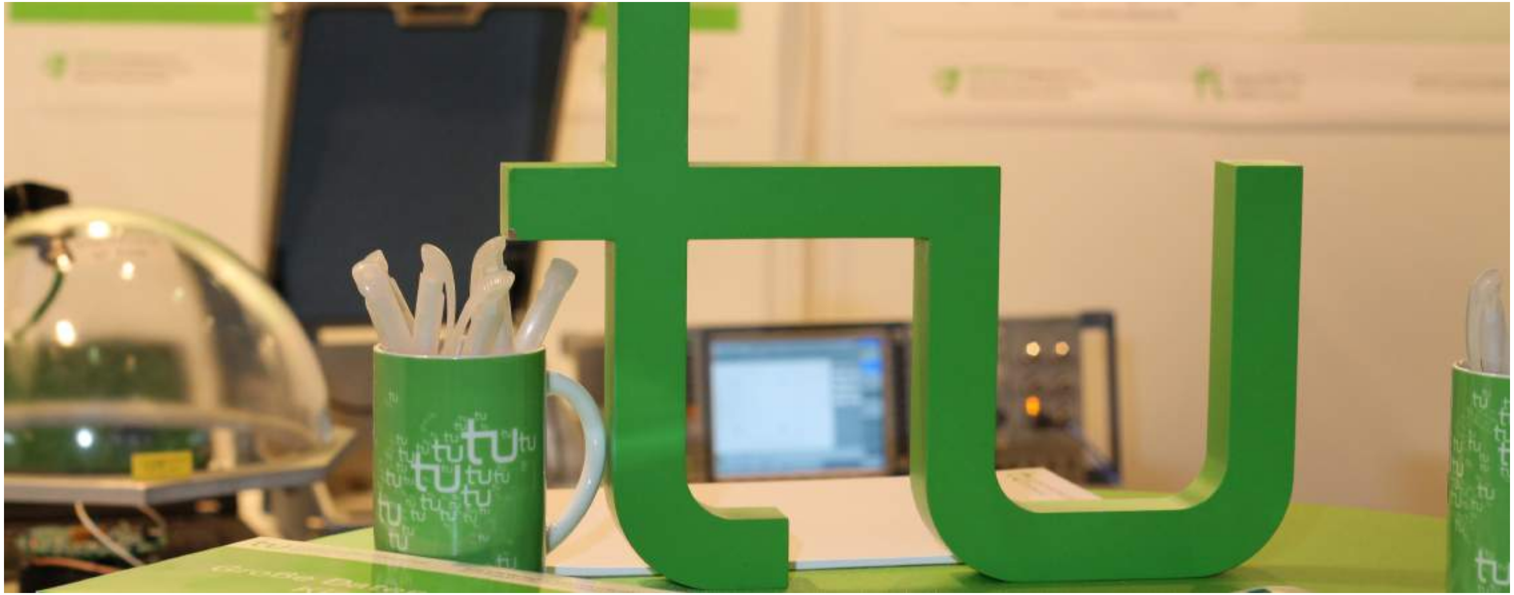
technische universität dortmund

SFB 876 Verfügbarkeit von Information durch Analyse unter Ressourcenbeschränkung

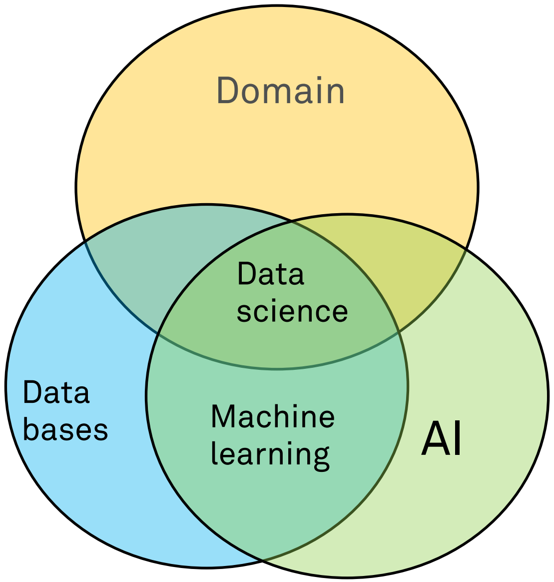


What is the impact of AI for industrial systems?

Katharina Morik, Computer Science 8, TU Dortmund

Overview

- Artificial Intelligence
- Machine Learning
- Applications
- What to do



Artificial Intelligence

- AI is about *computer* behavior that – if performed by *humans* – would be considered intelligent.
- It is about a <u>task</u> – humans perform many tasks!
- What we consider "intelligent":
  - Calculation, memorizing is no longer considered intelligent. Because *computers* are good at it?
  - Tying shoelaces is very difficult for robots. Are we as *humans* proud of this skill?

Artificial Intelligence

- AI is about *computer* behavior that – if performed by *humans* – would be considered intelligent.
- It is about a task – humans perform many tasks!
- What we consider “intelligent”:
    - Calculation, memorizing is no longer considered intelligent. Because *computers* are good at it?
    - Tying shoelaces is very difficult for robots. Are we as *humans* proud of this skill?

- AI since 1956 (USA), 1976 (Germany)
- AI as a model of cognition
- AI as a performer
- AI as an analogy or for science fiction

Artificial Intelligence Classics

- Planning
- Logical Inference, Reasoning
- Knowledge Representation
- Natural Language Processing
- Machine Learning
- Multi-Agent Systems
- Robotics
- Cognition
- Vision Understanding, Computer Graphics
- Games

Artificial Intelligence Classics

- Planning
- Logical Inference, Reasoning
- Knowledge Representation
- Natural Language Processing
- Machine Learning
- Multi-Agent Systems
- Robotics
- Cognition
- Vision Understanding, Computer Graphics
- Games

1996 DeepBlue defeats Kasparow – AI without machine learning

Machine Learning is

- Part of computer science with applications in all other areas.
- Based on data
  - Well acquired data (Excel)
  - Given data bases (SQL)
  - Big data (Hadoop distributed file system HDFS)
  - Structured data (graphs, facts)
- Implemented on an architecture
- Delivering an action.

Machine Learning is

- Part of computer science with applications in all other areas.
- Based on data
  - Well acquired data (Excel)
  - Given data bases (SQL)
  - Big data (Hadoop distributed file system HDFS)
  - Structured data (graphs, facts)
- Implemented on an architecture
- Delivering an action.

