

Machine Learning, Data Analytics und Co. Was ist das eigentlich ...

... und viel wichtiger: Was kann man damit anfangen?

- Lukas Hahn
- Assekuranzforum Lebensversicherung 1/2018
- Berlin, den 24. April 2018



Machine Learning, Data Analytics und Co.

Agenda

Was ist das eigentlich?

Wie geht das ungefähr?

Und viel wichtiger: Was kann man damit anfangen?

Fazit

Institut für Finanz- und Aktuarwissenschaften

Machine Learning, Data Analytics und Co.

Agenda

Was ist das eigentlich?

Machine Learning: (eine) Definition

Ein spielerisches Beispiel

Einordnung in der Statistik

Data Analytics und Co.

Buzzword-Domino

Wie geht das ungefähr?

Und viel wichtiger: Was kann man damit anfangen?

Fazit

Institut für Finanz- und Aktuarwissenschaften

Was ist das eigentlich?

Machine Learning: (eine) Definition

Machine Learning



Field of study that gives **computers**

the ability to **learn**

without being explicitly programmed.

Bildquelle:
Lexalytics, Inc.
[https://www.lexalytics.com/
technology/machine-learning](https://www.lexalytics.com/technology/machine-learning)

Arthur Samuel (1959), Pionier des Machine Learnings
Some Studies in Machine Learning Using the Game of Checkers
IBM Journal of Research and Development, 44(1.2)

Was ist das eigentlich?

Machine Learning: (eine) Definition

Was bedeutet **Lernen**?

Learning is any process by which a system improves performance from experience.

Herbert A. Simon, Wirtschaftsnobelpreisträger 1978

A **computer program** is said to **learn** from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T , as measured by P , improves with experience E .

Tom M. Mitchell (1997), Buchautor des Standardwerks „Machine Learning“, McGraw Hill

Voraussetzungen:

- klar spezifizierte Aufgabenstellung → Jochen soll für Alex tolle Socken stricken.
 - Gütekriterium zur Messung der Zielerreichung → Passen Alex die Socken? Geht das schnell?
 - Daten zum Lösen und Bewerten der Aufgabe → diverse Nadeln und Wolle (und Alex' Füße)
- Machine Learning (ML) bedeutet also nicht die klassische Konstruktion von Algorithmen, die durch statische und regelbasierte Instruktionen die Aufgabe lösen, sondern die **Konstruktion von Algorithmen**, die durch Daten so **einen Algorithmus selbst konstruieren können**.



- Charakteristische Eigenschaft** von Machine-Learning-Algorithmen ist, dass die
- Problemlösung mit mehr oder neuen Daten **immer besser wird** und
 - implizit konstruierte **besser als** eine direkt programmierte Lösung werden kann.

Was ist das eigentlich?

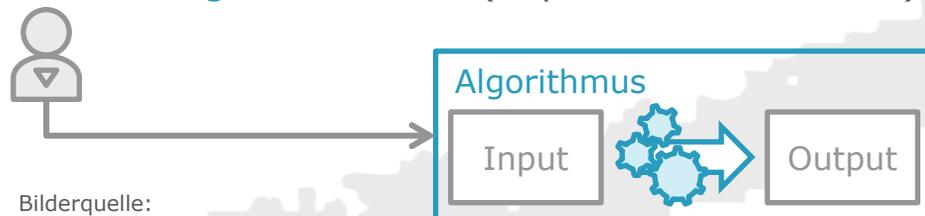
Ein spielerisches Beispiel

Deep Blue



intelligent
aber doch kein ML

- **Schachcomputer** von IBM
- Sieg über Weltmeister Kasparov im Jahr 1997
- Rezept: Algorithmus zur Problemlösung
 - Massives Brute Force
 - **Designte** Heuristik (Experten und Historie)



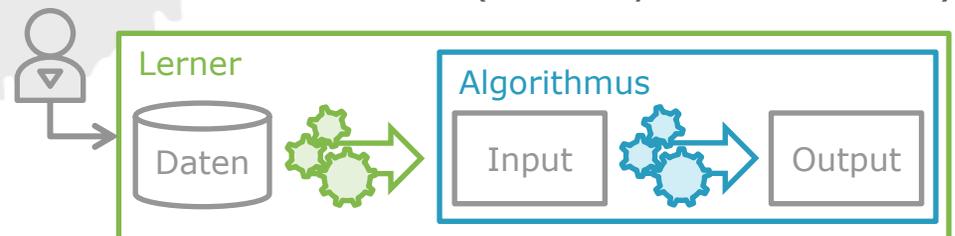
Bilderquelle:
Pixabay

AlphaGo



intelligent
Machine Learning!

- **Go-Computer** von Google DeepMind
- Sieg über „9-dan“ Lee Sedol im Jahr 2016
- Rezept: Erlerner eines solchen Algorithmus
 - Umfangreiches Brute Force
 - **Erlernte** Heuristik (Self-Play-Reinforcement)



Was ist das eigentlich?

Einordnung in der Statistik

Machine Learning bedeutet somit auch:

Gesetzmäßigkeiten in Daten identifizieren

Induktive Statistik

- Hypothesen vorhanden, z.B. Expertenwissen
- Bsp.: Wir testen die Hypothese „Je kürzer die Vertragslaufzeit, desto höher die Stornoquote.“
- gewisse Muster sind a priori bekannt
- Ziel: Induktive Statistik anwenden, um in den Daten vorhandene **Hypothesen zu verifizieren/falsifizieren.**

Top Down

Schließen

Explorative Statistik

- Hypothesen sind **nicht** erforderlich.
- Bsp.: Wir schauen in unserem Bestand, wer typischerweise storniert.
- Muster sind a priori **nicht** bekannt.
- Ziel: Explorative Statistik anwenden, um in den Daten **Muster (d.h. Strukturen und Zusammenhänge) zu erkennen**

Bottom Up

Suchen



- Die im Machine Learning zur Anwendung kommenden Methoden überschneiden sich häufig mit **komplexen Verfahren der explorativen Statistik.**
- Als Ergebnis werden i.d.R. Zusammenhänge sichtbar, die sich nicht intuitiv erklären lassen.

