technische universität dortmund

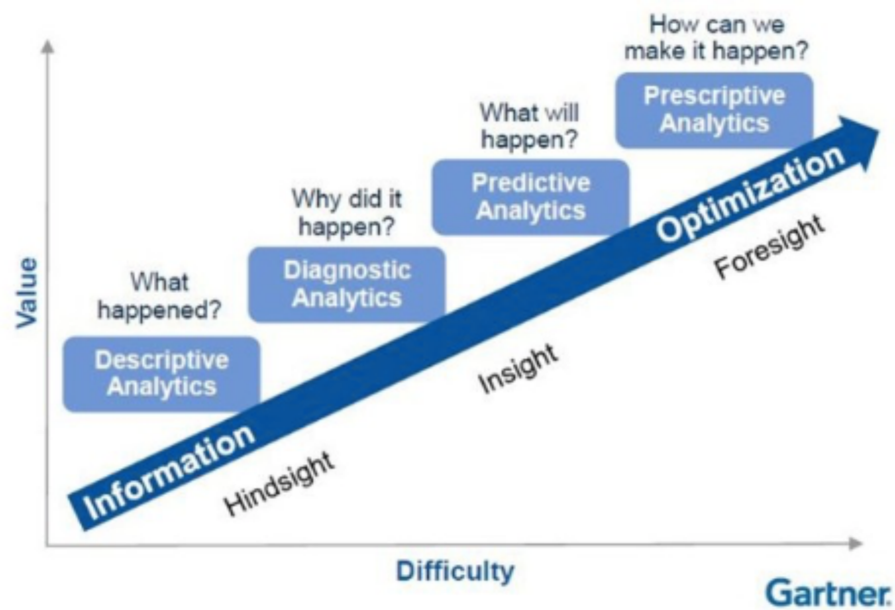
SFB 876 - Providing Information by Resource-Constrained Data Analysis

Machine Learning and Data Science – Research and Applications in Industry 4.0

Prof. Dr. Katharina Morik,
Künstliche Intelligenz,
TU Dortmund

Overview

- Introduction:
  Collaborative research center SFB 876
  - Big data and small devices
  - Streaming data
    Astrophysics
- Anomaly detection for diagnostic analytics
- Quality prediction as predictive analytics
- Quality control by prescriptive analytics



Collaborative Research Center 876:
Providing Information by
Resource-Constrained
Data Analysis

13 projects
20 professors
50 Ph D students

Integrated graduate school

2011 - 2018
4 more years are possible

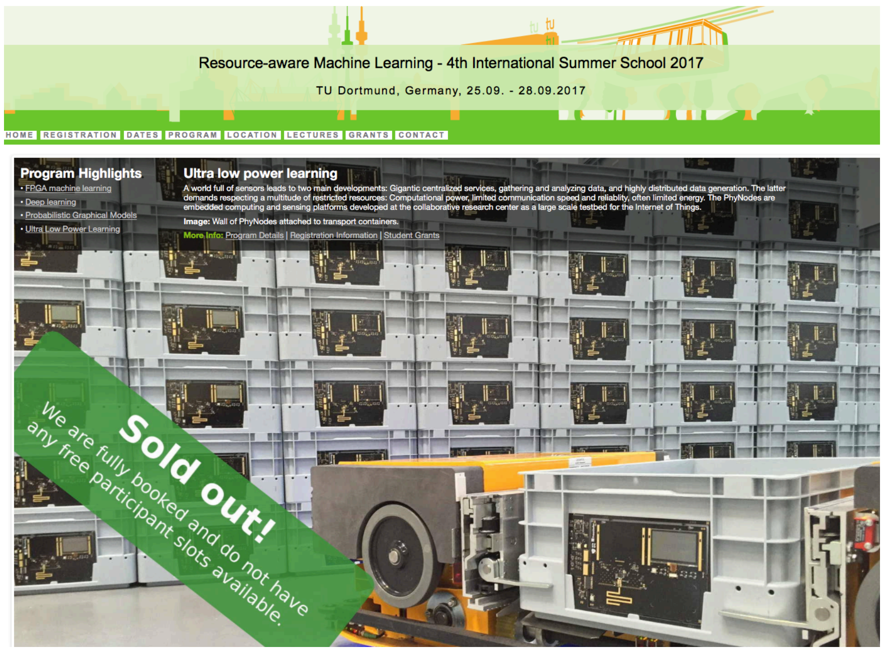
Internet of Things in Logistics

- Smart containers
  - Communication
  - Energy harvesting
- Small devices: logistics chips produced by SFB 876 produce big data.
- Analytics turn big data into smart data, here: enabling better routing.

Michael ten Hompel

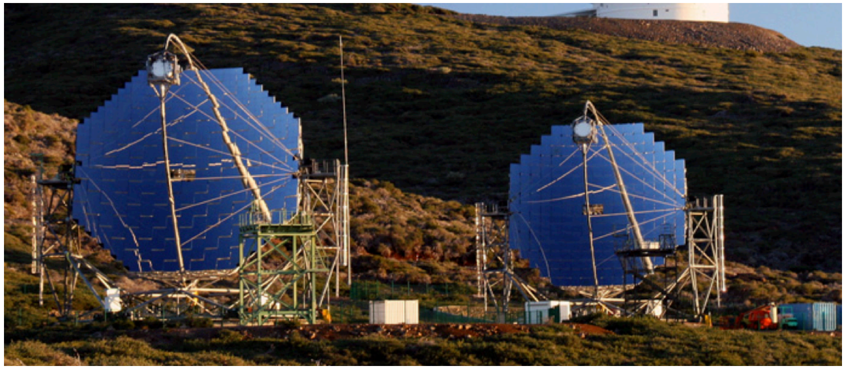
Project A4 in SFB 876

Test field of logistics
collaborative research center SFB 876

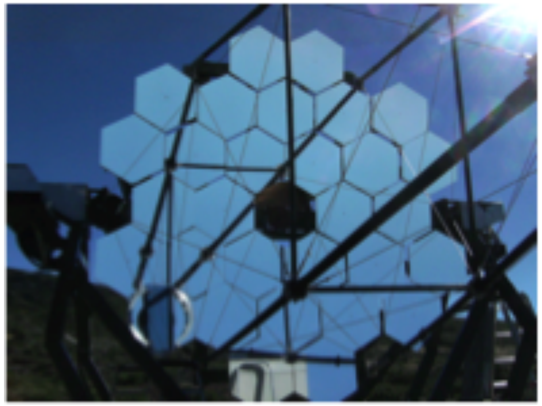


Massive data streams in astrophysics

- Imaging atmospheric Cherenkov telescopes (IACT) have mirrors and a camera to record the Cherenkov blue light produced by particle showers.
- A library of C++-programs, ROOT, and MARS programs store and preprocess the pictures.
- A simulator provides labeled observations.
- Gamma rays of high energy are rare events as opposed to hadrons, ratio 1 to 1000.