Classification

- Given:
  examples $\{(x_1, y_1), (x_2, y_2)\ldots\}$
- Find:
  predictions $f: X \rightarrow Y$,

  such that the quality $Q(Y, f(X))$ is maximized.

- Risk:

$$R(Y, f(X)) = \sum_{i=1}^{N} Q(y_i, x_i)p_i$$

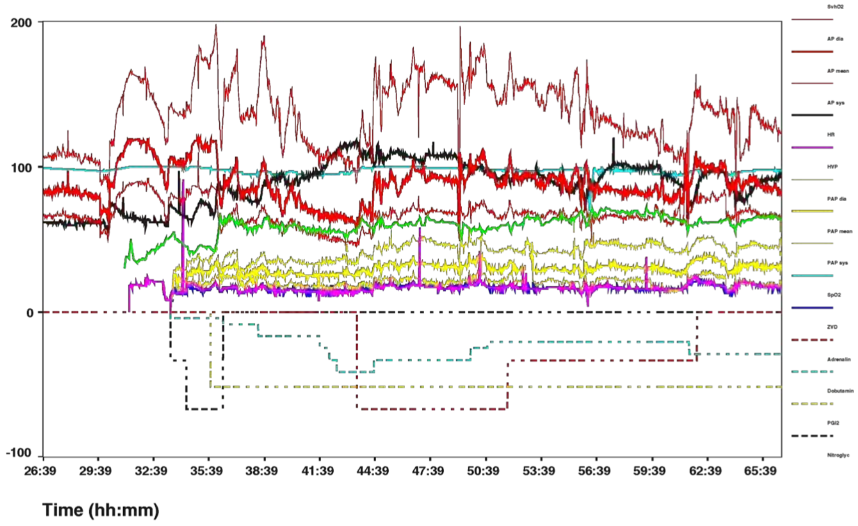
Machine Learning Algorithms

- Induction of Decision Trees
- Support Vector Machines
- Clustering
- Probabilistic Graphical Models
- Frequent Itemset Mining
- Reinforcement, Q Learning
- Neural Networks, Convolutional Neural Networks
- Time Series Classification, Clustering, Prediction

2016 alphaGo defeats Lee Sedol –
AI with machine learning

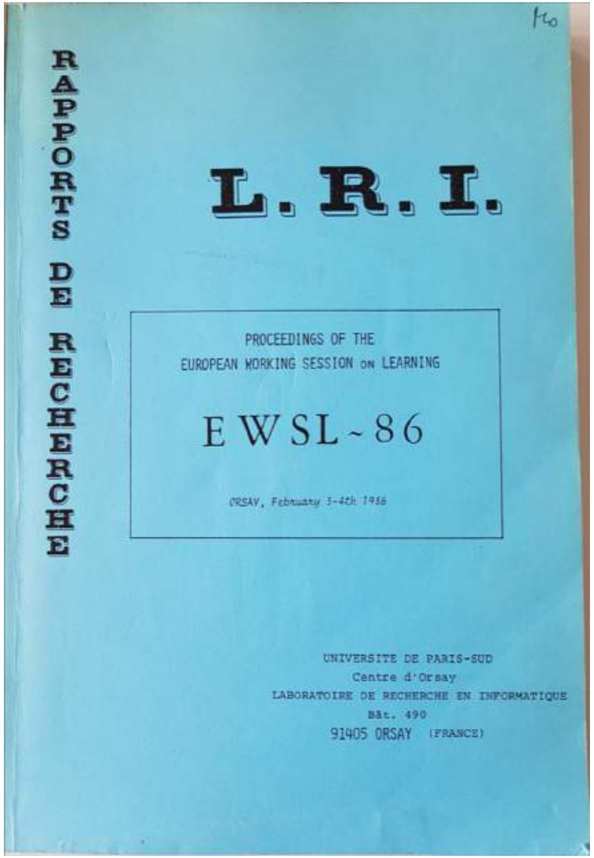
Why Classification is an action

- Learn from a large data base of intensive care treatments, when to give a certain drug.
- Discuss learned f(x) with the doctor.
- Apply learned classifier every minute
  - Given the vital parameters of a patient,
  - Find whether to give drug y=+1 or not y=-1,
  - such that the risk is minimal.



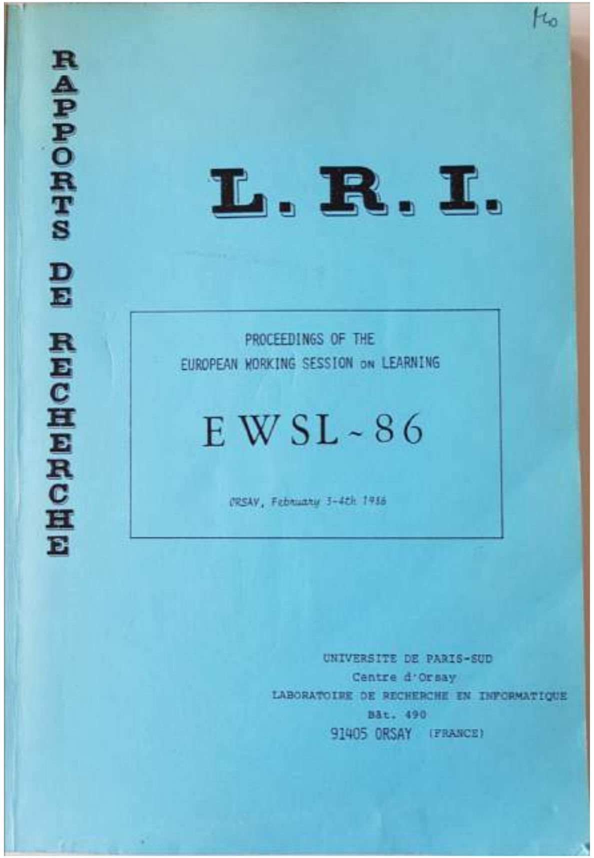
Short History of European Machine Learning – Logic

- Algorithms that generalize facts, form clusters, propagate weights in networks
  - Library of learning in Prolog
    Meta-programming in Prolog
- Summer school on learning with *participants' practical problems* (Les Arcs 1988)