Was ist das eigentlich?

Einordnung in der Statistik

Annahmen für die ML-Methoden sind oft nur **implizit** formuliert.

- Robustheit über die Zeit
 - Annahme bei einmaliger Modellierung: Statistische Zusammenhänge sind zeitlich stabil.
 - Strukturbrüche in den Vergangenheitsdaten müssen bereinigt werden.
 - **Beispiel: Änderung der Stornohaftungsdauer in der Vergangenheit bei Stornovorhersage**
 - Für die Zukunft werden keine Trends erwartet oder müssen separat berücksichtigt werden.
 - **Beispiel: (medizinische) Inflation bei der Vorhersage von Leistungshöhen**
 - Oder der ML-Algorithmus wird implizit befähigt auch Trends zu erlernen.
 - Das muss ganz bewusst in den Daten aufbereitet werden!
 - **Beispiel: Modellierung der historischen Inflation zur Prognose der zukünftigen Inflation**
- Unabhängigkeit der Erfahrungen, denn jede Erfahrung ist bei der Optimierung gleich viel wert.
- Hinreichend gute Datenqualität, denn aus schlechten Daten kann die Maschine nichts (Gutes) lernen.



Vor allem aber fehlen im **Gegensatz zur klassischen Statistik** typischerweise

- Annahmen zum datengenerierenden Prozess (Verteilungsannahmen),
- Vorgabe eines funktionalen Zusammenhangs in den Daten (z.B. Linearität).

Was ist das eigentlich?

Einordnung in der Statistik

Data Mining



- Bestmögliche Identifikation komplexer **Muster**
 - Ziel: Wissensgenerierung, z.B. Clustering
 - Aufgabe: Identifizierung systematischer Zusammenhänge in speziellem Datensatz
 - Vorgehen: oftmals einmalige Analyse unter Verwendung von ML-Methoden



Bilderquelle:
Pixabay

Predictive Modelling



- Bestmögliche individuelle **Vorhersage**
 - Ziel: optimale Entscheidungsfindung
 - Aufgabe: Identifizierung systematischer Vorhersageregeln für neue Daten
 - Vorgehen: oftmals manuelle, idealerweise tatsächlich selbstlernende ML-Anwendung



Was ist das eigentlich?

Data Analytics und Co.

Data Analytics

... is the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to **drive decisions and actions**.

Davenport, Thomas and, Harris, Jeanne (2007). Competing on Analytics. O'Reilly.

... is the discovery, interpretation, and communication of **meaningful patterns** in data.

Englischer Wikipedia-Eintrag zu „Analytics“, Stand 17.04.2018

- Prozess zur Einbettung der technischen Datenanalyse in kontextbasierte **Auswertung, Interpretation und Kommunikation** der gewonnenen Erkenntnisse bis hin zur abgeleiteten **Entscheidungsfindung**.
- Neben klassischen Ansätzen der Datenanalyse kommen dabei vermehrt **Methoden des Machine Learnings** zur Anwendung (**Advanced Analytics**).
- Durch die Überführung in eine verstärkt datengetriebene (ggf. nicht mehr intuitive) Entscheidungsfindung wird Data Analytics zusehends als ein Teil von **Data Science** verstanden, dem interdisziplinären Feld aus Daten Wissen zu extrahieren.

Bildquelle:
Pixabay

Was ist das eigentlich?

