Taming Time Series!

Data Science Day 2019 @ THU

A very short **tutorial** on **time series analysis** and **forecasting** in the context of **business analytics**!

Dr. Simon Müller, Joachim Rosskopf
Ulm, 22.11.2019.

Agenda for Today.

Time series are **everywhere** and especially **relevant** for data science in the **enterprise context**.

Today we give a very short **hands on intro** into **algorithmical forecasting** with R. Further we provide an **outlook** what is necessary to do **real world business forecasting**.

- Welcome & Introduction
- Time Series are Everywhere!
- Hands-On Introduction to Algorithmic Forecasting.
- Real-world Business Forecasting
- Discussion & Wrap Up

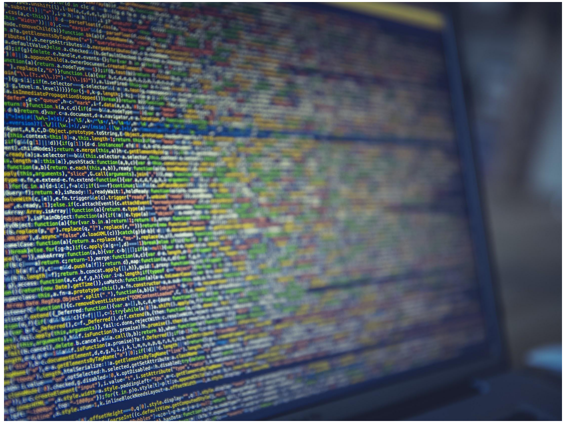
Dr. Simon Müller has a PhD in mathematics and has worked as a postdoc the field of biostatistics and is since six years a data science freelancer.

Joachim Rosskopf is lead of the DataTeam at Kärcher, where he tries to combine experience from software architecture and technology with business analytics and data science to build interesting, valuable data solutions.

Who is in the audience? Why are you here?

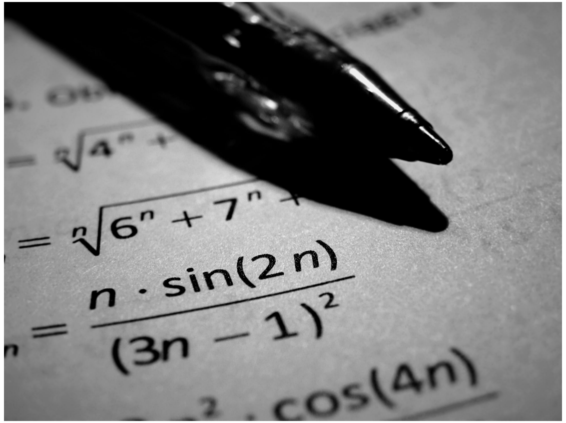
Data Analyst

__Analyze__ business, __discover__ data opportunities, __drive__ projects and __produce__ anchorage.



Data Engineer

Mainly concerned with __technology__ and __systems design__ and __operations__.

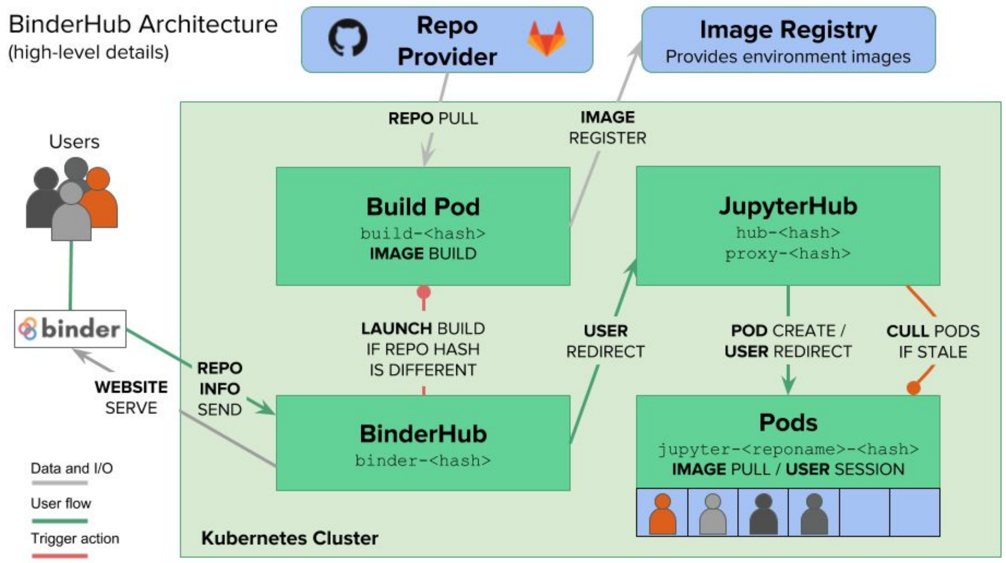


Data Scientist

Have deep __method understanding__, like __statistics__ or __machine learning__.