Short History of European Machine Learning – Logic

- Algorithms that generalize facts, form clusters, propagate weights in networks
  - Library of learning in Prolog
    Meta-programming in Prolog
- Summer school on learning with *participants' practical problems* (Les Arcs 1988)
- Balanced cooperative modeling of man and machine: explainability, revisability, inspectability. System MOBAL
  Morik, Emde, Kietz, Wrobel 1993



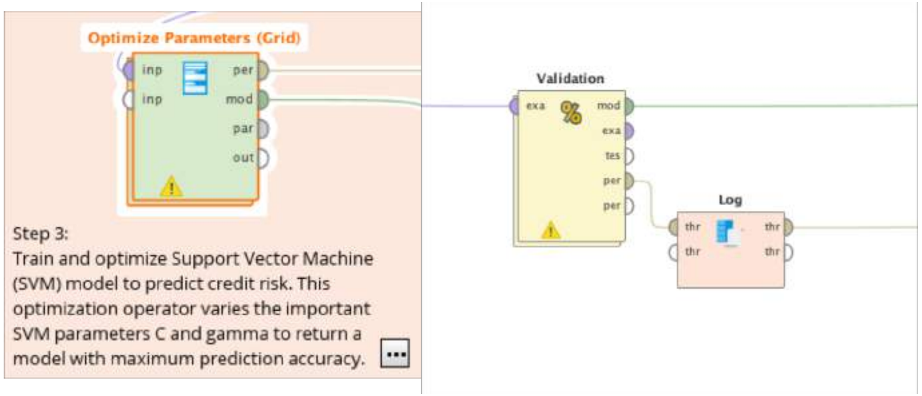
Short History… – Learning to Learn

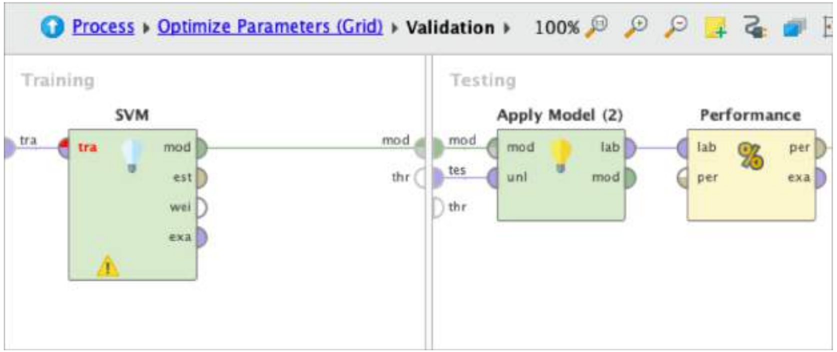
- Meta-data driven processes
  - Learning new meta data
    means new processes

Short History… – Learning to Learn

- Meta-data driven processes
  - Learning new meta data means new processes
- Learning for a learning algorithm
  - best parameters
  - best examples
  - best features
  - …
- Embedded learning processes in learning processes

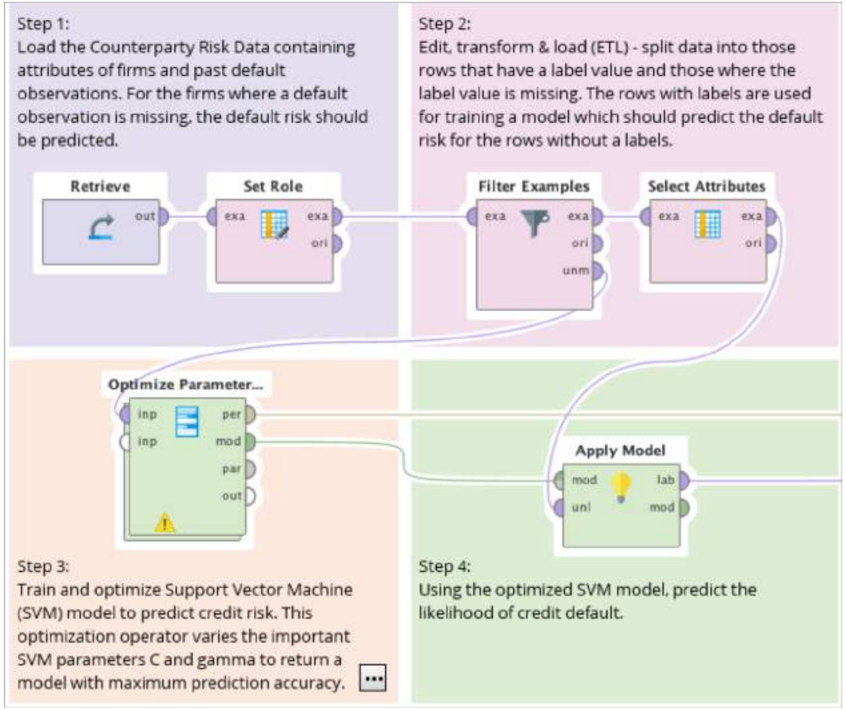
RapidMiner 2002





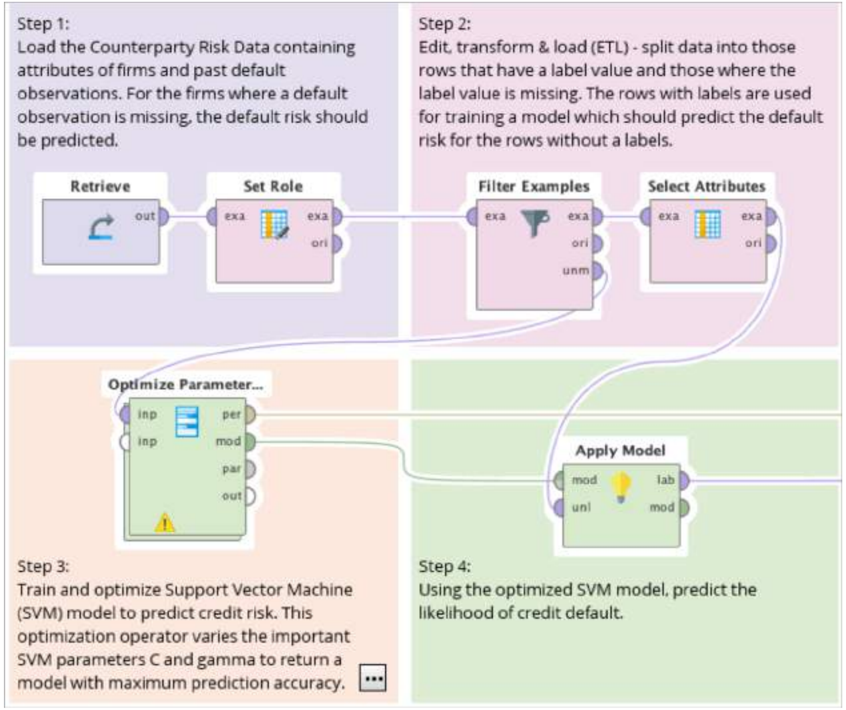
Short History of European Machine Learning – Pipelines

- Long pipeline from entering data to delivering result.



