Business Process Management

Process Mining

Challenges with BPM (1/4)

- Creation of “current state” processes: a description of how a business process is being performed today
  - “as-is”
  - organizations are primarily interested in an improved “to be” process
- Lack of connections between business processes and an organization’s enterprise information systems (SAP, Navision, …).

Challenges with BPM (2/4)

- As-is status is critical to knowing
  - whether it is worth investing in improvements
  - where performance problems exist
  - how much variation there is in the process across the organization

Challenges with BPM (3/4)

- Some enterprise systems (e.g. SAP) are process-oriented:
  - they support processes like order-to-cash or procure-to-pay
  - but there is rarely an easy way to understand how the process is being executed from the information system.

Challenges with BPM (4/4)

- Other technologies (such as Microsoft's Visio or Software AG's Aris) support aspects of process design.
  - But if you want information about how your process is performing day to day, you require a difficult set of manual steps to gather and synthesize data.
- Many process improvement approaches (e.g. Lean and Six Sigma) have not emphasized on information technologies as enablers of processes or of process management.

Solution: process mining

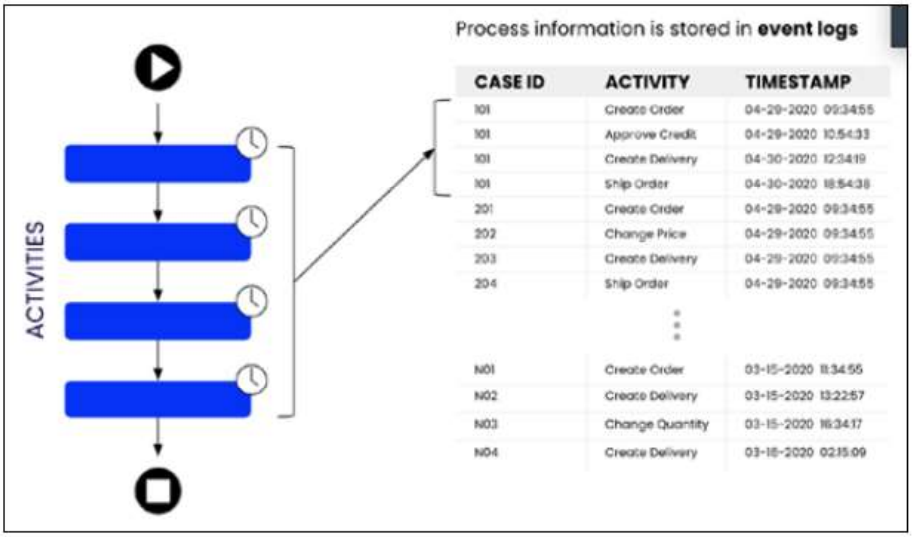
- Has been an academic topic for many years.
- 2011: Celonis (Munich) was founded
  - Collects data from SAP, Oracle, Salesforce, and ServiceNow
  - Other systems through APIs.
- Other vendors:
  - Fluxicon (NL): most popular stand-alone analysis focused tool
  - QPR (Finland)

Process Mining Software

- Capture information from enterprise transactions.
- Provides detailed — and data-driven — information about how key processes are performing.
- Creates event logs as work is done: an order is received, a product is delivered, a payment is made.

Process Mining: Event logs

Event logs make visible how computer-mediated work is really happening, including who did it, how long it takes, and how it departs from the average.