Interactive Part:

Start your session on Binderhub



Go To : https://github.com/anofox/data_science_tag_2019

Real-world Business Forecasting.

By **observing** and **forecasting** key **business processes** one can **optimize** to **gain efficiency** and **enable new business models.**

The challenge is to provide this at a **scale**, **automation** and **quality** which is relevant to enterprises.

Where does this matter.

Demand Forecasting

- Inventory Control
- Demand Planning
- Stock Optimization
- OTIF Optimization

Sales & Resource Planning

- Resource Allocation & Positioning
- Customer Behavior
- Price Adaption
- Marketing Allocation

IoT & Predictive Maintenance

Process & Condition Monitoring

Predictive analytics on enterprise processes is a greate data opportunity.

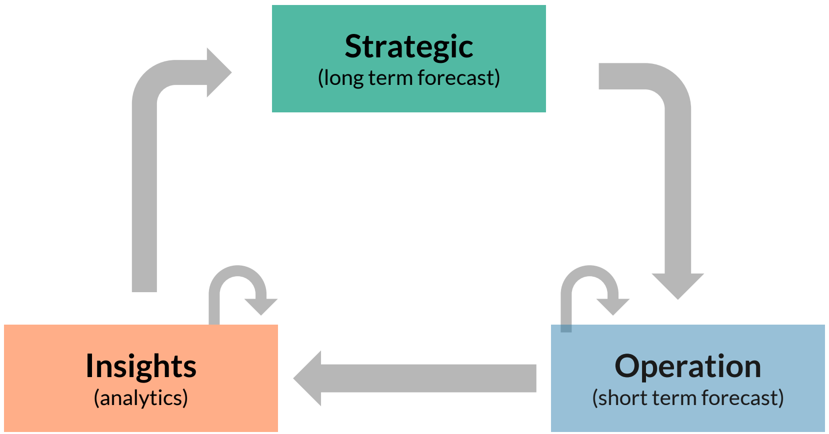
Business Forecasting

<u>Key questions:</u>

1. **Planning** and **demand forecasting** in industry is mainly a **data-driven task**! ✓

2. Since almost **15 years data** on **processes** is captured core processes! ✓

3. On a **tactical level** enterprises have **thousands of products** and **tens-** to **hundred-thousand customers**! ✓



Planning is a task which is **core** to a lot of central **processes** in an **enterprise**. Meaning it is **complicated**, hence has a lot of **potential** for ML.

Planning is a optimization problem. Forecasting is the beginning of the scenarios.

Business Forecasting

1

Enterprises manage **thousands** of **products** and **customers**

Forecast product demand

Forecast necessary spendings

2

Based on **demand forecasting**, **safety stock** must be set per material number.

Hierarchical forecasts

Optimization due to constraints

3

Effects of the optimization

Cost optimization in processes

Reduction of capital binding

Happier customers and new business

Reduction of planning effort and focus

For a real world solution several challenges have to be solved.

Integration

1 Observation of business key figures over time

Advanced Models

4 Modern AI tech. for practical forecasting

Automatic combination

5 Ensemble of models tailored to individual SKU or customer

Process Automation

8 Scalable cloud environment and APIs for automation & process integration

Data Preparation | Data Science / Business Analytics | Data Solution

2 Cleaning, adding public data, representation

Automatic feature extraction

3 Selection of domain specific metrics and KPIs

Business metrics

6 Bring forecasts together with constraints and optimize quantity, cycles or placement

Resource Optimization

7 Actionable results like forecasts, alerts and optimal values for core processes

Prescriptive Model

Planning is a optimization problem. Forecasting is the beginning of the scenarios.

Business Forecasting

1

Enterprises manage **thousands of products** and **customers**

- Forecast product demand
- Forecast necessary spendings

2

Based on **demand forecasting**, **safety stock** must be set per material number.