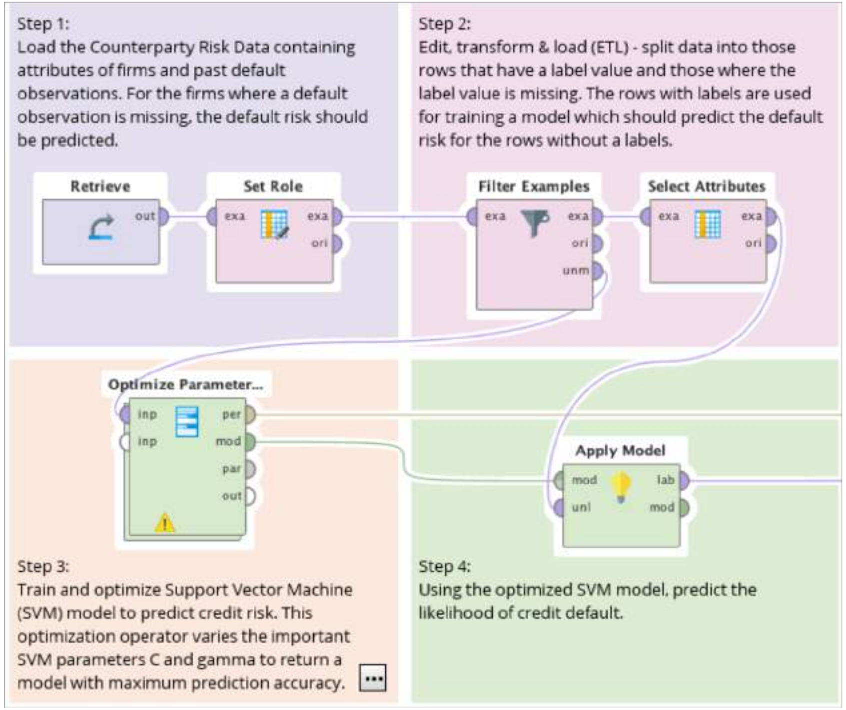
Short History of European Machine Learning – Pipelines

- Long pipeline from entering data to delivering result.
- Every step of the pipeline can be enhanced through learning.
  - Sampling through active learning
  - Missing value learning
  - Feature construction
  - Feature selection
  - ...



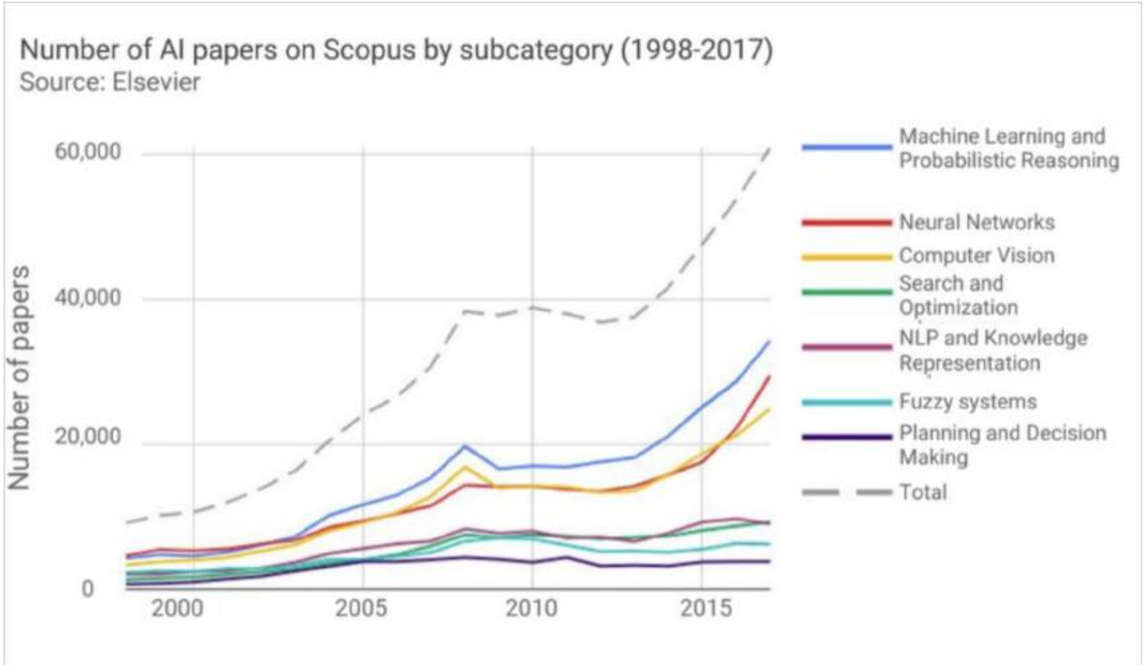
Short History of European Machine Learning – Pipelines

- Long pipeline from entering data to delivering result.
- Every step of the pipeline can be enhanced through learning.
  - Sampling through active learning
  - Missing value learning
  - Feature construction
  - Feature selection
  - ...
- Learning the best learnig algorithm for a task and data.