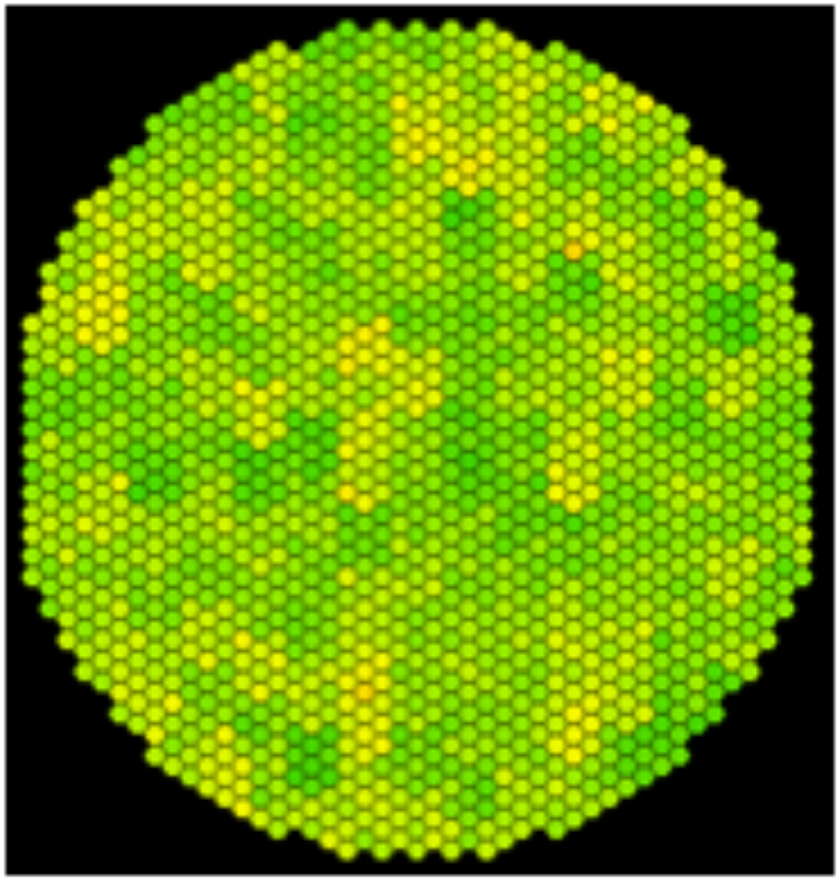


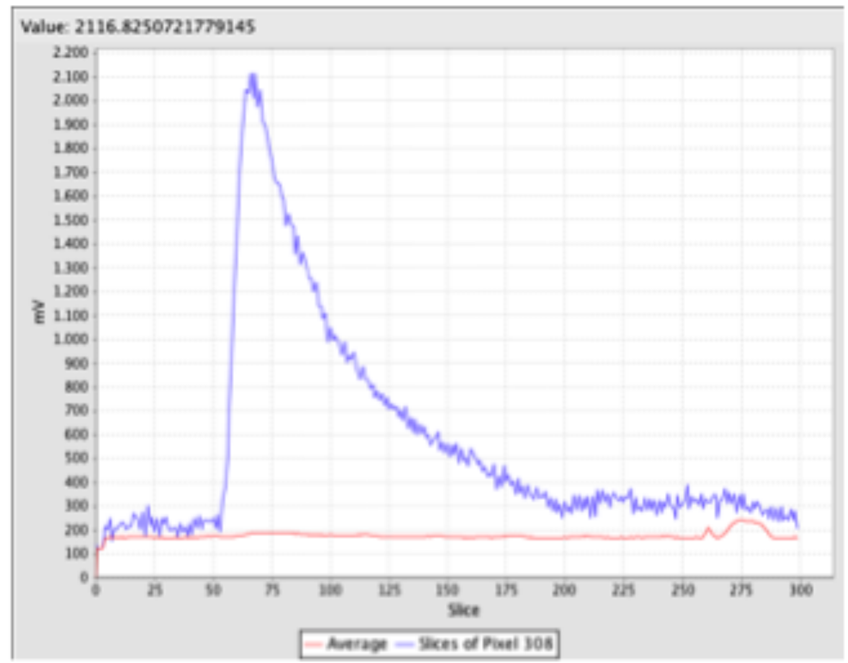
MAGIC I (2003) and MAGIC II (2009)
La Palma, Roque de los Muchachos



FACT (2011) same place

FACT-Viewer




Value: 2116.8250721779145
mV
Slice
Average
Slices of Pixel 308

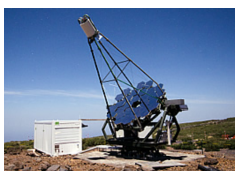
Integrating analytics in streaming environments: *streams*

Abstraction of various streaming data as data flow graphs by *streams* framework, which accesses Storm, Spark, RapidMiner,...

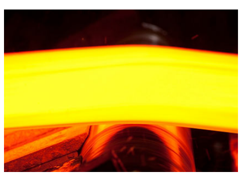
Christian Bockermann "Mining Big Data Streams for Multiple Concepts" 2015, TU Dortmund University, https://eldorado.tu-dortmund.de/handle/2003/34363

Applications

Realtime IP TV statistics



FACT Telesope sensor data



Steel Production sensor data

City Traffic sensor data (bus)

Preprocessing

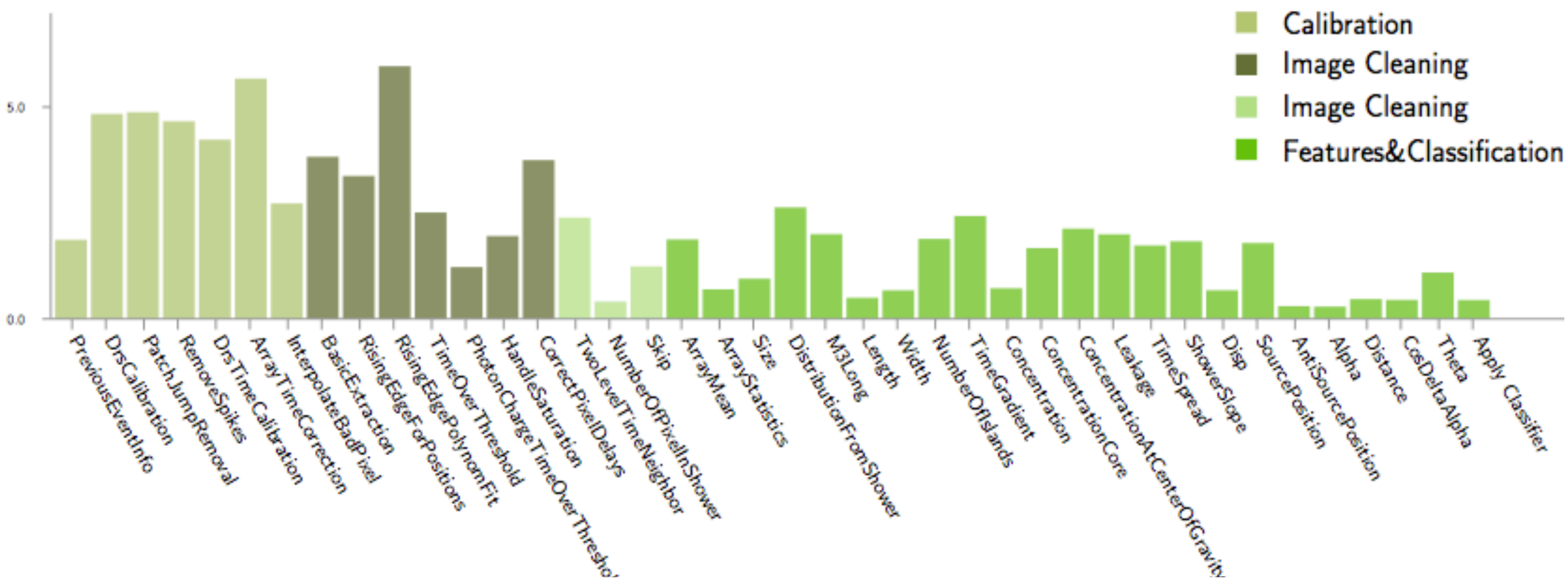
- Feature Extraction
- Transformation
- Filtering

Machine Learning

- **RapidMiner**
- Weka
- MOA

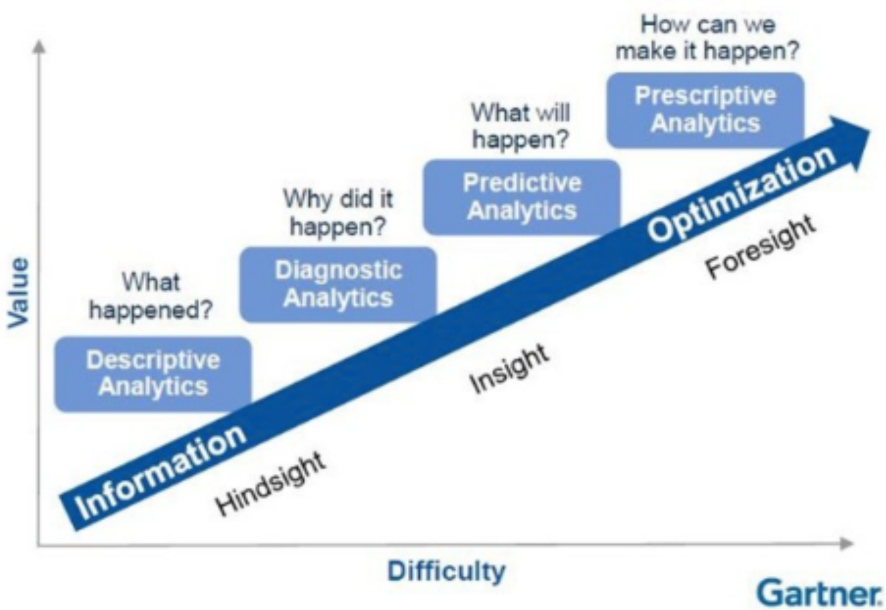
Preprocessing streaming data

- FACT records 60 events per second.
- Each events amounts to 3 Megabyte of raw data.
- ➢ 180MB/second are to be processed!
- ➢ Average processing time in milliseconds at a log scale shows the overall process ending with a classifier application.