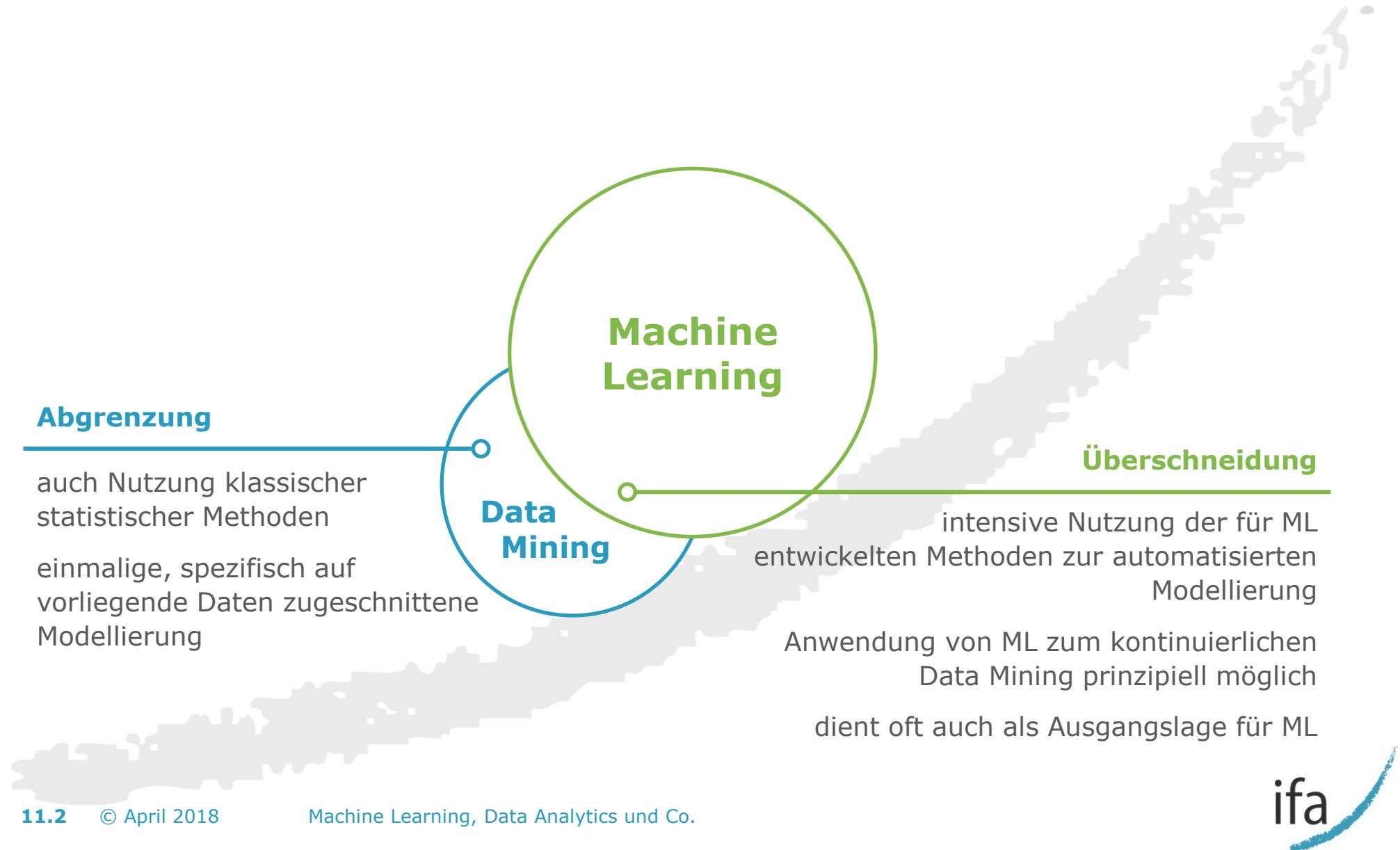
Buzzword-Domino



**Machine
Learning**

Was ist das eigentlich?

Buzzword-Domino



Was ist das eigentlich?

Buzzword-Domino

Machine Learning

Überschneidung

ML behandelt i.d.R. Vorhersagen und Entscheidungsfindungen, wird daher oft synonym verwendet.