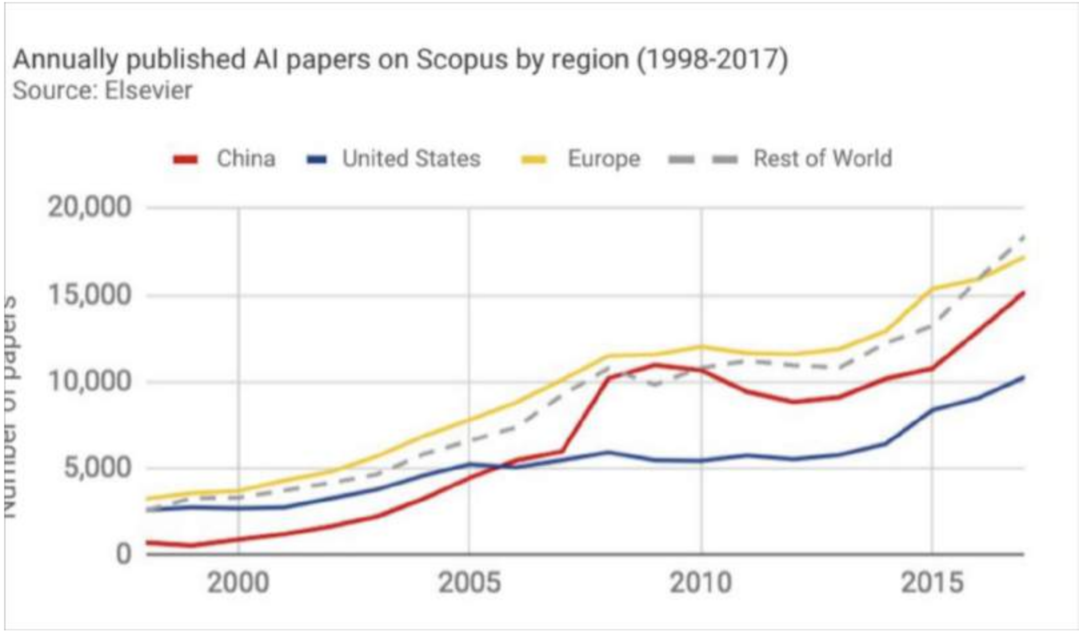
Stanford AI Index 2018 – AI Subcategories

- Machine learning even without neural networks has the largest number of publications and a steep ascent.
- Neural networks has the steepest ascent.
- Within the growing computer vision papers, many deal with learning.



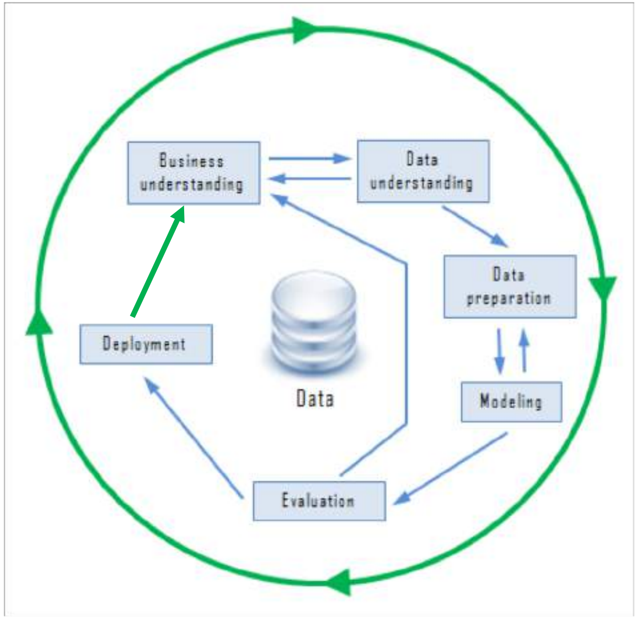
Stanford AI Index 2018 – Regions

- Europe is extremely good at publishing.
- That is: our research is excellent.



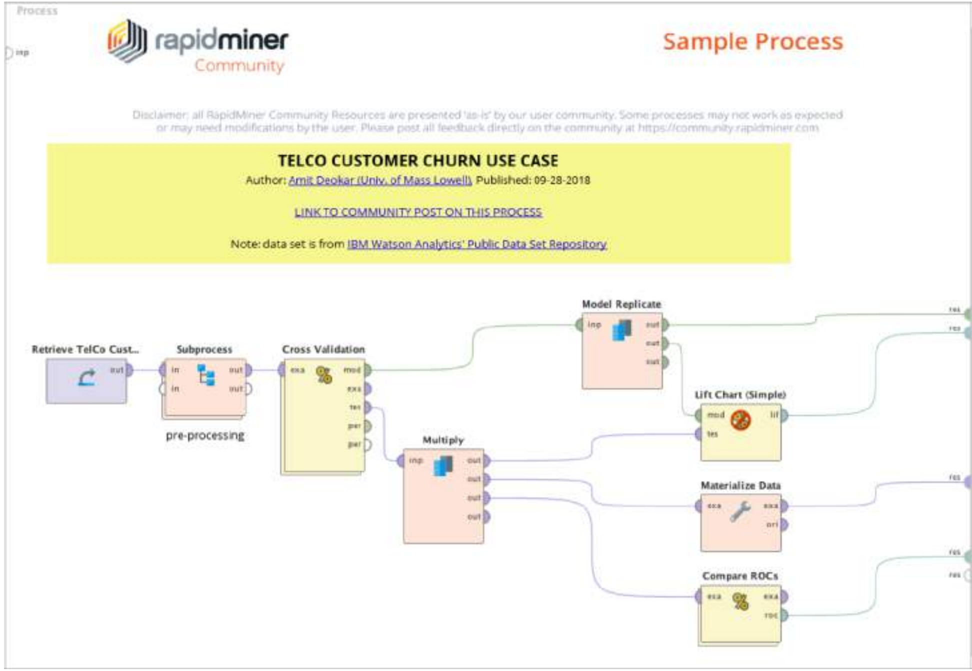
Learning is everywhere

- In each step of the data analysis process, learning can be applied.
- Hence, many learning algorithms together deliver the performance.
- This is the reason of success.
- This makes explainability, revisability, inspectability a problem.

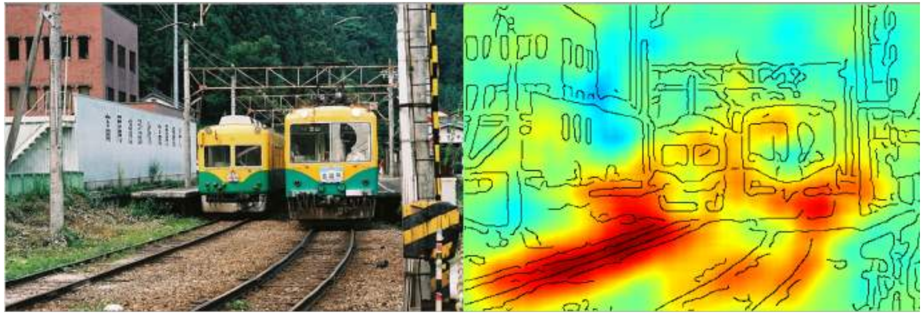


Toolboxes and libraries make learning easy for everyone

- RapidMiner allows to design an overall pipeline for your data, fixes bugs in it, verifies the results and uses implementations that are strictly based on scientific terms.
- Libraries like PyTorch or frameworks like TensorFlow ease programming.
- No experts needed?
  For designing a new application – you do need experts!