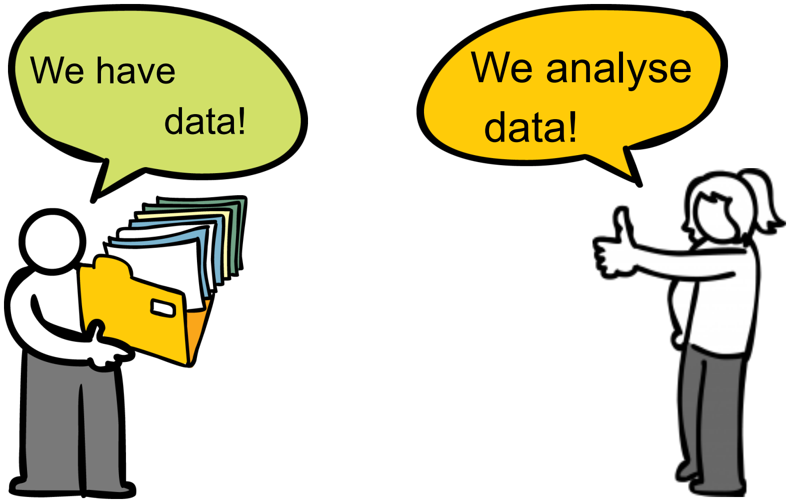


Overview

- Introduction:
  Collaborative research center SFB 876
  - Big data and small devices
  - Streaming data
    Astrophysics
- Anomaly detection for diagnostic analytics
- Quality prediction as predictive analytics
- Quality control by prescriptive analytics



The “we have data problem”



- RapidMiner
  - eases preprocessing,
  - supports interdisciplinary work,
  - demands expertise, experience.
  - Easy to change!
  - Easy to maintain!

- Preparing the data for the analysis is a hard problem:
  - time-consuming
  - requires knowledge of machine learning, statistics
  - requires domain knowledge.

Change parameters!

Press play!

Anomaly detection

- Feature extraction
- Feature selection
- Single class SVM,
  Core Vector Machine
- Clustering of observations
- Using many clustering for determining the certainty of an anomaly
- Reporting anomalies to the user

Injection Molding – Supervised feature selection

- Minimum Redundancy Maximum Relevance feature selection requires data with labels: <**x**, y>
- Most observations are not labeled.
- Using domain knowledge by asking the expert? Each **x**?!
- Using domain knowledge indirectly!
  - Known causalities label observations f(**x**) = y
    e.g., y=max injection pressure
  - Features are ranked according to their contribution to correct predictions.

- Dataset 1:
  5.2 Mio. observations from 1154 processes
  varying material wetness
- Dataset 2:
  4.3 Mio. Observations from 721 processes
  varying injector size.
- Structured according to component groups:

  Schnecke,

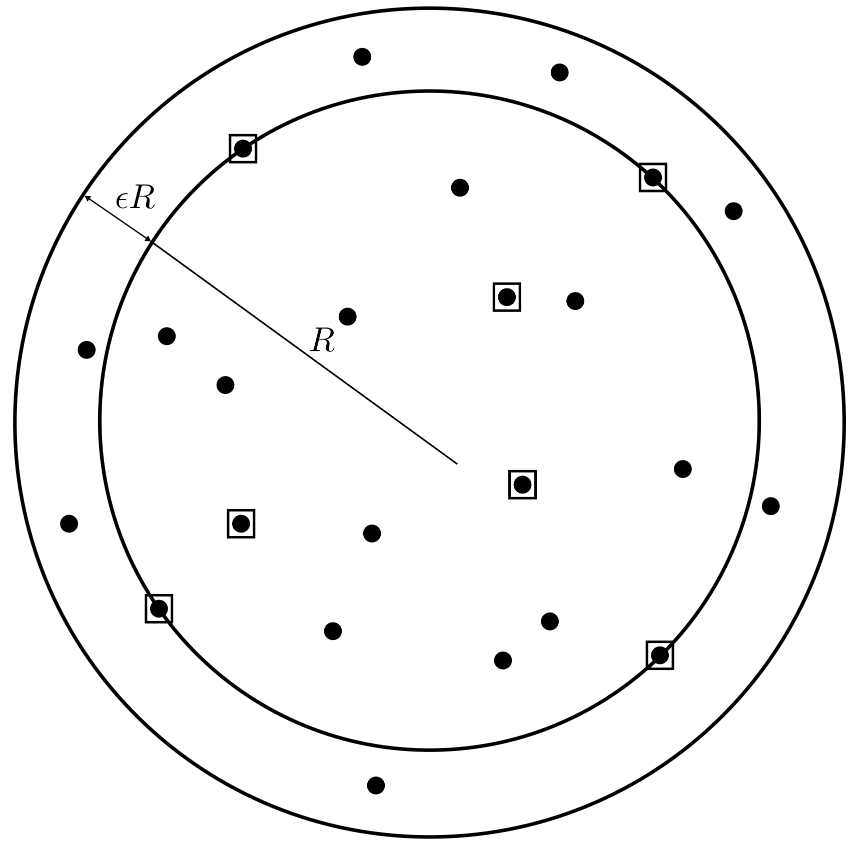
  Werkzeug,

  Heizung

Johannes Wortberg, Alexander Schulze-Struchtrup, Chen-Liang Zhao (2017): Digitalisierung der Spritzgießproduktion – Intelligente Maschinen für effiziente Prozesse nutzen. In. Spritzgießen, VDI Jahrestagung, VDI-Verlag, 55-65

Injection Molding -- Unsupervised feature selection

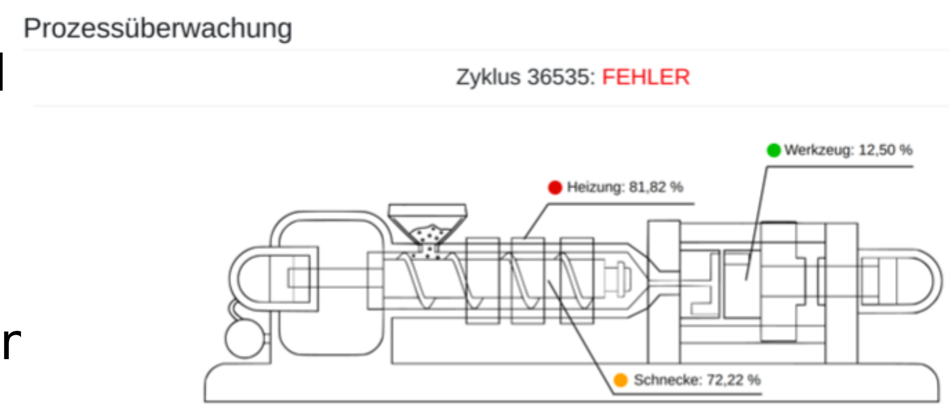
- Single class SVM
  - Outliers are anomalies
  - SVM ranks features according to their contribution to the decision
- Multi-objective optimization clustering
  - Members in a cluster are close to each other
  - Few clusters
  - Few not assigned observations
  - Members of different clusters are very different



Single class SVM, minimum enclosing ball

Weighting of features, weighting of anomalies

- Evolutionary process delivers several feature sets, each is used for clustering.
- For all clusterings:
  Large clusters are considered normal.
  Small clusters show anomalies.
- Features that are often used ir large clusters
  receive a higher weight.
- Anomalies that are found by many clusters
  receive a higher weight.