teils auch bei Data Mining eingeordnet als spezielle Art der Mustererkennung

Data Mining

Predictive Modelling

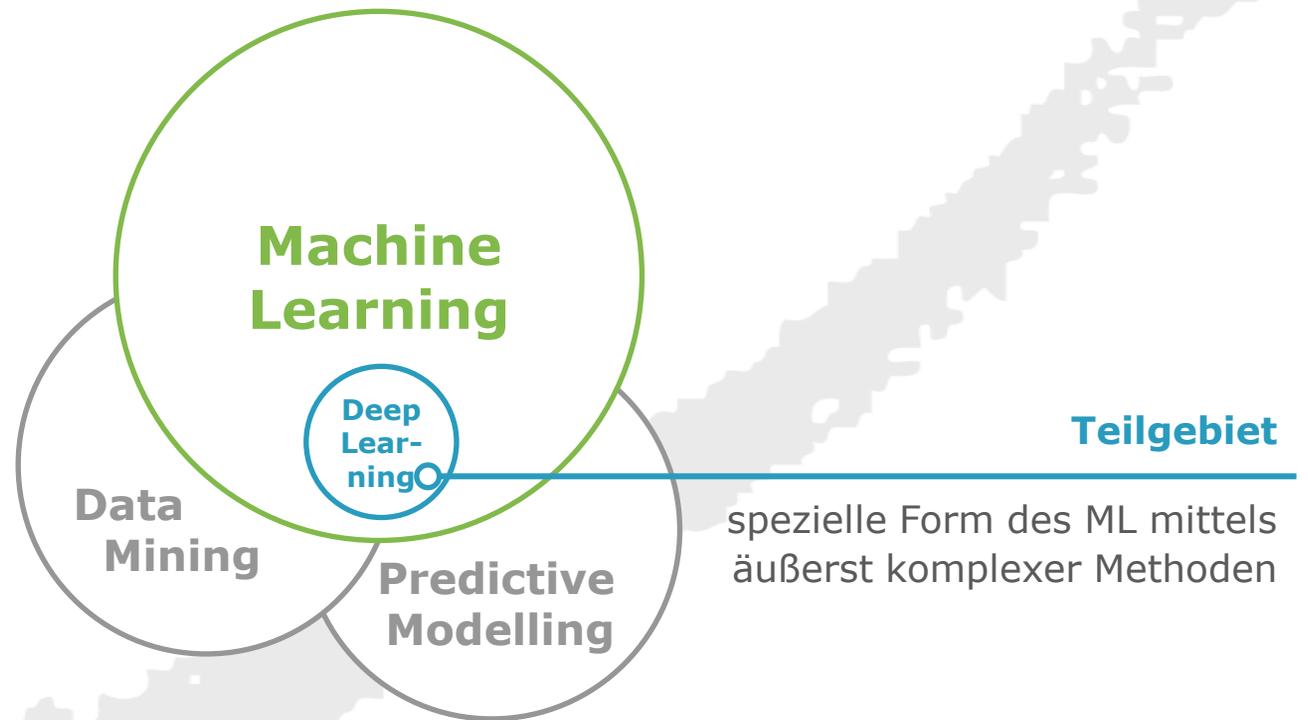
Abgrenzung

Nutzung klassischer statistischer Methoden

oftmals einmalige Anwendung von ML-Methoden für statisches Modell

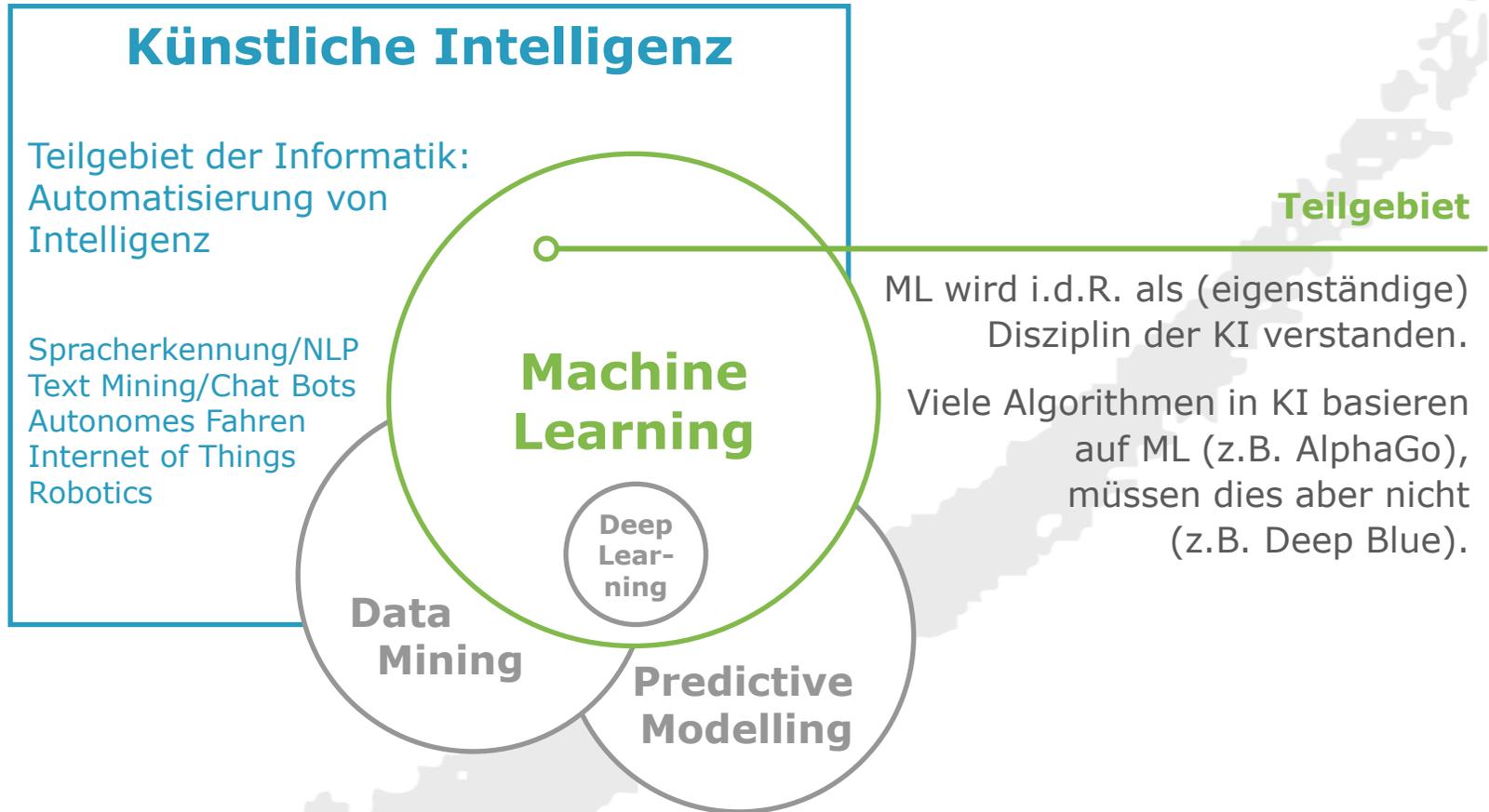
Was ist das eigentlich?

Buzzword-Domino



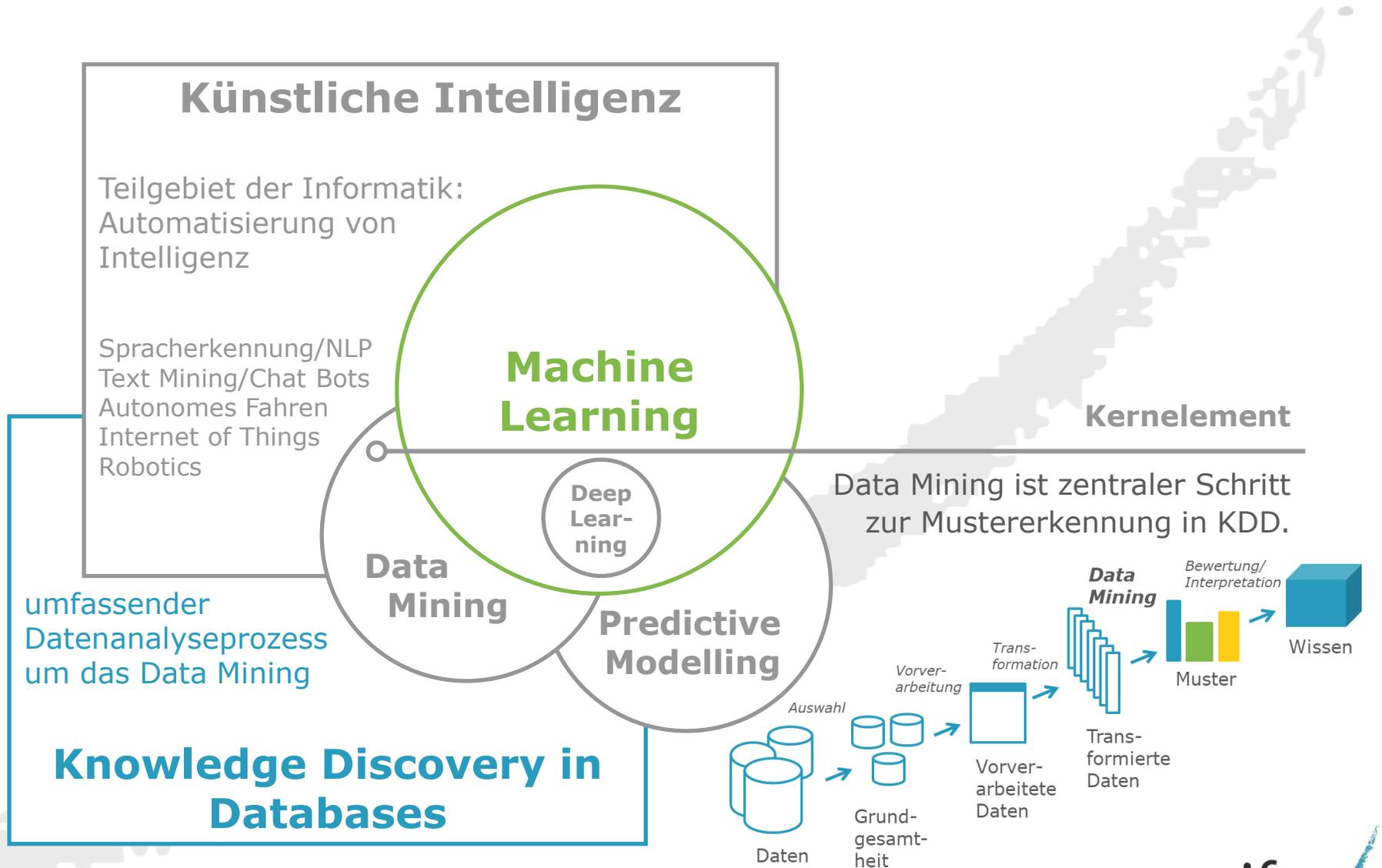
Was ist das eigentlich?