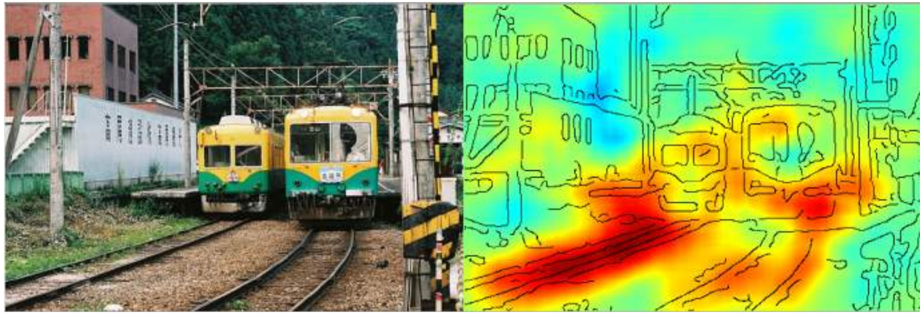
Veracity, Trustworthiness, Robustness



Klaus-Robert Müller (TU Berlin), Wojciech Samek (HHI, Berlin) developed a certification method for Deep Learning. Here, it shows by heat maps that the classification is based on features of the environment, not of the object. This phenomenon is well known for DeepNeural Networks. DOI: 10.1038/s41467-019-08987-4

Veracity, Trustworthiness, Robustness

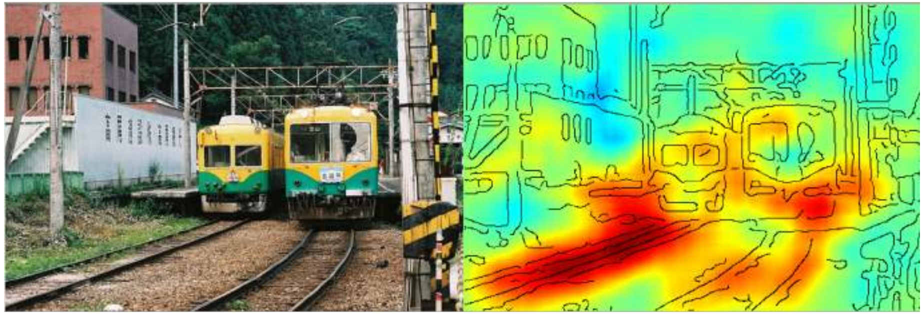
- Machine learning algorithms need a sound theoretical basis.
  - Proven tight quality guarantees.
  - Robust against minor changes in the input data.
  - Verified combination of methods.



Klaus-Robert Müller (TU Berlin), Wojciech Samek (HHI, Berlin) developed a certification method for Deep Learning. Here, it shows by heat maps that the classification is based on features of the environment, not of the object. This phenomenon is well known for DeepNeural Networks. DOI: 10.1038/s41467-019-08987-4

Veracity, Trustworthiness, Robustness

- Machine learning algorithms need a sound theoretical basis.
  - Proven tight quality guarantees.
  - Robust against minor changes in the input data.
  - Verified combination of methods.
- Applying machine learning requires to know the theoretical foundations of machine learning.



Klaus-Robert Müller (TU Berlin), Wojciech Samek (HHI, Berlin) developed a certification method for Deep Learning. Here, it shows by heat maps that the classification is based on features of the environment, not of the object. This phenomenon is well known for DeepNeural Networks. DOI: 10.1038/s41467-019-08987-4

Overview

- Artificial Intelligence
  - Classical fields continue and now use machine learning
- Machine Learning
  - Many algorithms
  - Long pipelines
  - Complex interactions of learning (learning to learn, ensembles)
  - Theory, certification needed!
- **Applications**
- What to do

Applications

- Medicine
- (Astro-)Physics
- Telecommunication
  - Security policy
  - Customer churn
- Internet of Things
- Recommender systems
- Chemical plants configuration
- Injection molding
- Steel production and milling
- Tunneling
- Turning and milling

Injection Molding – Supervised feature selection

- 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

Injection Molding – Supervised feature selection

- Minimum Redundancy Maximum Relevance feature selection requires data with labels: $\langle \mathbf{x}, y \rangle$
- Most observations are not labeled.
- Using domain knowledge by asking the expert? Each **x**?!

- 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

Injection Molding – Supervised feature selection

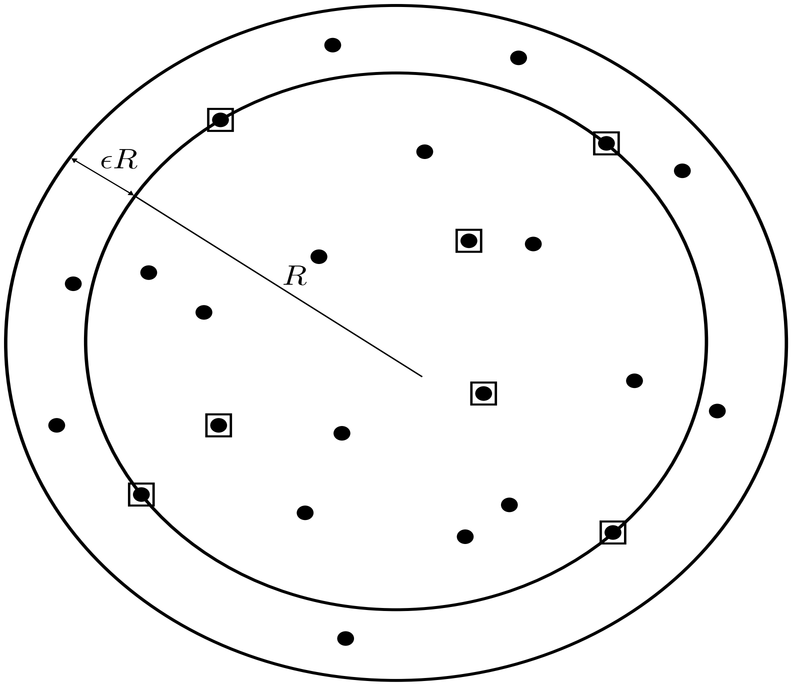
- Minimum Redundancy Maximum Relevance feature selection requires data with labels: ‹**x**, y›
- Most observations are not labeled.
- Using domain knowledge by asking the expert? Each **x**?!
- Using domain knowledge indirectly!
  - Known causalities label observations f(**x**) = y
    e.g., y=max injection 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