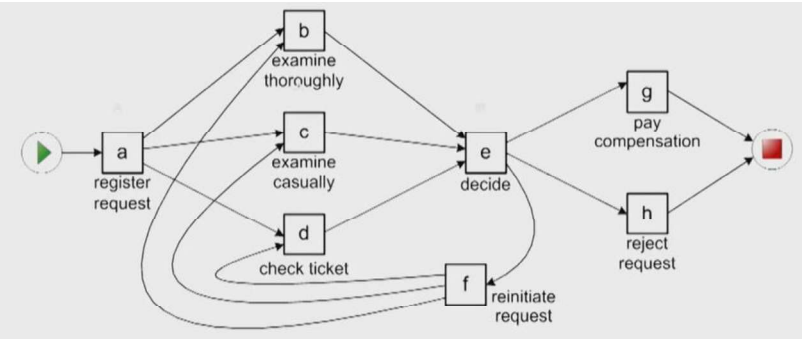
Applications of process mining

- Process discovery
- Conformance checking
- Deep dive data mining

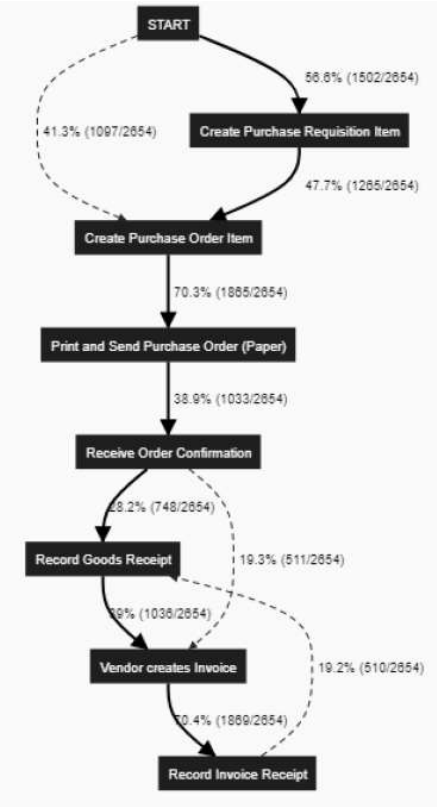
Process Discovery

- Process flow (all possible paths) are reconstructed.
- Based on event logs.
- Alternative for interviews of stakeholders.
- Works only for processes that are mainly digitized.
- Can serve as basis for BPMN diagram

Process Discovery: examples

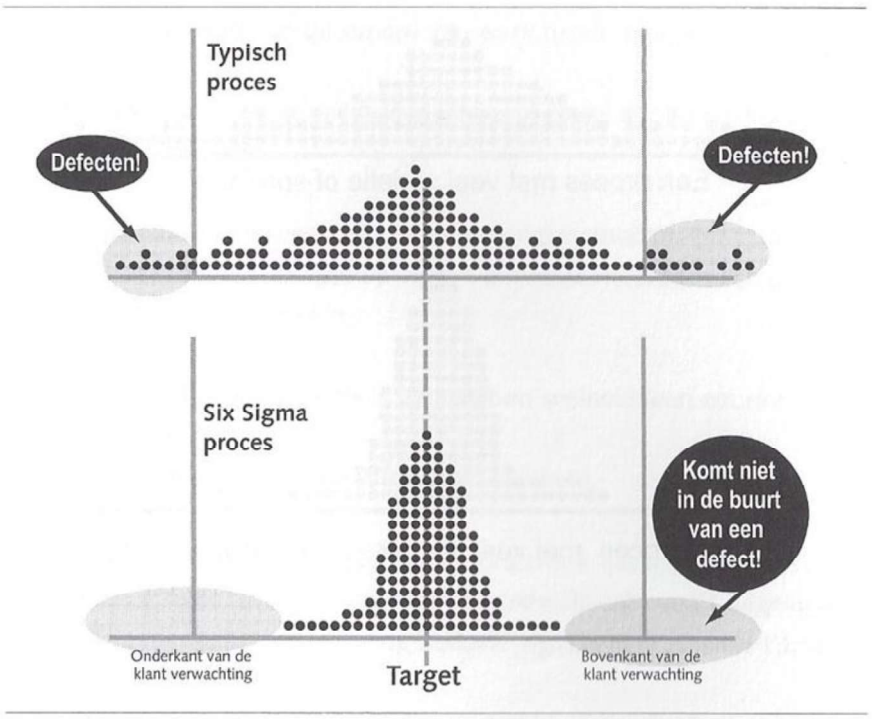


Fuzzy model: shows all possible paths



Activity percentage threshold 15%

Conformance checking: six sigma



Figuur 3.2 Variatie in procesuitkomst

Conformance checking:

- Compare each “instance” (or “case”) of process with “happy path”.
- Violations are discovered.
- Examples:
  - In 47 % of all cases no purchase item is created.
  - If city = “Antwerp” probability of conformity is 3.69 times bigger than average.