Buzzword-Domino



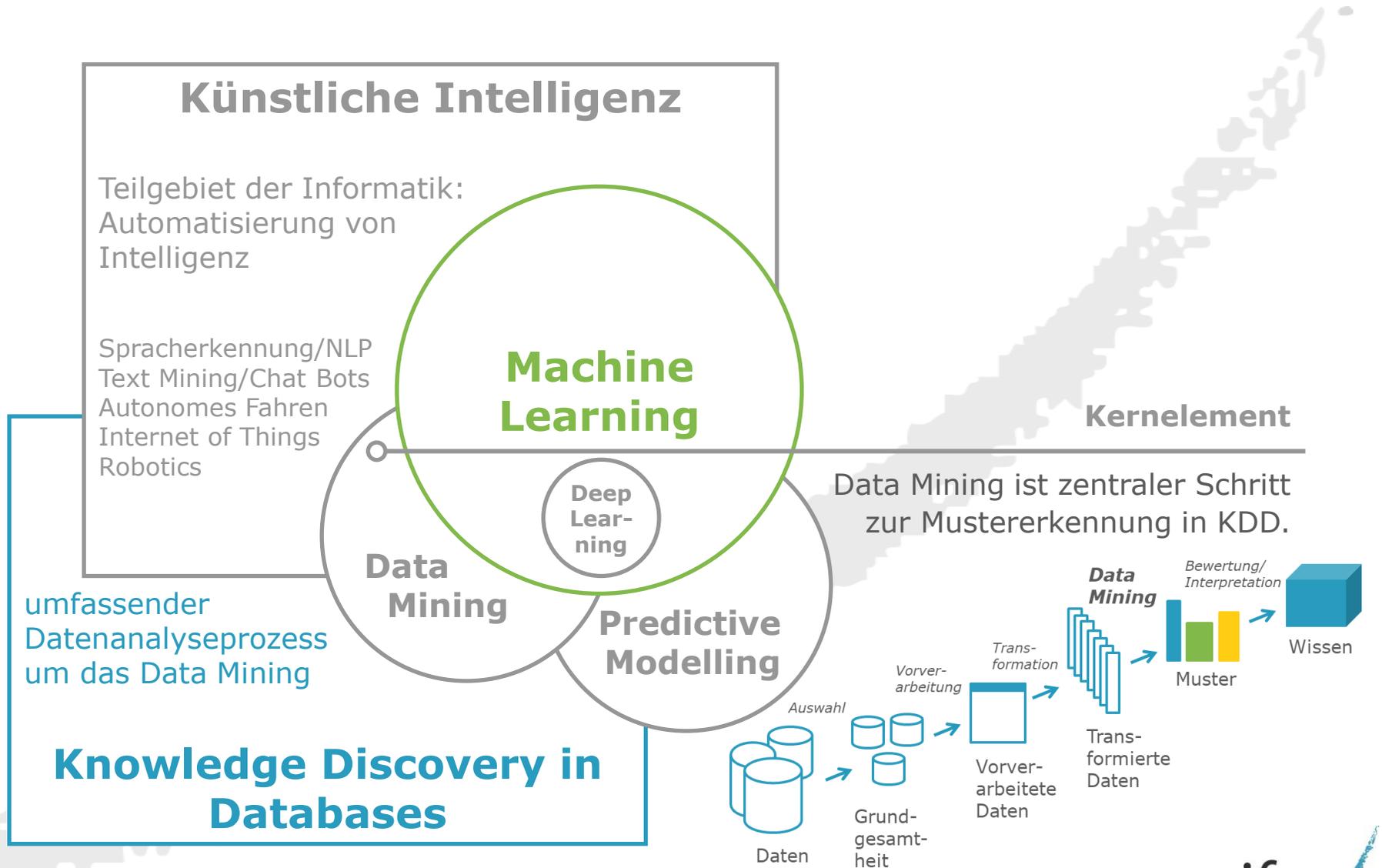
Was ist das eigentlich?