Injection Molding -- Unsupervised feature selection

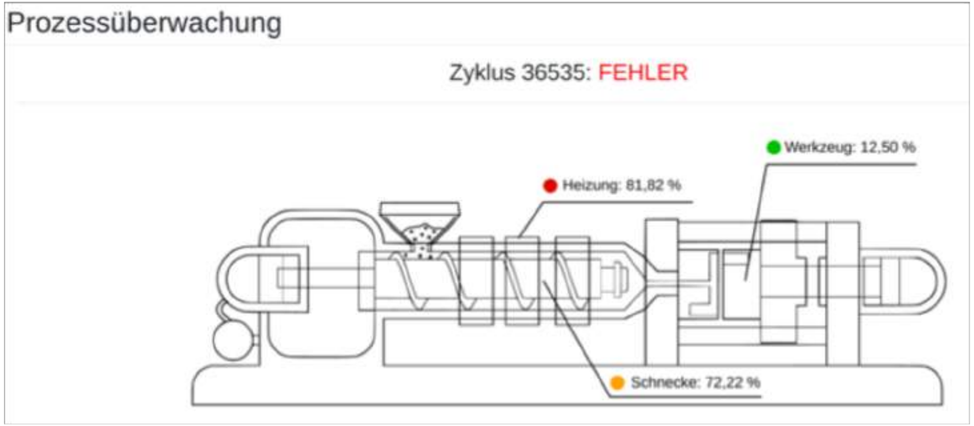
- Single class SVM
  - Outliers are anomalies
  - SVM ranks features according to their contribution to the decision
- Multi-objective optimization clustering
  - Members in a cluster are close to each other
  - Few clusters
  - Few not assigned observations
  - Members of different clusters are very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



Single class SVM, minimum enclosing ball

Weighting of features, weighting of anomalies

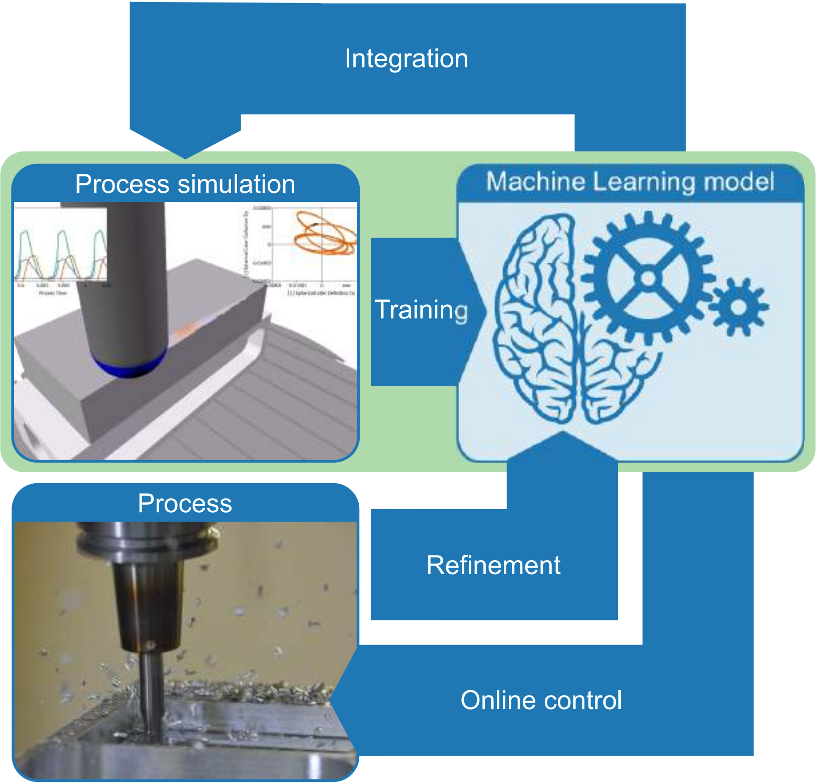
- Evolutionary process delivers several feature sets, each is used for clustering.
- For all clusterings:
  Large clusters are considered normal.
  Small clusters show anomalies.
- Features that are often used in large clusters
  receive a higher weight.
- Anomalies that are found by many clusters
 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



Reinhard Schiffers, Katharina Morik,
Alexander Schulze-Struchtrup, Philipp-Jan Honysz,
Johannes Wortberg (2018) „Anomaly detevtion in injection molding process data based on unsupervised learning“
Journal of Plastics Technology, No. 5, p 301 - 347