- Hierarchical forecasts
- Optimization due to constraints

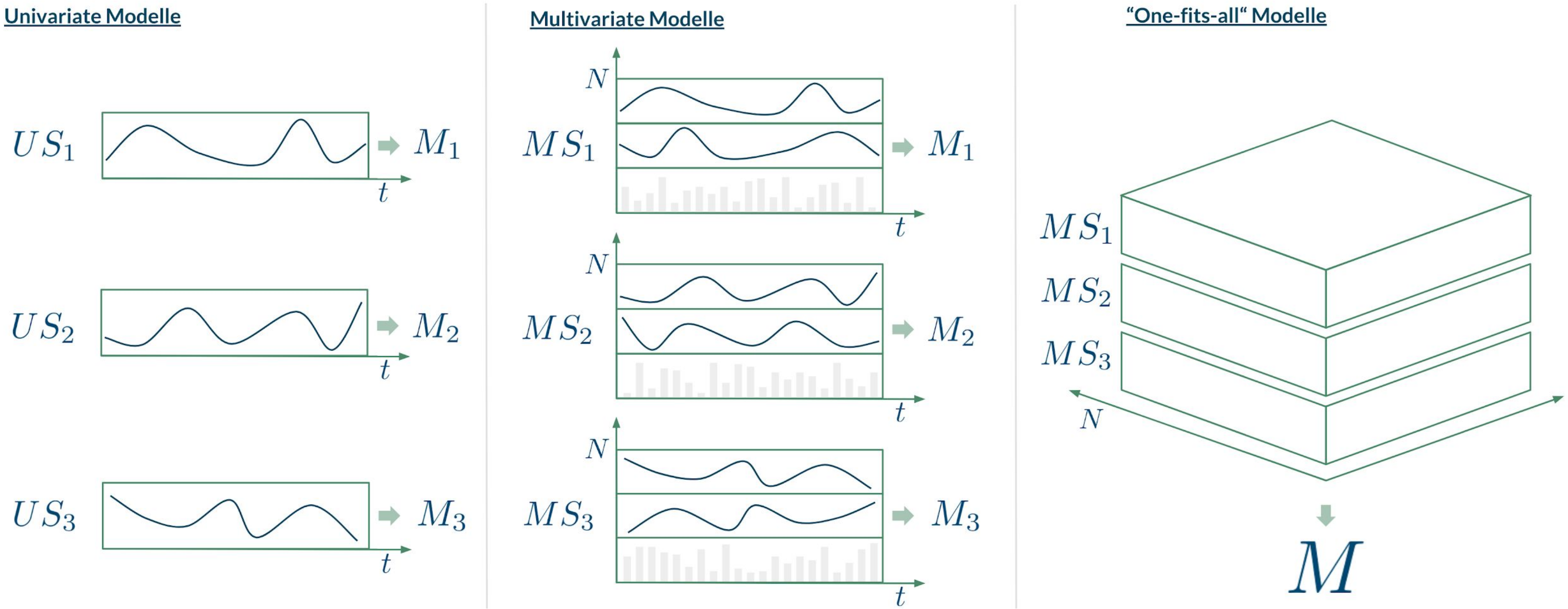
3

Effects of the optimization

- Cost optimization in processes
- Reduction of capital binding
- Happier customers and new business
- Reduction of planning effort and focus

Forecasting is the **starting point** for a lot of **complex planning** problems. The better its **results** are, the easier is good **downstream planning** quality. Especially for stochastic planning predictions are key.

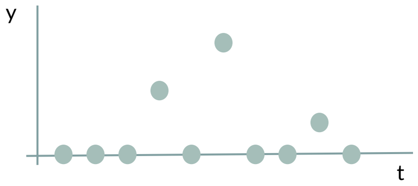
Advanced models improve quality for intermittent series with additional features.



Academia and the research institutions of digital giants like Amazon, Uber or SAP publish parts of their research results. Based on these results, one direction is to **build one model**, that **fits all data concurrently**, instead of one model per time series.