Buzzword-Domino



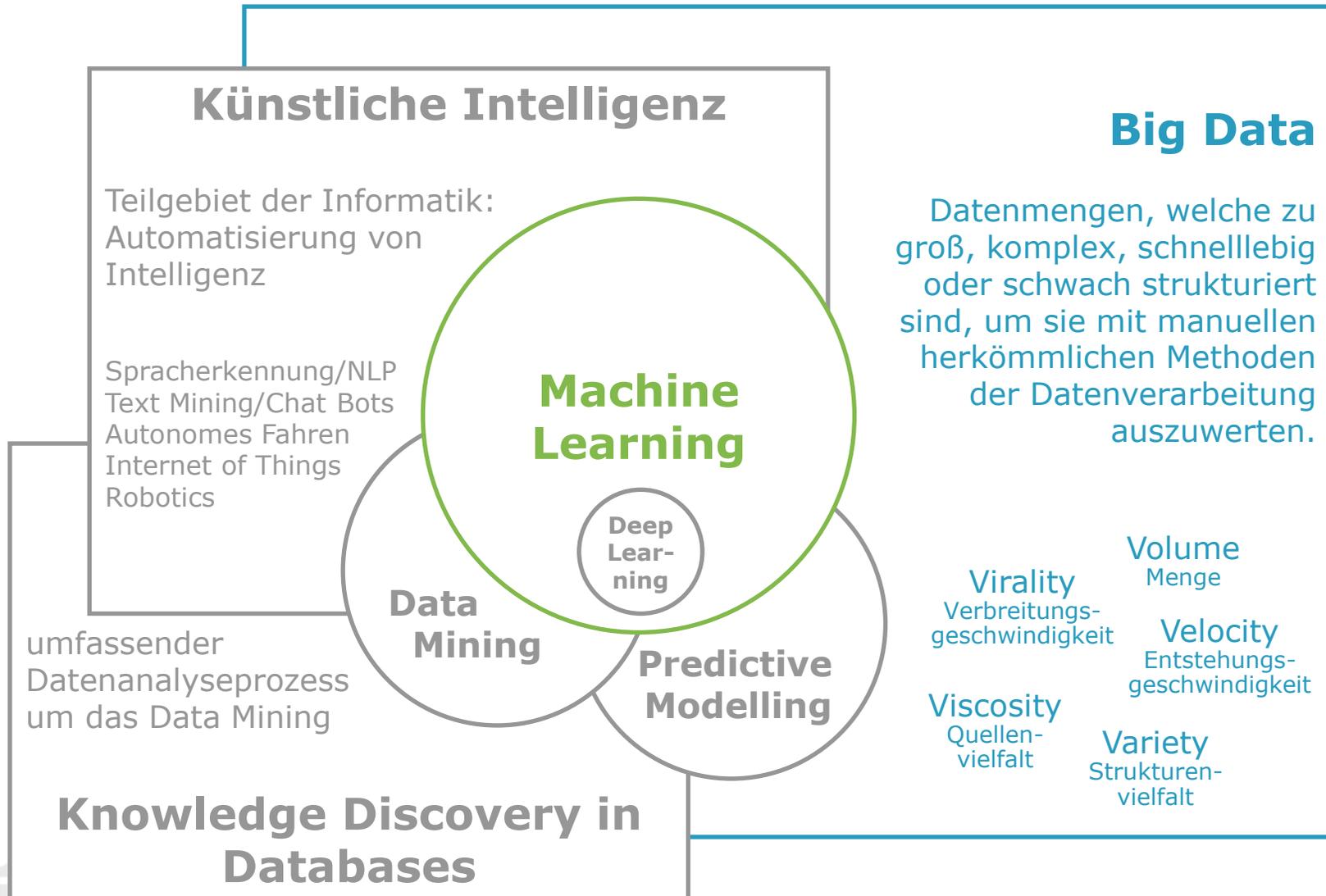
Was ist das eigentlich?

Buzzword-Domino



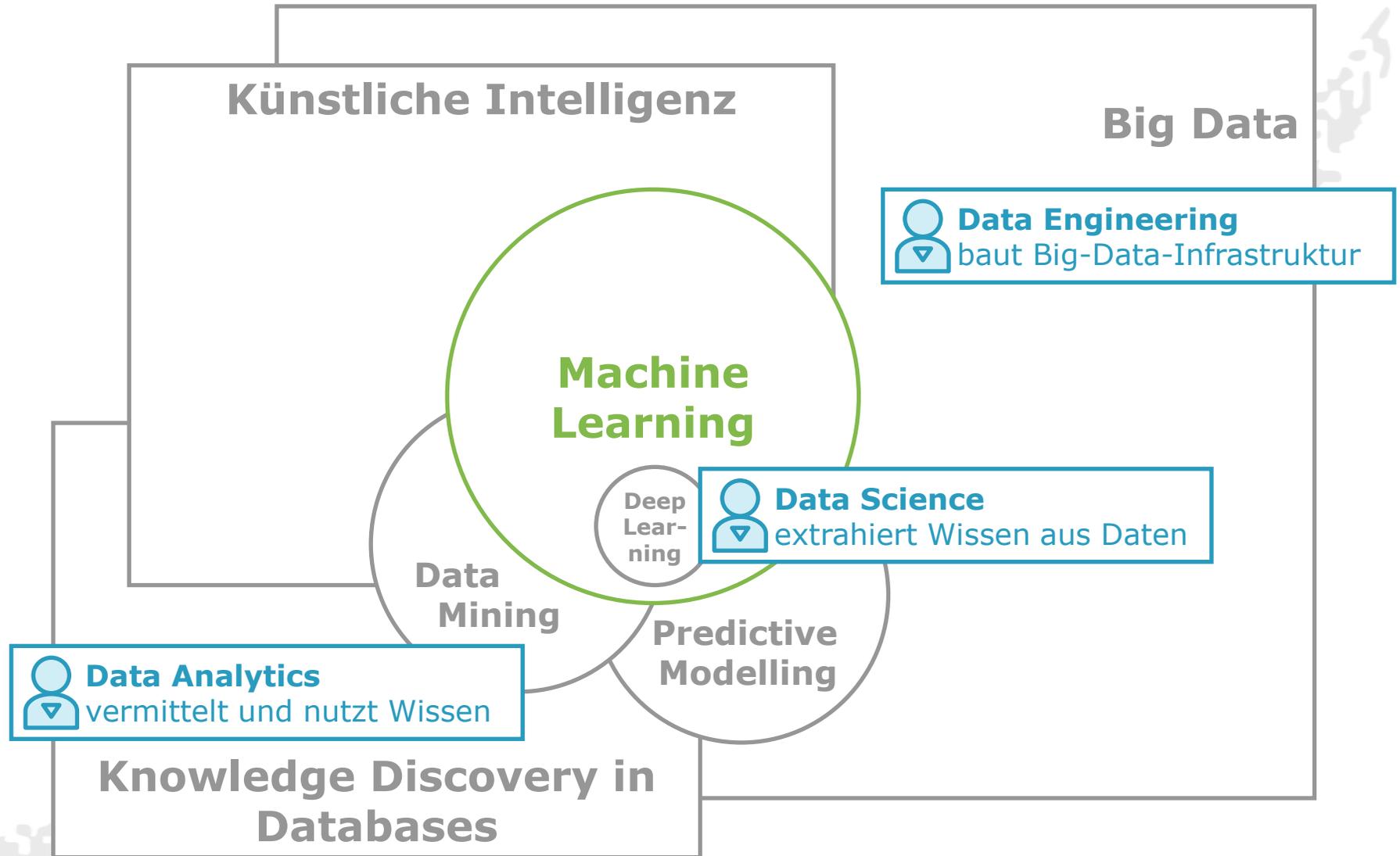
Was ist das eigentlich?