Deep dive data mining

- Create key performance indicators for process
- Enables to focus on the priority steps to improve
- AI algorithms can detect the root causes of variation
  - for example, they might point out that every time a new customer needs a credit check, the process is slowed down considerably.

Deep dive data mining

- Works best for processes
  - that have been digitized (i.e. supported by an IT system)
  - where there is still some unstructured work (i.e. reviews and approvals) that happen outside the IT system.
- “Process mining as a method is dependent on case and event data in the logs, and it can accomplish only things allowed by that data.” (CEO QPR)

AI, Machine learning and data mining

- AI = automating tasks we associate with human thinking:
  - Taking decisions
  - Problem solving
  - Learning

without being explicity programmed

- Explicit programming: step-by-step instructions
- AI: computer discovers what to do → learning
- “Let the data talk”

AI = science fiction? NO!

*Still in the early days of its expansion, but **AI's impact on so many different spheres of life is already obvious**. Artificial intelligence is the result of scientific and technological progress in the fields of computing and mathematics, **driven by big data and machine learning**, but also fed by insights in brain research, neuroscience and cognitive psychology. **Quantum computing** will soon dramatically move the frontiers of computer capabilities. And integration with **synthetic biology** is just around the corner.*

Dirk Van Damme, voormalig diensthoofd van het Centre for Educational Research and Innovation van de OESO, 4 februari 2021

What is Machine Learning: definitions

Machine Learning is the science (and art) of programming computers so they can learn from data.

More general:

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.

—Arthur Samuel, 1959

More engineering-oriented:

A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

—Tom Mitchell, 1997

Example: SPAM filter

- = Machine Learning program
- Given:
  - Examples of spam emails (flagged by users)
  - Examples of regular (nonspam = “ham”) emails
- Learns to flag spam
- Training set = example the systems uses to learn

Why use Machine Learning?

- Traditonal approach for SPAM filter
  1. Find common words in spam emails: 4U, credit car, free, amazing, …
  2. Flag all emails that contain (a combination of) these words as spam.
  3. Test and repeat 1+2 until it’s good enough

Why use Machine Learning?

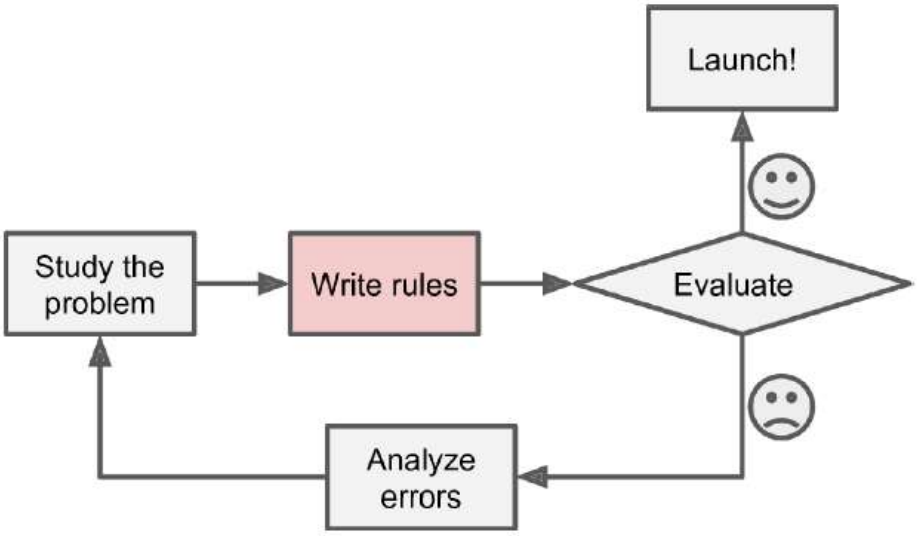


Figure 1-1. The traditional approach

Why use Machine Learning?

- Machine learning approach for SPAM filter
  - automatically learns which words and phrases are good predictors of spam by detecting unusually frequent patterns of words in the spam examples compared to the ham examples
  - The program is much shorter, easier to maintain, and most likely more accurate.