Example Algorithm:
Hierarchical Bayesian Methods (HBM)

Intermittent demands



1

2

Mathematical model

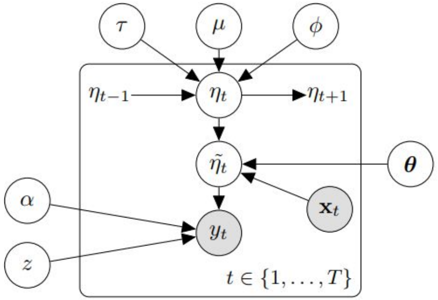
$$y_t \sim \mathrm{NB}(\exp \eta_t, \alpha).$$

Discrete distribution to model "0"

$$\begin{aligned}
\eta_1 &= \mu + \epsilon_1, \\
\eta_t &= \mu + \phi(\eta_{t-1} - \mu) + \epsilon_t, \quad & t > 1, \\
\epsilon_1 &\sim \mathcal{N}(0, 1/\tau_0 + 1/\tau), \\
\epsilon_t &\sim \mathcal{N}(0, 1/\tau), & t > 1,
\end{aligned}$$

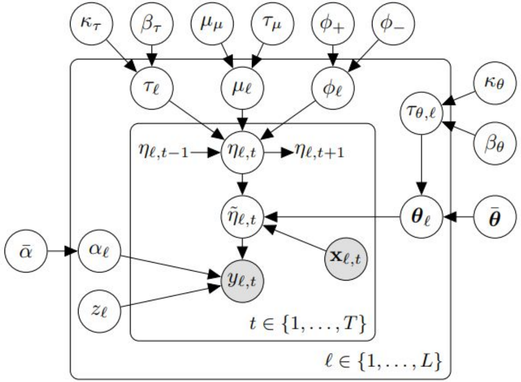
Autoregressive model for demands ">0"

Single spare part



3

Multi spare part (one fits all)



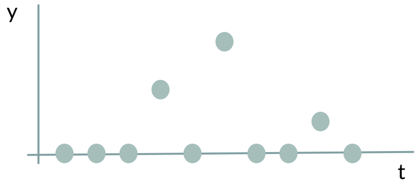
"**One fits all**" approach done via **sharing latent variables** between part time series.

The algorithm is a **mixture of Bayes nets** and **autoregressive models** and shows great promise, and practical use in **retail- & warehouse forecasting**.

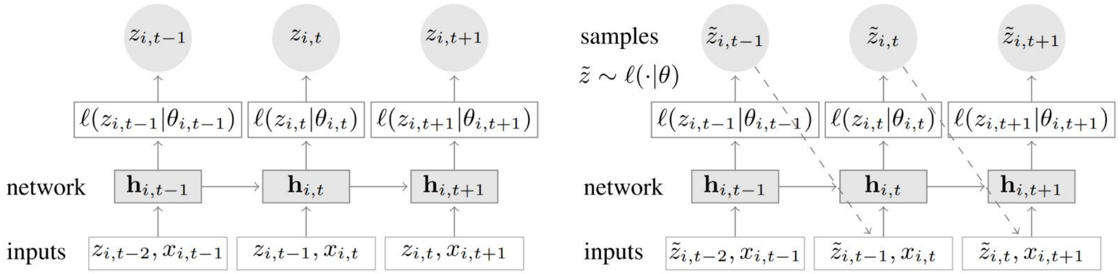
Business Forecasting

Example Algorithm: Autoregressive Recurrent Networks

Intermittent demands



Recurrent Neural Network



"**One fits all**" approach optimizes the loss function that depends on all N series and is considering the intermittent data structure

$$\mathcal{L} = \sum_{i=1}^{N} \sum_{t=t_0}^{T} \log \ell(z_{i,t}|\theta(\mathbf{h}_{i,t}))$$

The algorithm uses **recurrent neural networks** with a special **quantile loss** function to predict future demands. It is developed by Amazon Labs in Berlin.

Automatic combination reduces effort while improving robustness of forecasts.

Business Forecasting

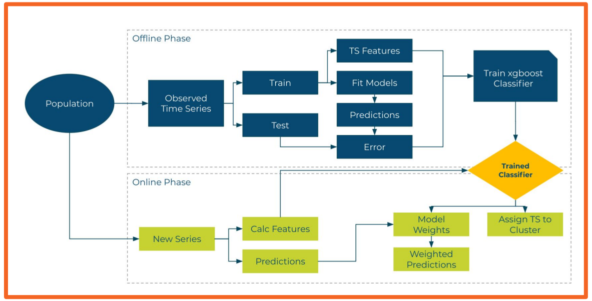
Learns to build an ensemble of algorithms automatically:

Input

Historical Demands

Forecast Models

Meta-Learning with Ensembling

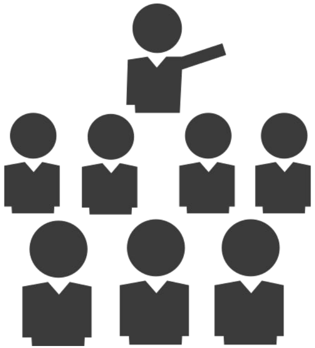


Output

Ensemble Forecasts

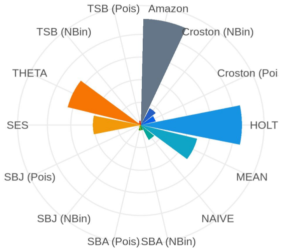
Demand classif.

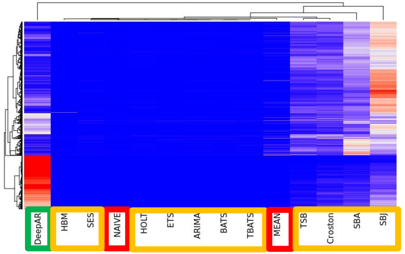
An ensemble is the optimal combo of individual models:



- Each algorithm has strength
  - Dependant on situation
  - No free lunch
- Combination of learners is beneficial
  - Better precision
  - More robust
  - Automatic training
- Can be used to classify / group series

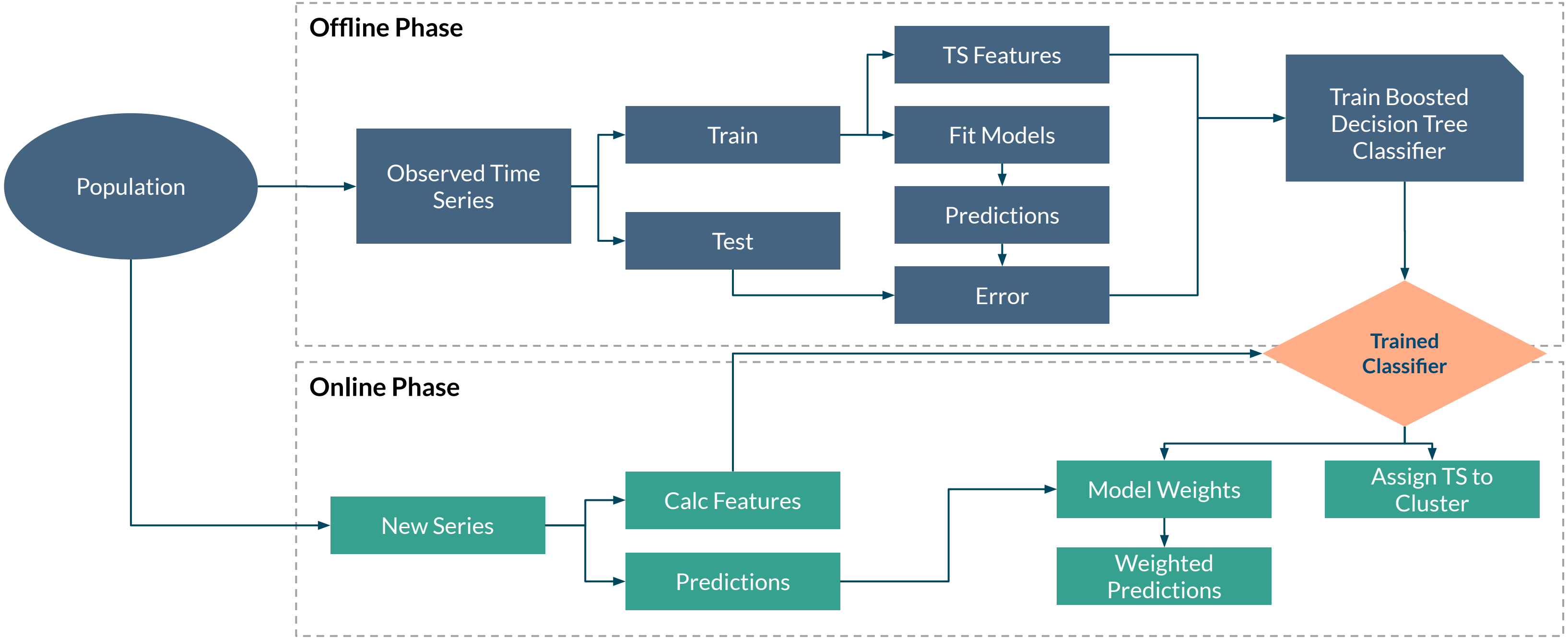
$$g_{ens}(\cdot) = \sum_{k=1}^{N} c_k g_k(\cdot)$$





Automatic combination reduces effort while improving robustness of forecasts.