Buzzword-Domino



Was ist das eigentlich?

Buzzword-Domino



Machine Learning, Data Analytics und Co.

Agenda

Was ist das eigentlich?

Wie geht das ungefähr?

Ein Fallbeispiel aus der Schadenregulierung

Spezifikation der Zielsetzung

Modellierung

Beispiel: Klassifikationsbaum

Deployment

Und viel wichtiger: Was kann man damit anfangen?

Fazit

Institut für Finanz- und Aktuarwissenschaften

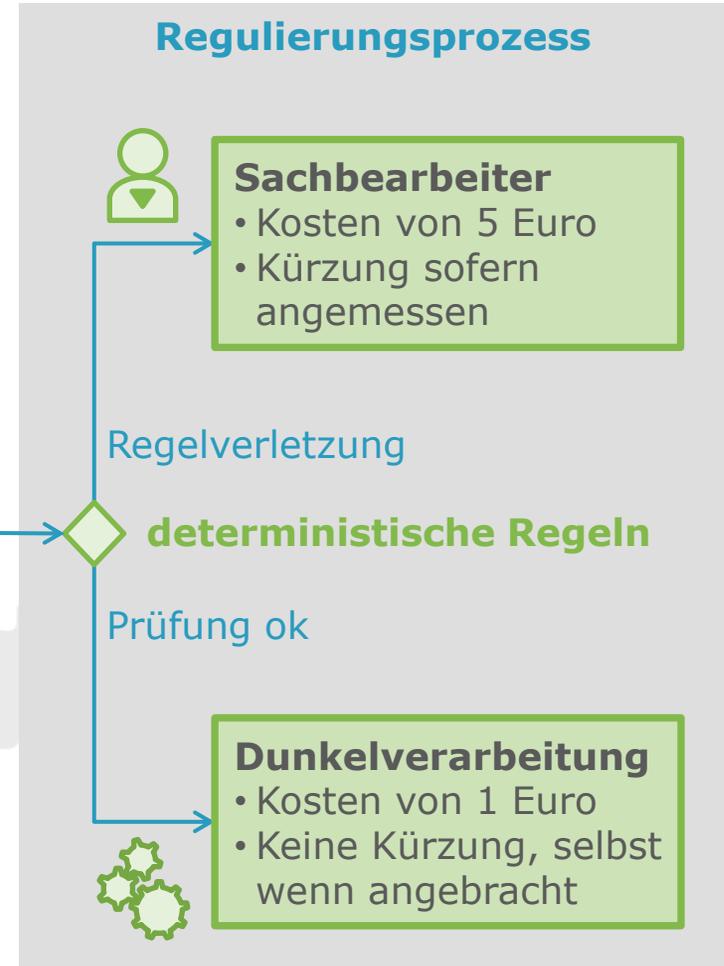
Wie geht das ungefähr?

Ein Fallbeispiel aus der Schadenregulierung

illustrativ

Ausgangslage

- Aussteuerung von Belegen zur vollautomatisierten Regulierung mit deterministischen Regeln
 - Beispiel: „keine Pille für den Mann“
- Annahme von jährlich 1 Mio. Arzneimittelbelegen
 - 15% der Belege sind zu kürzen
 - im Schnitt um 30 Euro
 - 50% der Belege dunkel verarbeitet
 - alle ohne Kürzung
 - Kosten pro Beleg
 - Sachbearbeiter: 5 Euro
 - Dunkelverarbeitung: 1 Euro
- Saldo der Leistungsprüfung von **1,5 Mio. Euro** aus
 - Bearbeitungskosten: 3,0 Mio. Euro
 - Einsparung durch Kürzung: 4,5 Mio. Euro



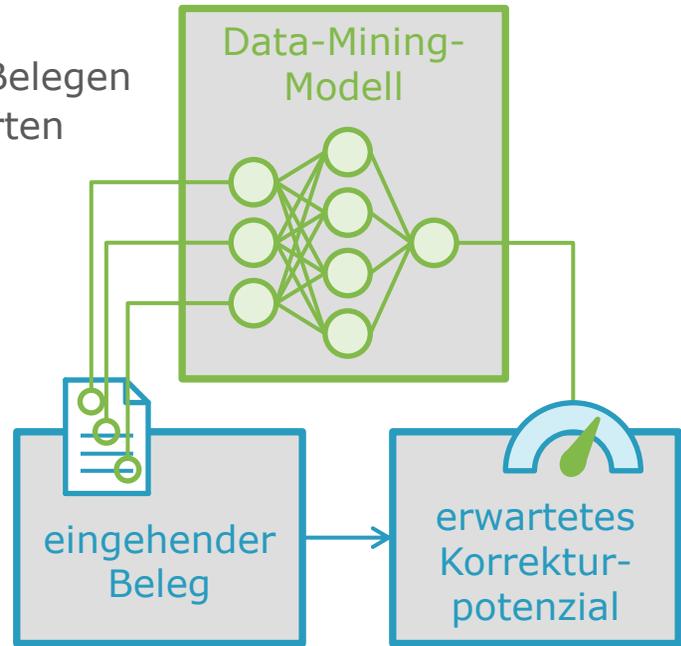
Wie geht das ungefähr?