Why use machine learning

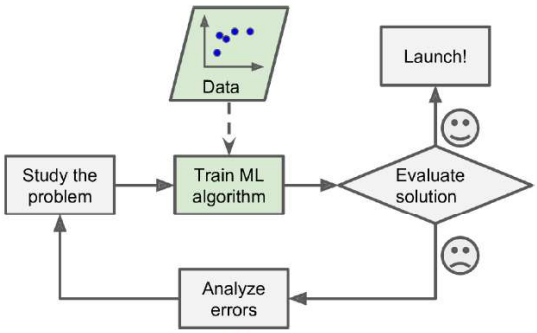


Figure 1-2. The Machine Learning approach

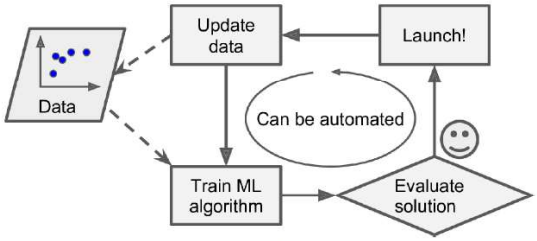


Figure 1-3. Automatically adapting to change

Data mining

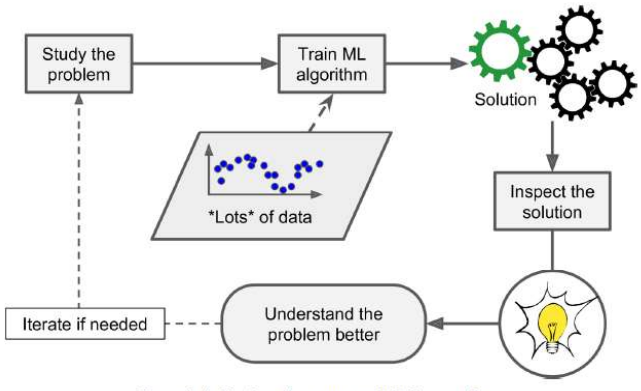


Figure 1-4. Machine Learning can help humans learn

Once a spam filter has been trained on enough spam, it can easily be inspected to reveal the list of words and combinations of words that it believes are the best predictors of spam.

Data mining = applying ML techniques to dig into large amounts of data to help discover patterns that were not immediately apparent.

Examples of Applications

- Analyzing images of products on a production line to automatically classify them
- Detecting tumors in brain scans
- Automatically classifying news articles as sports, financial news, …
- Automatically flagging offensive comments on discussion forums
- Summarizing long documents automatically
- Creating a chatbot or a personal assistant
- Forecasting your company's revenue next year, based on many performance metrics
- Making your app react to voice commands
- Detecting credit card fraud
- Segmenting clients based on their purchases so that you can design a different marketing strategy for each segment
- Representing a complex, high-dimensional dataset in a clear and insightful diagram
- Recommending a product that a client may be interested in, based on past purchases
- Building an intelligent bot for a game