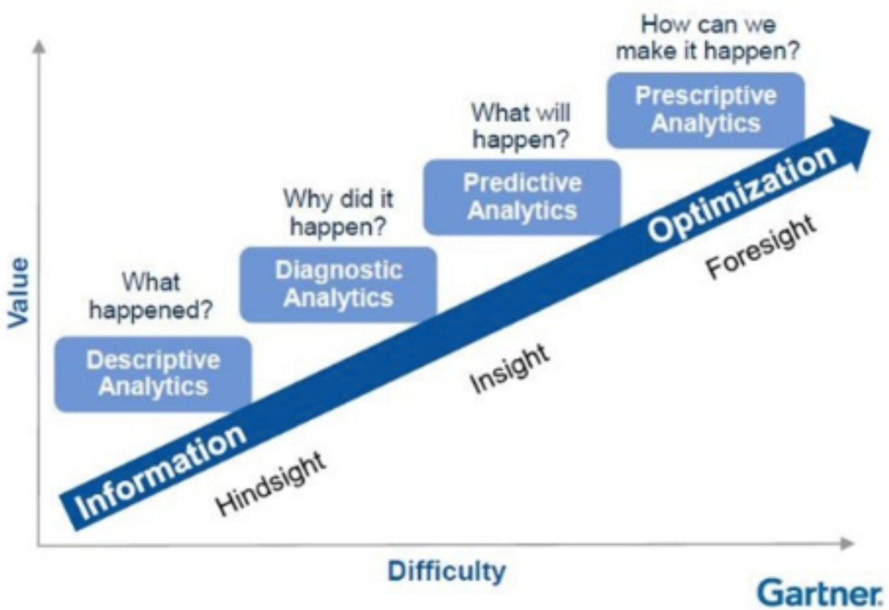


Anomaly detection

- Feature extraction
- ✓ Feature selection
- ✓ Single class SVM,
  Core Vector Machine
- ✓ Clustering of observations
- ✓ Using many clustering for determining the certainty of an anomaly
- ✓ Reporting anomalies to the user
- ➢ Experiments show, that a pre-selection based on domain knowledge may enhance or decrease feature selection.

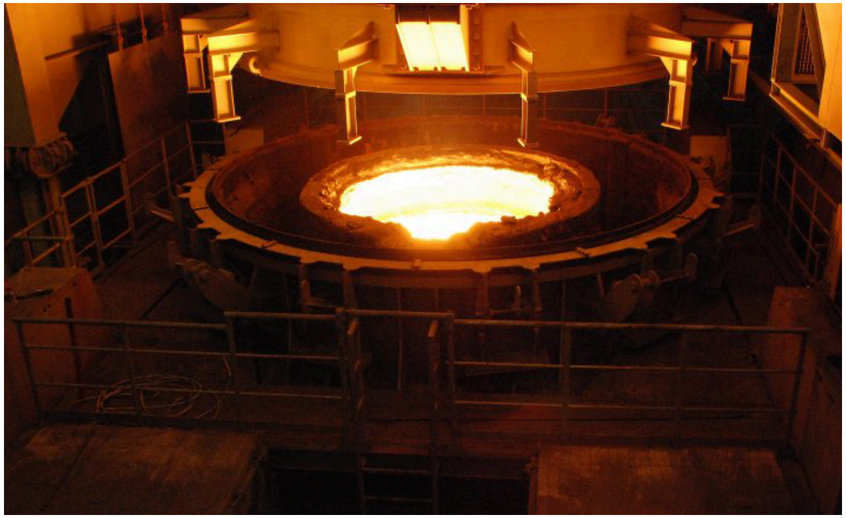
Overview

- Introduction:
  Collaborative research center
  SFB 876
  - Big data and small devices
  - Streaming data
    Astrophysics
- Anomaly detection for diagnostic analytics
- Quality prediction as predictive analytics
- Quality control by prescriptive analytics



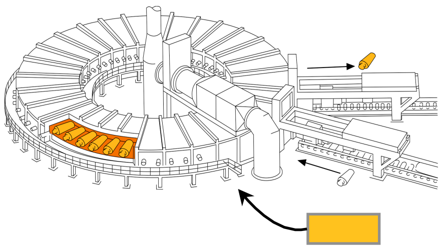
Quality prediction as predictive analytics

- Making the data smart: RapidMiner for
  - Preprocessing of time series data
  - Aggregation, feature extraction
  - Prediction
- Project B3 in SFB 876 with Jochen Deuse
- Collaboration with Deutsche Edelstahlwerke on quality prediction in a rolling mill.



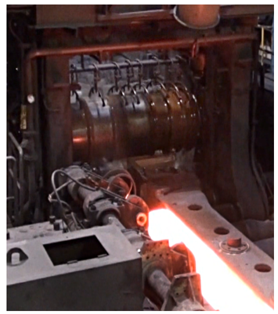
Smart data for smart factories

Rotary Hearth Furnace

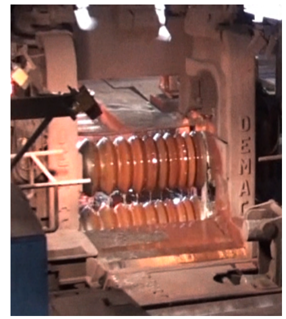


Steel bars

Block roll

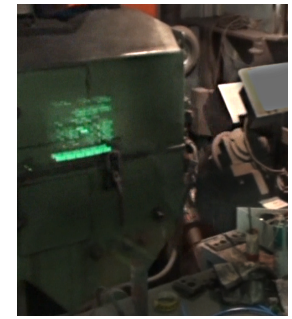


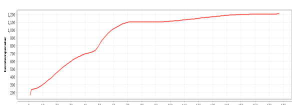
Finishing roll 1/2

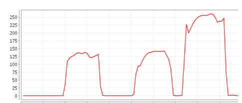


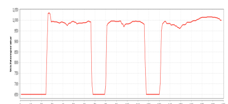
Cutting

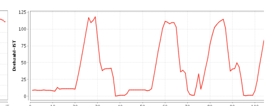
Ultrasonic tests











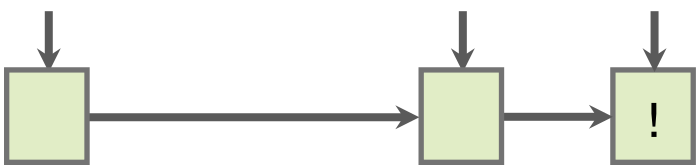
Temperature

Force

Temperature

Speed

Test results



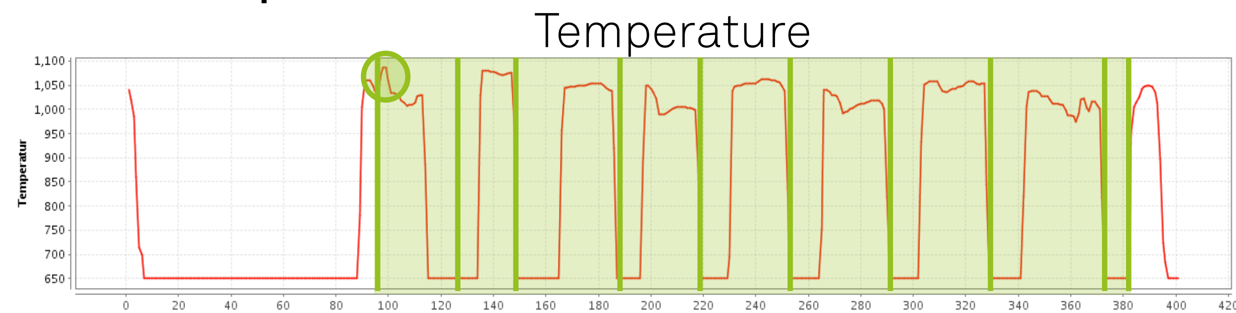
- Recording of parameters at different processing stations
- Learning of distributed models across processing stations
- Early prediction of product quality during the process

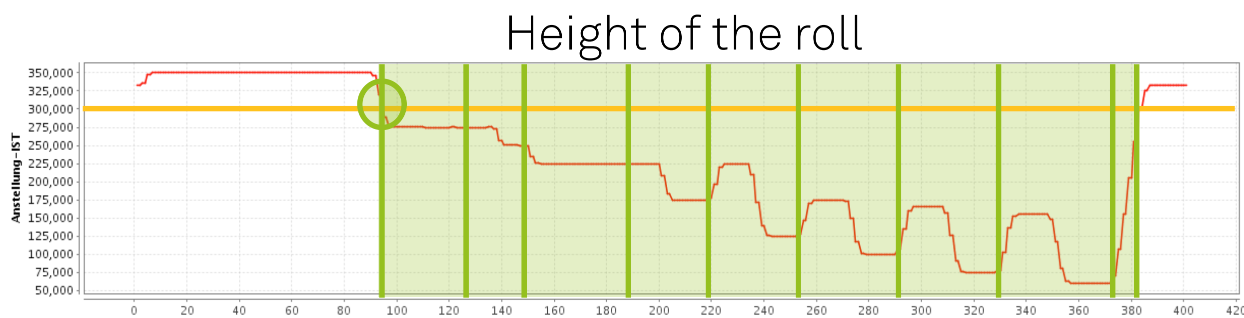
Preprocessing of time series per station

