecasting is an important tool for demand side management in industry
- Forecasting can be framed as a supervised learning problem and can be solved with known supervised learning algorithms
- Time series forecasting has some peculiar data preprocessing steps:
  - Split into training and test data
  - Target preparation: Time shift
  - Feature Engineering: Time lag and moving average
- Data understanding is very important for result interpretation and model evaluation

Agenda

Demo



Demo:

- Data understanding
- Target preparation
- Feature Engineering

Ausblick weitere Forschung
Energy Forecasting

Ausschreibungen:

- https://eta-fabrik.de/aktuell/studentische-arbeiten/
- RNN: https://eta-fabrik.de/aktuell/studentische-arbeit/entwicklung-von-lastprognosemodellen-fuer-werkzeugmaschinen-basierend-auf-recurrent-neural-networks/
- Zeitserien-Analyse: https://eta-fabrik.de/aktuell/studentische-arbeit/datenvorbereitung-und-bewertung-von-machine-learning-basierten-lastprognosemodellen/
- Wärmebedarfsprognose: https://eta-fabrik.de/aktuell/studentische-arbeit/entwicklung-einer-anlagenspezifischen-modellierungsstrategie-fuer-ki-basierte-waermebedarfsprognosen-in-der-eta-fabrik/
- U.v.m.

Lehrveranstaltungen der ETA-Fabrik:

- Master-Vorlesung und Tutorium
- https://eta-fabrik.de/bildung/lehrveranstaltungen/

Literature and Links

Basics of Machine Learning application:

Géron, Aurélien (2017): Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems: O'Reilly Media, Inc.

Further information on different Machine Learning problems and how-to's:

https://machinelearningmastery.com/

Ausschreibungen:

https://eta-fabrik.de/aktuell/studentische-arbeiten/

https://eta-fabrik.de/aktuell/stellenangebote/

ETA-Fabrik:

https://eta-fabrik.de/

https://www.youtube.com/watch?v=eY2kjUZB1oM (ETA-Video)

https://www.youtube.com/watch?v=vo8w4sOBv_4 (SynErgie Kurzfilm)

Thank you for your interest!

For further questions we are happy to be at your disposal.

TECHNISCHE UNIVERSITÄT DARMSTADT

PTW TU DARMSTADT

Bastian Dietrich, M.Sc.

b.dietrich@ptw.tu-darmstadt.de

+49 6151 / 16 22 0 98

Prof. Dr.-Ing. Eberhard Abele
Prof. Dr.-Ing. Joachim Metternich
Prof. Dr.-Ing. Matthias Weigold
Institute for Production Management, Technology and Machine Tools
Technische Universität Darmstadt

Otto-Berndt-Straße 2
64287 Darmstadt

Tel.: +49 61 51 | 16 2 00 80
Fax: +49 61 51 | 16 2 00 87
E-Mail: info@ptw.tu-darmstadt.de
Internet: www.ptw.tu-darmstadt.de