Business Forecasting



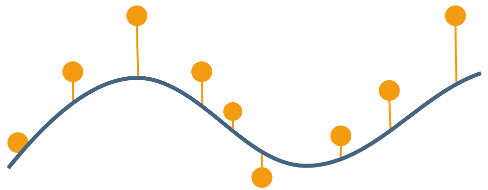
Again **machine learning** is used to learn in which situation which **combination** of **models** should be used. This can also be used to **classify customers** or **parts**.

A core challenge to successfully solve the business problem is to choose the right error metric.

Business Forecasting

1

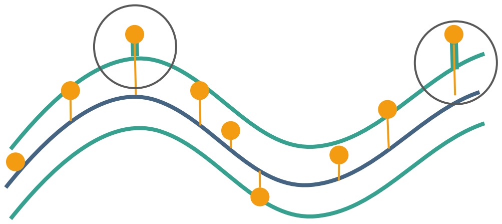
Try to predict the expected value of the time series as good as possible.



Suitable for high-freq. IoT predictions. Limited value for logistics

2

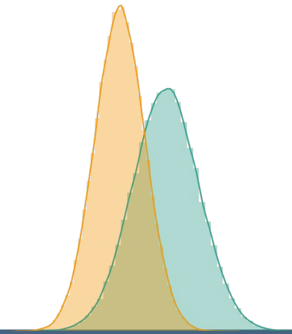
On top of expected value Penalize predictions not fitting inside the the error bounds.



Already contains the central idea, that logistics is more concerned on variability and intervals.

3

Optimize to make the ML algorithm predict the right distribution of demands.



For a low amount of data, this measure really captures the key questions: What are the demands to be prepared for?

Traditional **objective function** for time series like MAPE or MASE **don't** fully **capture** the **problem** of demand forecasting. Alternative measures concentrate on optimizing the **distribution** of **demands**. This is especially useful for **stochastic demand planning.**

Meet the AnoFox!

Key features:

1. Scalable, unsupervised **time series forecasting** and **anomaly prediction** on business and IoT data (machine learning and statistical methods)**.**

2. **Quickly deployable** as **building block** into the **virtual private clouds** of customers. We come, where the data already resides!

3. We rely on **cloud services**, **open source** software and **modern data science** methods. At its core we rely on **battle tested data analysis**. For higher level intelligence we utilize **state of the art machine learning** research.

Use cases in the wild:

Optimize inventory and **forecast part or service demands**. Focus on tricky intermittent demands (e.g. spare parts, customer purchases)

Predict customer demand and **steer sales and marketing**. Focus on recurrent buying behavior for consumables and spares.

Do **condition monitoring** and **anomaly detection** on **internet of things (IoT)** devices or **group customers** by device utilization**.**

> Focus of **AnoFox** are **data driven core processes** of **manufacturing companies** with **high business value** and a **clear return on investment.**

AnoFox: enterprise grade forecasting out of the box

1 *Integration*

- **SAP BW, ERP, HANA** integration (bi-dir) via **SparrowBI**
- **AWS Abstraction**
- **Azure Abstraction**
- **GCP Abstraction**

2 *Pipeline*

- **Pipeline Orchestration**
- **R & Python Blocks**
- **Metadata**
- **Parallelization & Distribution**

3 *ML & Tools*

- **Forecasting Workbench**
- **Forecasting Framework**
- **Deep Learning Forecasters**
- **Op. Res. Forecasters**

4 *Blueprints*

- **Demand Planning**
- **Anomaly Detection**
- **Predictive Maintenance**
- **Process Monitoring**

5 *Operationalization*

- Feeding back into **SAP BW, ERP, HANA via SparrowBI**
- **Reporting & Dashboarding**
- **Alerting & Messaging**
- **Forecasting via API**

Data Preparation | Data Science / Business Analytics | Data Solution

Thank you for the opportunity to present our ideas!

Contact us:

Telephone: +49 177 30 73 73 7
Email: contact@data-zoo.de