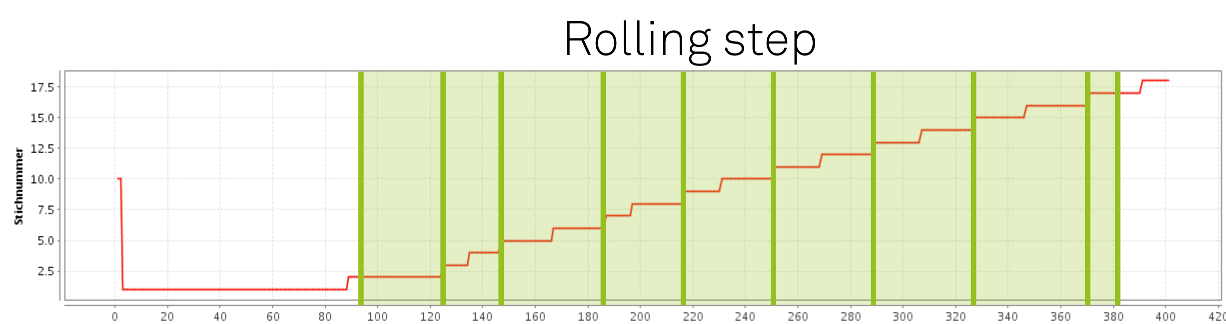
- **Outliers**
  Replace values > x

- **Cleansing**
  Focus on intervals
  roll height < 300cm

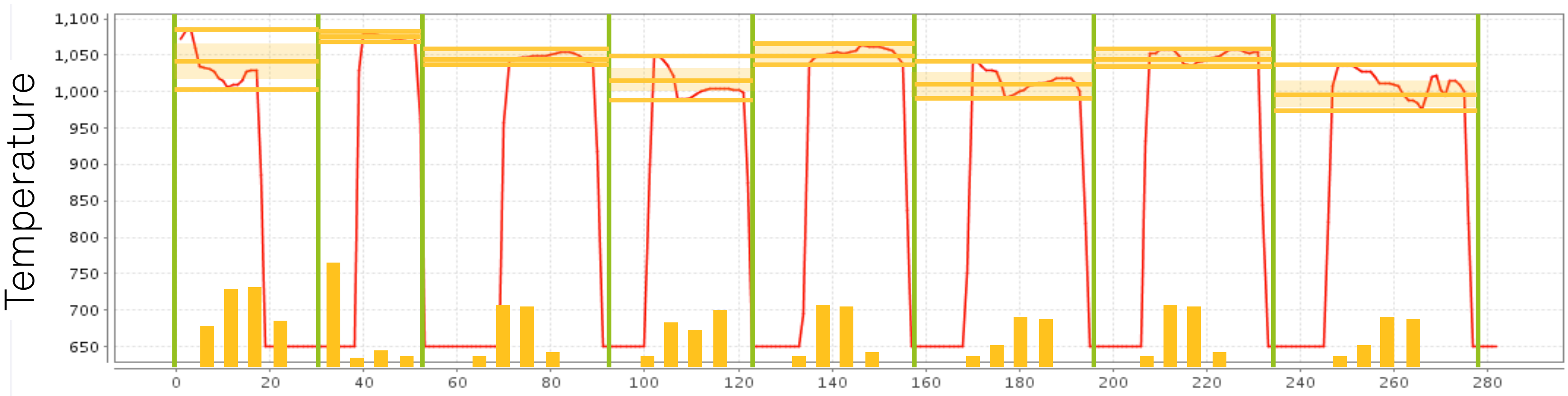
- **Segmentation**
  Divide time series
  according to
  series of
  rolling steps





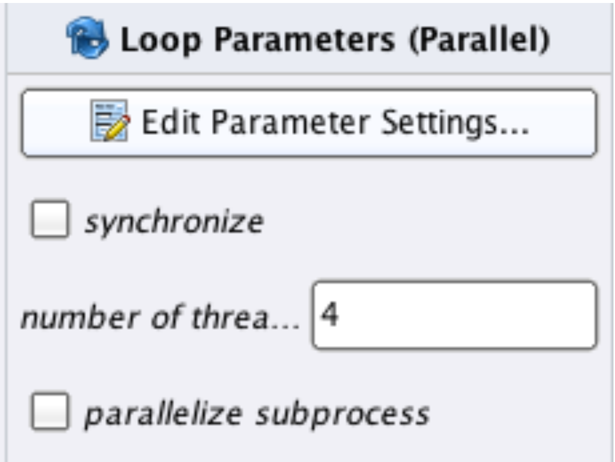


Aggregation and feature extraction

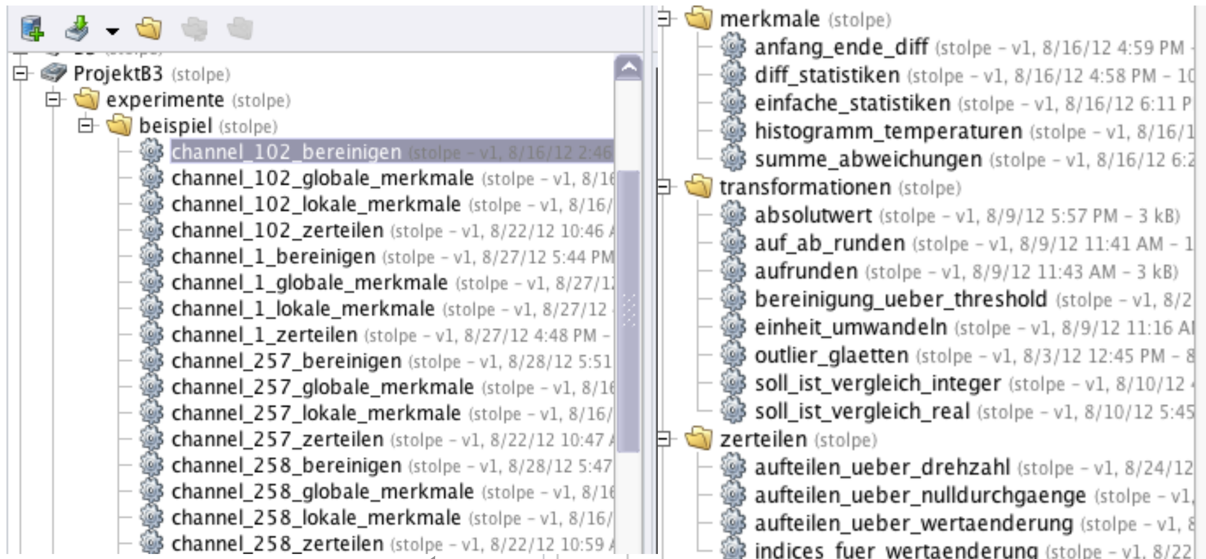


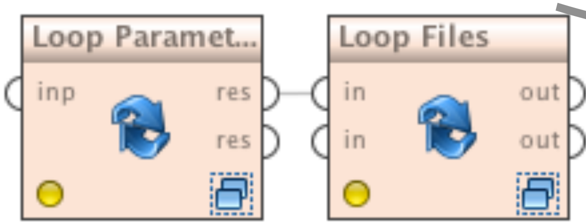
- RapidMiner offers several methods for value series:
  - Min, max, average, variance of values
  - Length, distances, frequencies of segments
  - Statistics of changes
  - Gradients
- Automatically created 60 000 features aggregated to 2 170 features, automatically selected 218 features based on classification accuracy.