Types of Machine Learning

- Supervised learning
  - Classification
  - Regression
- Unsupervised learning

Supervised Learning: classification

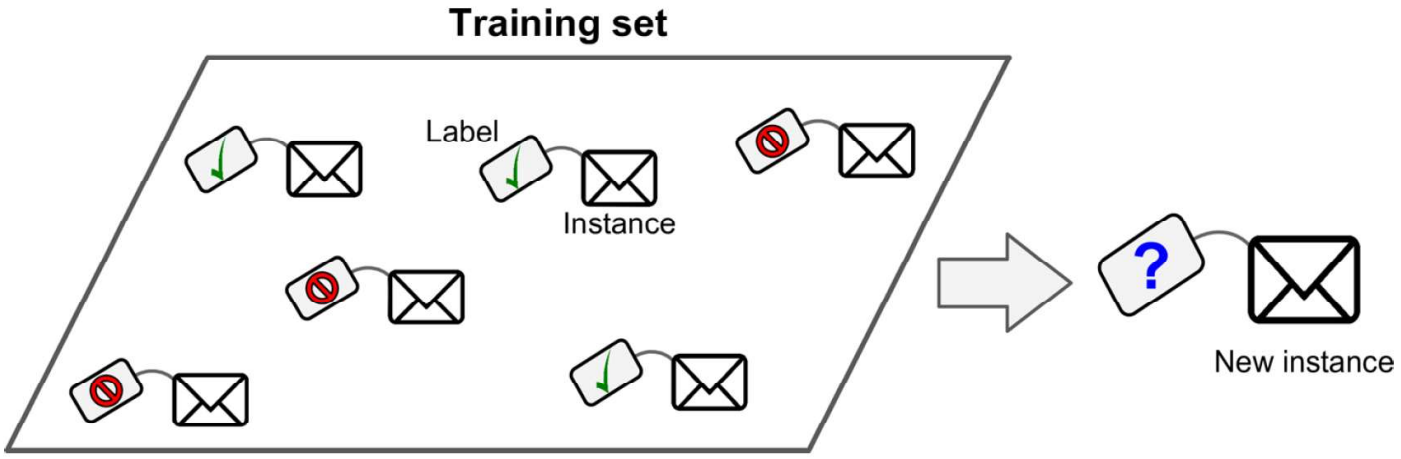


Figure 1-5. A labeled training set for spam classification (an example of supervised learning)

Supervised Learning: regression

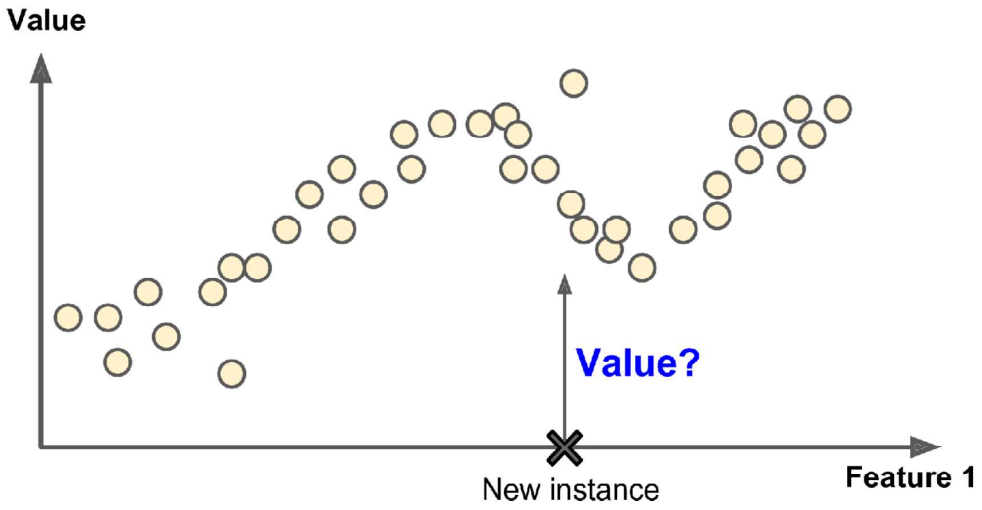


Figure 1-6. A regression problem: predict a value, given an input feature (there are usually multiple input features, and sometimes multiple output values)