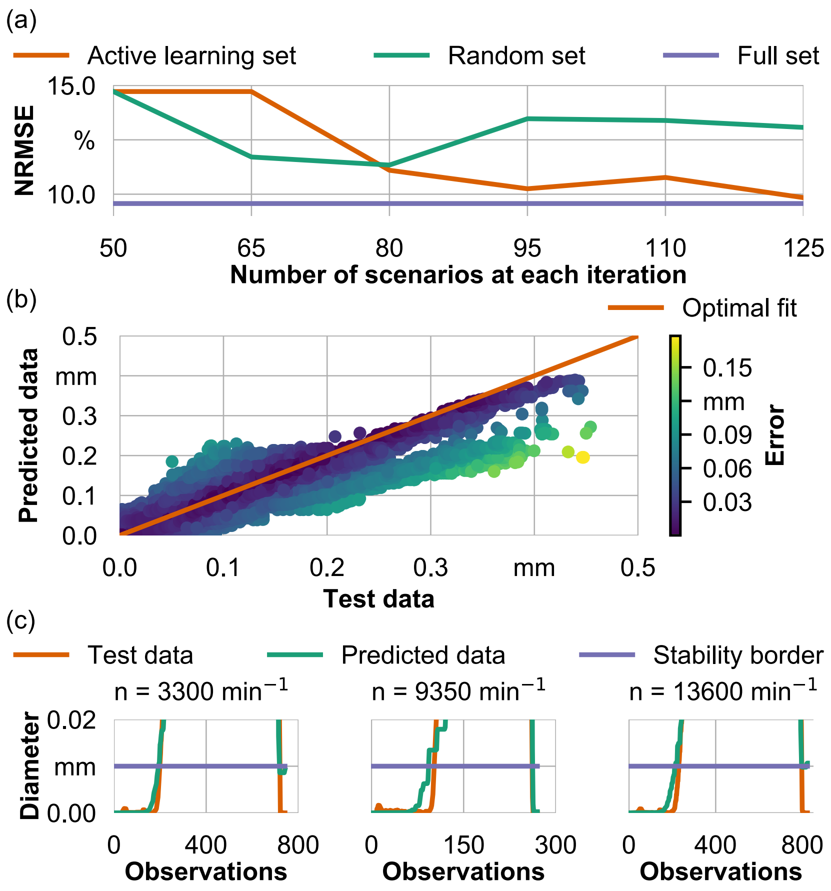
Interaction of Simulation and Milling

- Predict cutting forces or tool vibrations in order to stabilize the milling by a multilayer perceptron.
- Simulation delivers the label y for a process.
- Active learning orders few, well selected simulations.
- Learned model is applied in real-time.
- Amal Saadallah, Felix Finkeldey, Katharina Morik, Petra Wiederkehr „Stability prediction in milling processes using a simulation-based machine learning approach“ CIRP 2018



Results

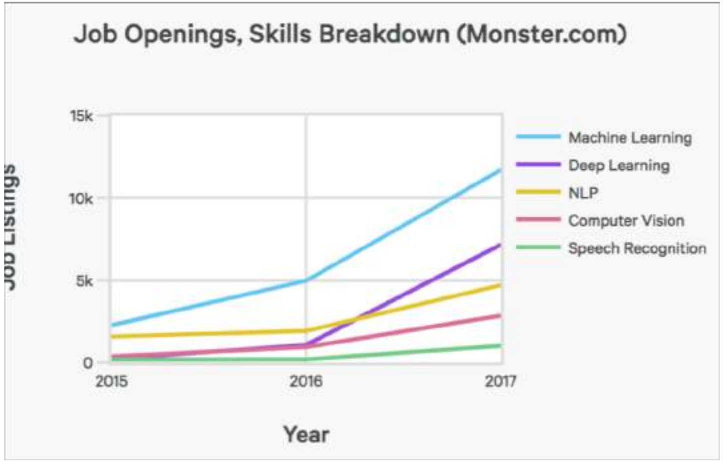
- Error is reduced with less simulation runs.
- Prediction is near optimal.
- The new method can be applied in real time.
- The instability is predicted early or in time for three different spindle speeds.



Overview

- Artificial Intelligence
- Machine Learning
- Applications
  - Anomaly detection for injection molding
  - Complex control of milling
    - Active learning for acquiring the right labels
    - Ensembles of learned models for predicting forces
- **What to do**

Growth of the Field, Demand of Skills



https://aiindex.org

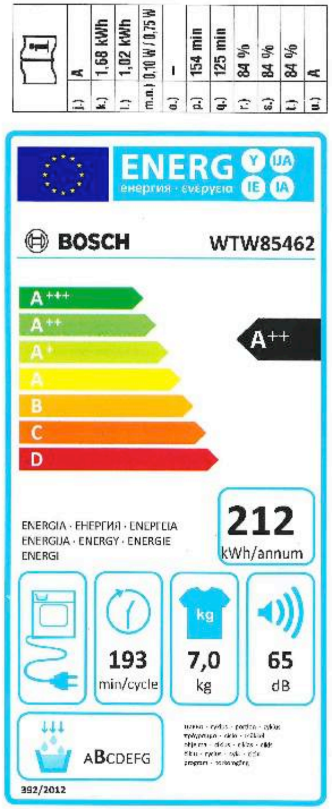
- 457 Open positions in Germany Maschinelles Lernen
- 470 Open positions in Germany Machine Learning
- 154 Deep Learning
- 118 Natural Language Processing incl. Speech
- 200 Image, Vision

(Monster.de am 23.9.2018)

- From 2652 professors for computer science less than 60 professors for machine learning in Germany (2017).

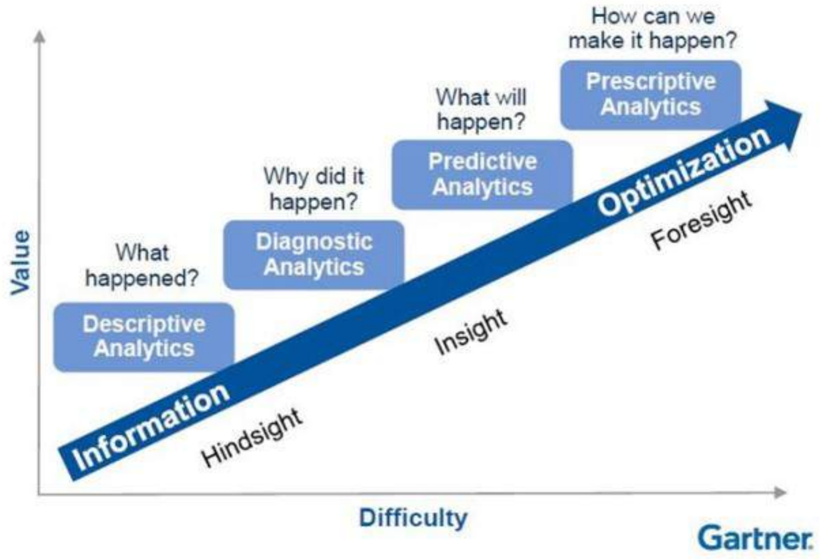
What to do in research?

- Making learning robust and inspectable
- Verifying machine learning at all levels
  - Statistical models
  - Algorithms
  - Software frameworks
  - Computer architectures
- Balancing resource and error bounds
- ➢ Certify algorithms like other electronic devices.



What to do in applications?

- Paying machine learning experts for the first case study
  - Determining where prediction brings highest benefit
    (in terms of ROI, business understanding)
  - Gathering data and taking time for data preparation
- Hiring machine learning experts
- Neither over- nor underestimate AI!



Overview

- Artificial Intelligence
- Machine Learning
- Applications
- What to do:
  - Research is needed.
  - Education at all levels!
  - Sustainable investments offer perspectives for scientists, experts and allows to manage the change in companies.