RapidMiner as a tool for structured programming



Parallel processing

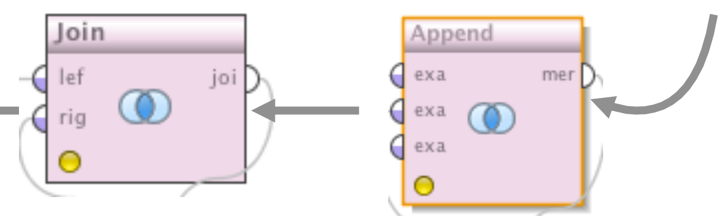




For each channel
and each time series

Call processes for cleansing
and feature extraction

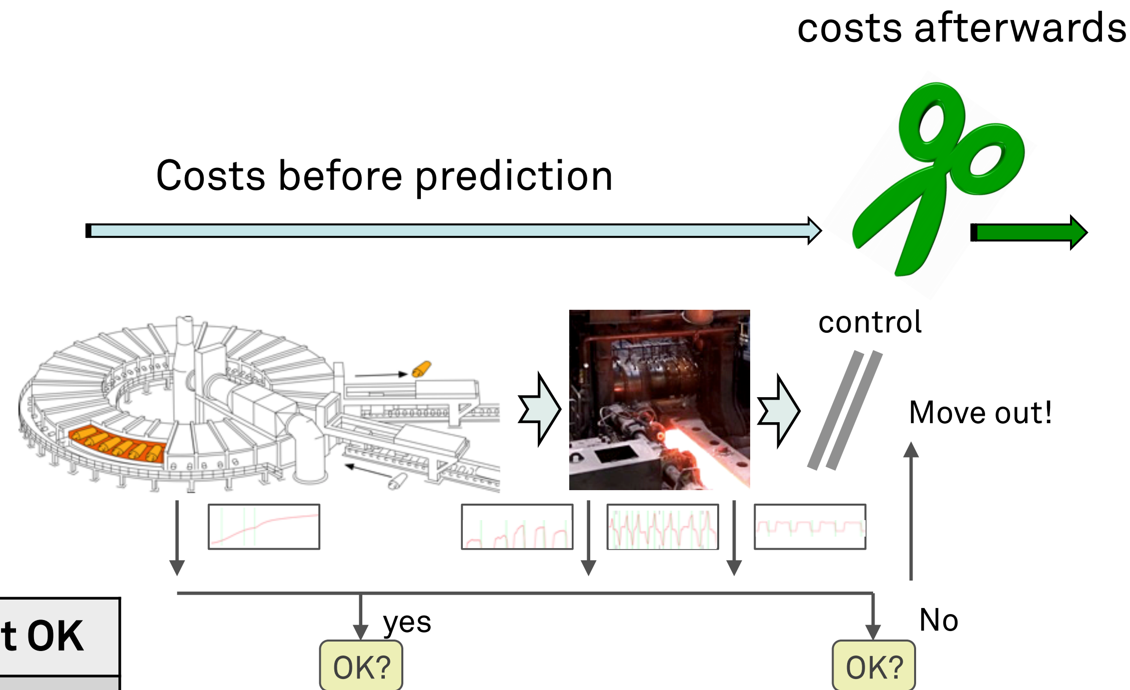
102_length	102_min	102_max	102_mean	102_stddev
198	232	1208	912.232323232	307.41643
197	177	1209	914.989847715	308.62437
195	235	1209	922.235897435	304.31125
83	234	1178	735.469879518	317.62246
84	214	1177	724.369047619	320.99717



all channels single channel

Quality prediction

- Conservative estimate:
  - If ok, say ok;
  - Minimize wrong not ok
- Future work: not only moving parts out, but adapt processing!

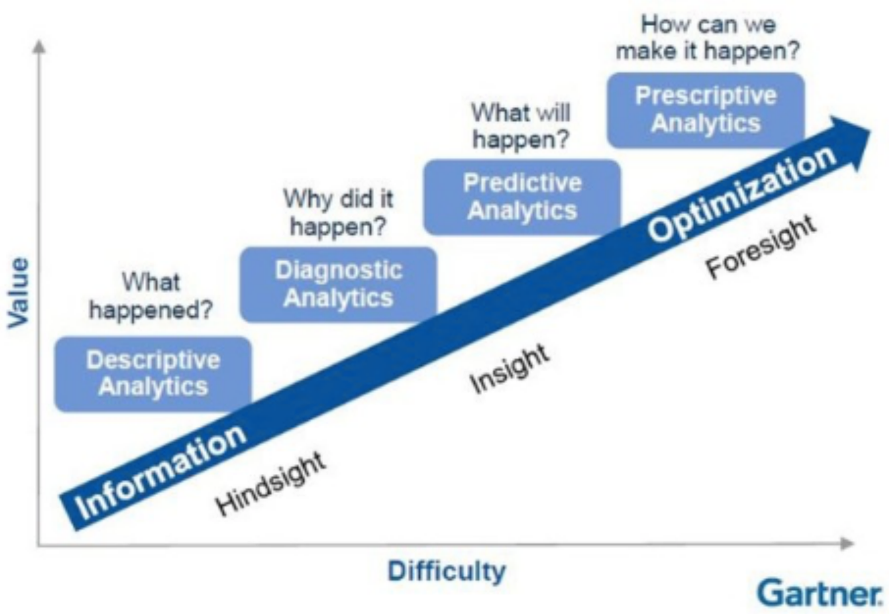


	True OK	**true not OK**
OK predicted	82%	14%
Not OK predicted	1%	**3%**

Konrad, Lieber, Deuse 2013
"Striving for Zero Defect Production: Intelligent Manufacturing Control through Data Mining in Continuous Rolling Mill Processes", in: Windt (ed) Robust Manufacturing Control, 215—229
Stolpe, Blom, Morik 2016
"Sustainable Industrial Processes by Embedded Real-Time Quality Prediction" in: Lässig, Kersting, Morik (eds) Computational Sustainability, 201—243

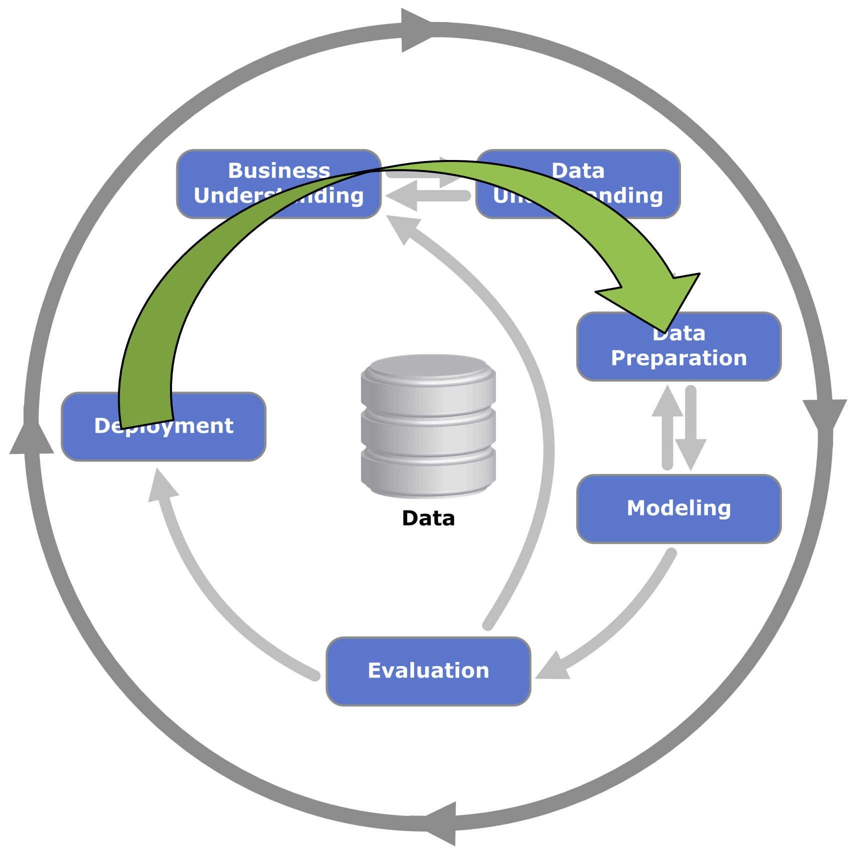
Overview

- Introduction:
  Collaborative research center
  SFB 876
  - Big data and small devices
  - Streaming data
    Astrophysics
- Anomaly detection for diagnostic analytics
- Quality prediction as predictive analytics
- Quality control by prescriptive analytics