Unsupervised learning

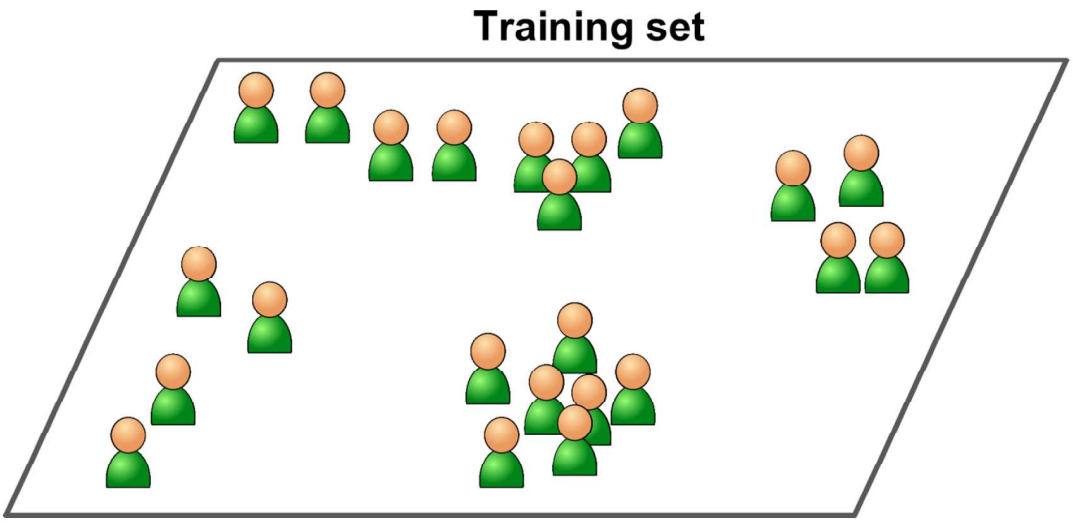
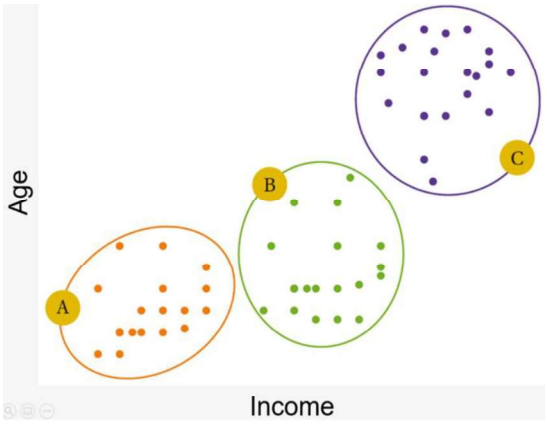


Figure 1-7. An unlabeled training set for unsupervised learning

Unsupervised learning

- Clustering
- Anomaly detection and novelty detection
- Association rule learning (ex. Market Basket Analysis)



Example of clustering

Unsupervised learning: anomaly detection

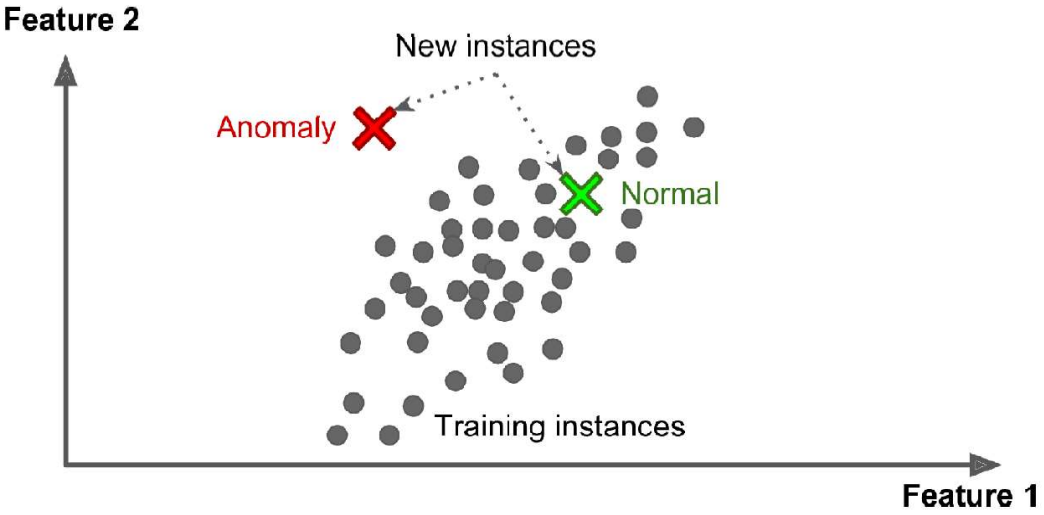


Figure 1-10. Anomaly detection

Data mining and process mining: examples

- Find anomalies (unsupervised learning):
  - Cases or activities taking too much time
  - Too many feedback loops (until it's "ok")
- Predict lead time for a combination of inputs (supervised learning/regression)
- Cluster cases to discover patterns, e.g. purchases from a certain group of countries for a group of products lead to longer lead time (unsupervised learning)
- Predict wether a case (based on certain inputs) will fail or not (supervised learning / classification)
- Find root cause of process failures (using "whith box" algorithms, explainable AI)

Exercise

- Which data for your selected process do you already have?
- Which extra data could be easily captured?
- What possible applications of process mining do you see?