Prescriptive analytics – Managing many models

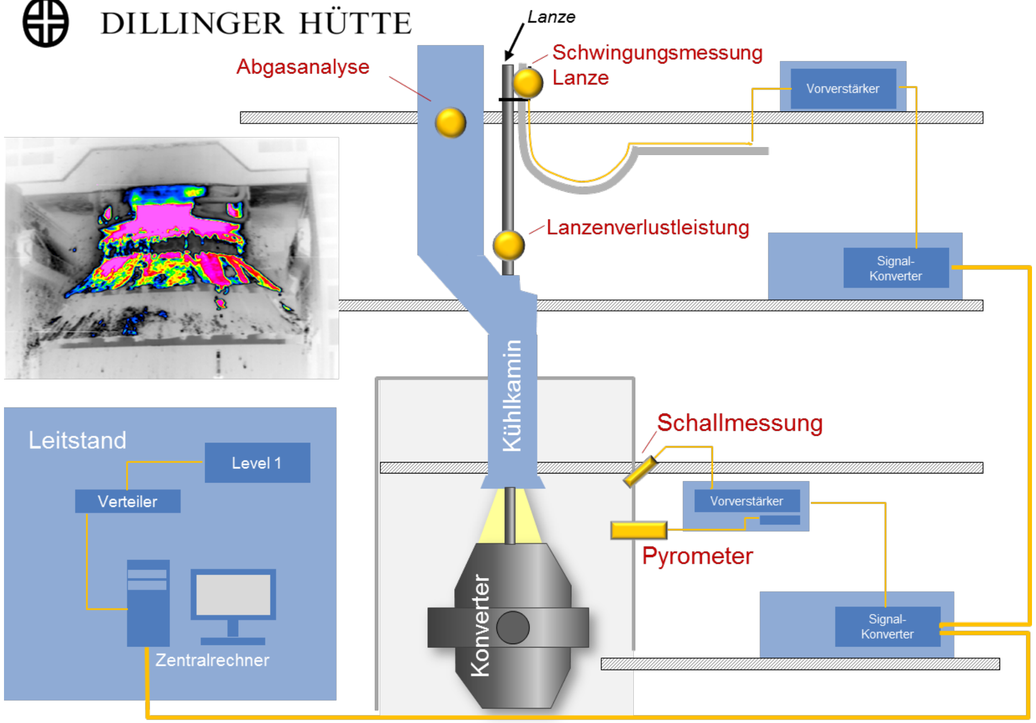
- Real-time prognosis
  - Data streams
  - Feature extraction
  - Prognosis of 4 targets each second → Process stop/continue
- Use past process data
  - Curate the process data as cleansed streams
  - Run stored process data 8000 times faster
- Use many learned models
  - Concept drift
  - Process changes



End point prediction of Basic Oxygen Furnace (BOF) converter processes

- Collaboration with SMS Siemag, Dillinger Hütte
- Converter must achieve good values of the key features T, [%C], [%P], [%Fe]
- The features cannot be measured during the process.
- Prediction of the features every second of the process.

DILLINGER HÜTTE



Model learning and validation

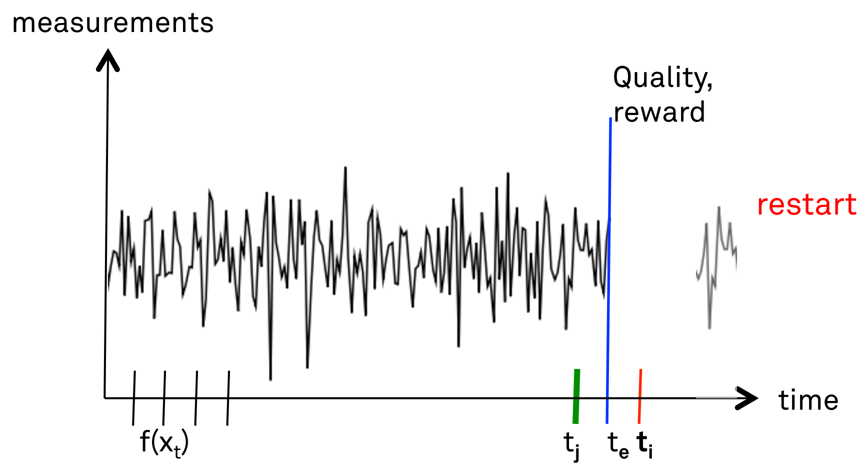
- Data Dillinger Hütte
  - 350 GB (1 year production)
  - 922 (553) charges
- Feature extraction, selection
- SVM learning offline
- Model application online
- Feature extraction online
- ONE representation for online and offline experiments, i.e. always working on streams!

- Fe: error: 2,17 %
- Temperature T: error: 18,38
- C in PPM: error: 63,36
- P in PPM: error: 29,44

- Excellent learning results – but what does it mean in terms of money?

Validation according to business impact

- $t_e < t_i$ actual before predicted
  - if process has been restarted
    – prediction right
- $t_e > t_j$ predicted before actual
  - if quality not ok
    – prediction probably right
  - if actual quality ok
    – prediction wrong?
    possibly even better?
- Look at outcome of similar cases in the past process data!
- Calculate the savings due to machine learning.



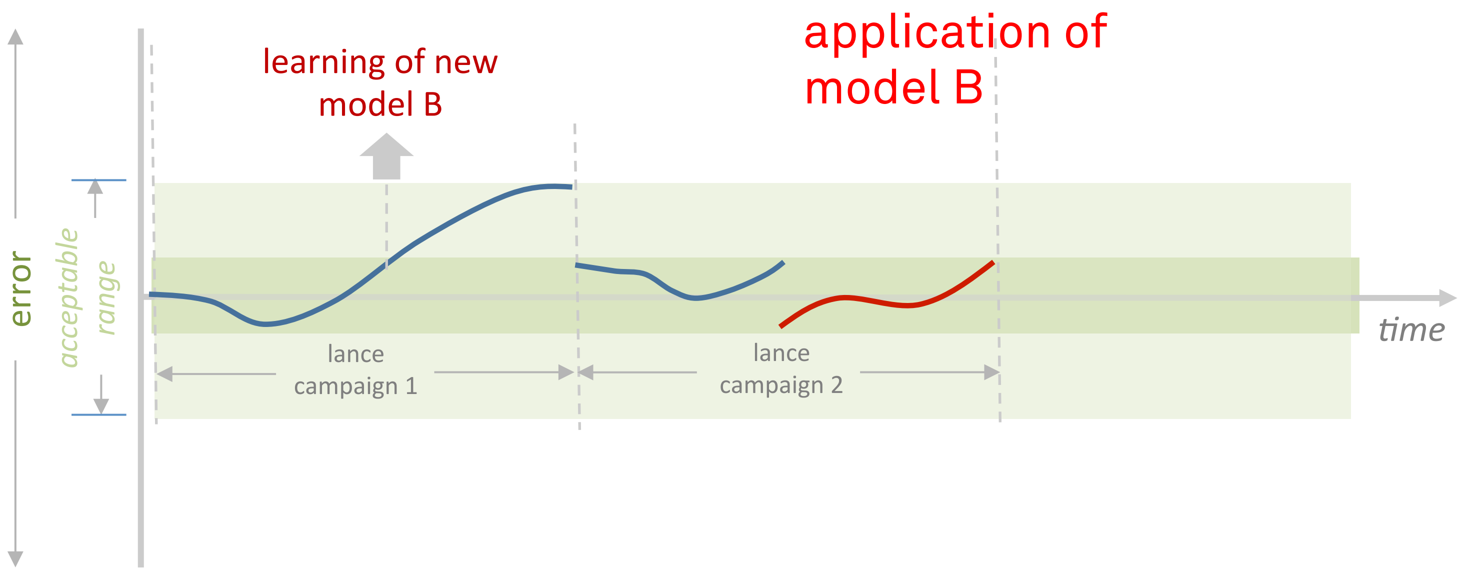
Concept drift

- Sensors break down.
- Wear of tuyeres or lance tip (age of converter lining)
  slow concept drift over a series of BOF processes.
- During processing, learning is impossible.

- Many models must be available -- ready to use!
  - Train models for missing features offline.
  - Switch to a model that does not use the missing feature online .
- Model selection in real-time!
  - Cyclic concept drift – exploit the repetition!

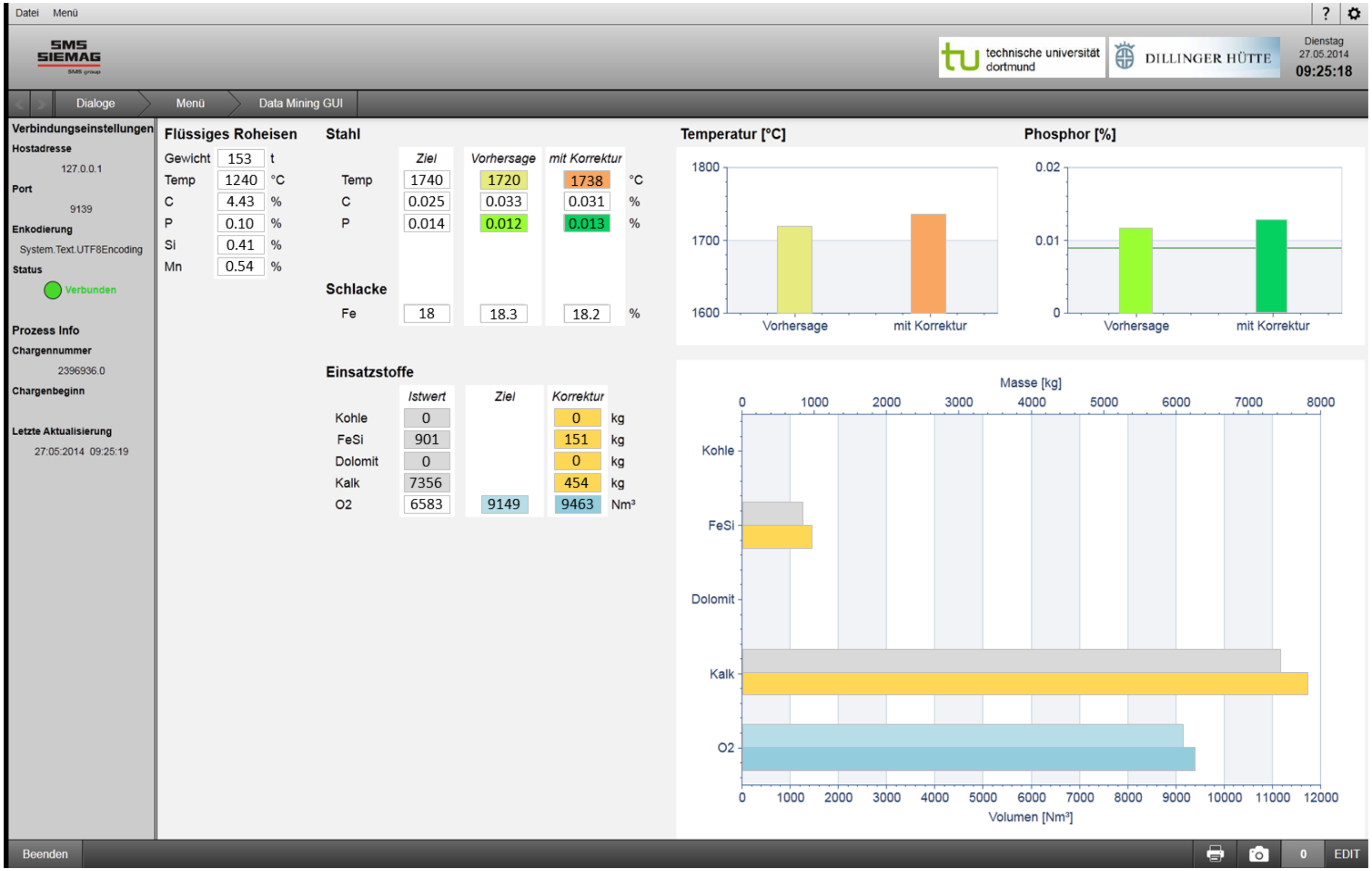
Model Management in the steel production

- When the error of predictions exceeds the acceptable range, the processes are used for learning a new model from the newer data.
- We assume a cyclic concept drift, e.g. after several processes, model A decreases and model B is better suited for the aged lance.



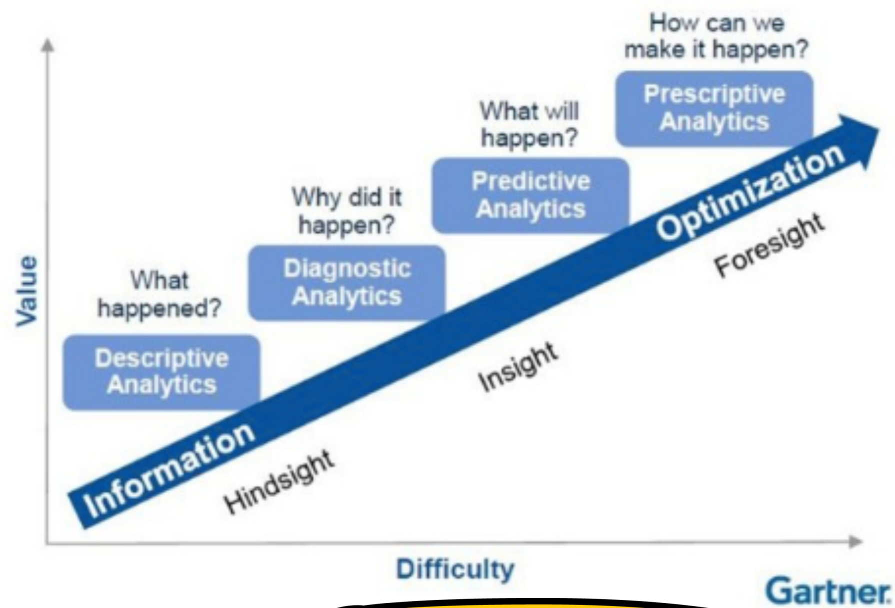
Changing process parameters accordingly

- What is the prediction, if some material is added?
- Online optimization in real-time using the set of learned models!



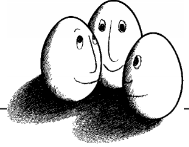
Summary

- Collaborative research center SFB 876 basic research for
  - Streaming data
    Astrophysics
- Anomaly detection
  - Feature and anomaly weighting
    Injection molding
- Quality prediction
  - RapidMiner preprocessing
    Rolling mill
- Quality control
  - Model management for concept drift and process optimization.



Literature

- Sebastian Buschjäger, Katharina Morik (2017): Decision Tree and Random Forest Implementations for Fast Filtering of Sensor Data. In: IEEE Transactions on Circuits and Systems, Vol. PP, No.99
- Marco Stolpe (2016): The Internet of Things: Opportunities and Challenges for Distributed Data Analysis. In: SIGKDD Explorations
- Marco Stolpe, Hendrik Blom, Katharina Morik (2016): Sustainable Industrial Processes by Embedded Real-Time Quality Prediction. In: Lässig, Kersting, Morik (eds) Computational Sustainability, 201—243
- Marco Stolpe, Kanishka Bhaduri, Kamalika Das, Katharina Morik (2013): Anomaly Detection in Vertically Partitioned Sensor Data by Distributed Core Vector Machines. In: ECML PKDD, 321 - 336
- Benjamin Schowe, Katharina Morik (2011): Fast-Ensembles of Minimum Redundancy Feature Selection. In: Okun, Valentini, Re (eds): Ensembles in Machine Learning Applications, 75 - 96
- Marco Stolpe, Katharina Morik (2011): Learning from Label Proportions by Optimizing Cluster Model Selection In: Procs. European Conf. Machine Learning, Springer, 349 – 364
- Katharina Morik, Michael Kaiser, Volker Klingspor (eds) (1999): Making Robots Smarter -- Combining Sensing and Action through Robot Learning, Kluwer



Literature

- Mario Wiegand, Marco Stolpe, Jochen Deuse, Katharina Morik (2016): Prädiktive Prozessüberwachung auf Basis verteilt erfasster Sensordaten. In: at Automatisierungstechnik, Vol.64, No. 3, 521 - 533
- Katharina Morik, Hendrik Blom, Nobert Uebber, Tobias Beckers, Hans-Jürgen Odenthal, Jochen Schlüter (2014): Reliable BOF Endpoint Prediction by a Real-Time Data-Driven Model. In: AIST Indianapolis
- Norbert Uebbe, Hans Jürgen Odenthal, Jochen Schlüter, Hendrik Blom, Katharina Morik (2013): A novel data-driven prediction model for BOF endpoint. In: The Iron and Steel Technology Conference and Exposition (AIST) Pittsburgh
- Katharina Morik, Hendrik Blom, Hans-Jürgen Odenthal, Norbert Uebbe (2012): Resource-Aware Steel Production Through Data Mining. In: SustKDD workshop at KDD, Peking, August 2012
- Daniel Lieber, Benedikt Konrad, Jochen Deuse, Marco Stolpe, Katharina Morik (2012): Sustainable Interlinked Manufactoring Processes through Real-Time Quality Prediction. In: CIRP
- Norbert Uebbe, Hans Jürgen Odenthal, Jochen Schlüter, Hendrik Blom, Katharina Morik (2012): A novel data-driven prediction model for BOF endpoint. In: 30th International Steel Industry Conference in Paris (JSI)