Ein Fallbeispiel aus der Schadenregulierung

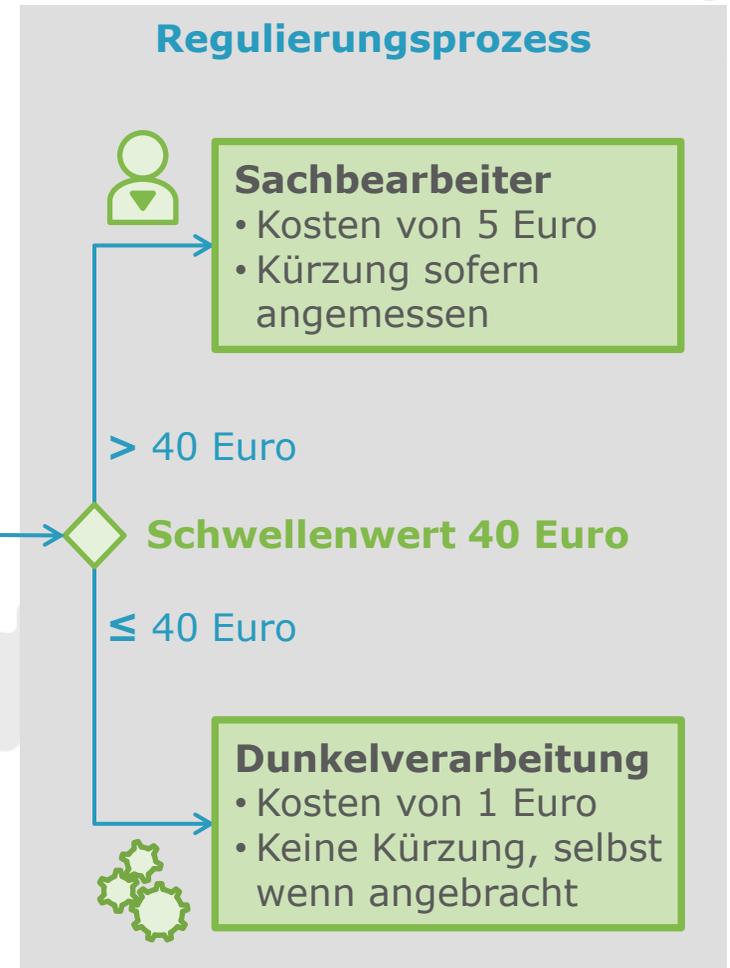
illustrativ

Ökonomisches Ziel

- Aussteuerung von Belegen zur vollautomatisierten Regulierung bei geringem Korrekturpotenzial



- belegindividuelle Vorhersage durch ML-Methode
- Zielgröße: erwartetes Korrekturpotenzial des Belegs anhand vergangener Belegs- und Personendaten
- Performance: Fehlsteuerung unter Berücksichtigung des Tradeoffs von Kosten und entgangener Kürzung



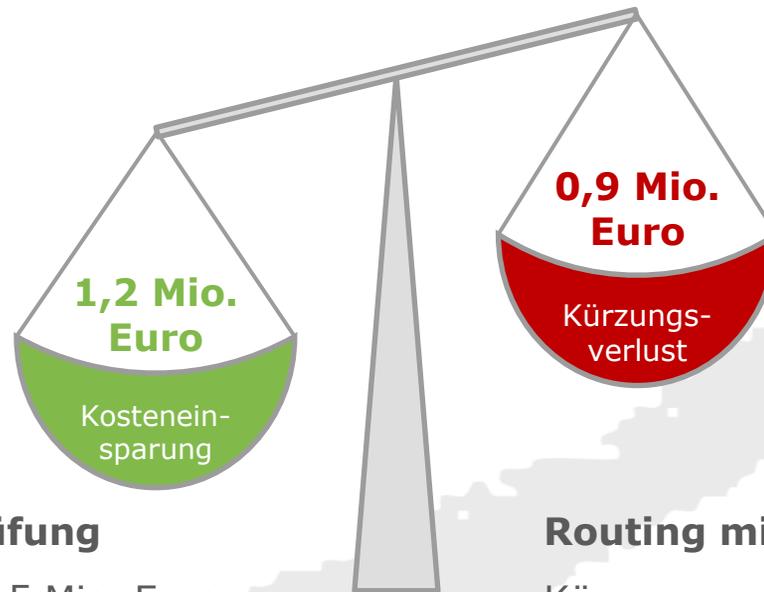
Wie geht das ungefähr?

Ein Fallbeispiel aus der Schadenregulierung

illustrativ

Insgesamt steigt die Profitabilität

- Annahme: 80% statt 50% Belege ausgesteuert
 - darunter aber auch 1/5 aller Kürzungsbelege



Regelbasierte Belegprüfung

Kürzungen: 4,5 Mio. Euro
Bearbeitungskosten: 3,0 Mio. Euro

Saldo: 1,5 Mio. Euro



Routing mit Machine Learning

Kürzungen: 3,6 Mio. Euro
Bearbeitungskosten: 1,8 Mio. Euro

Saldo: 1,8 Mio. Euro

Wie geht das ungefähr?

Spezifikation der Zielsetzung

Aufgabenstellung

Performancemessung

Datenanforderung

Ausgangsfrage: Was soll der erlernte Algorithmus tun können?

Zunächst muss das **übergeordnete Ziel** klar sein:

- Was ist die zugrundeliegende Motivation?
- Welche Maßnahme soll ergriffen werden?
- Was erhofft man sich davon konkret?
- Worauf wird sich dabei fokussiert?
- Was ist Aufgabe des zu lernenden Algorithmus?
- Was ist nicht Teil des zu lernenden Algorithmus?

Beispiel Hochkostenfälle

übergeordnetes Ziel: Leistungsausgaben in der PKV durch Prävention steuern

- Motivation: Kostensenkung, Kundenservice
- konkrete Maßnahme: (medizinische) Prävention für **zukünftige** Hochkostenfälle
- Ziel: Hochkostenfälle verhindern oder abmildern
- Fokus: „Volkskrankheiten“, z.B. Diabetes
- Aufgabe: Hochkostenfälle **rechtzeitig** vorhersagen
- aber nicht: Präventionsmaßnahme vorschlagen oder durchführen



- Die Zielsetzung muss **genauestens konkretisiert** werden!
- Ohne eindeutige Klärung dieser Fragen lässt sich für die Maschine keine eindeutige Aufgabe ableiten, die **zielgerichtet** gelöst werden soll.

Wie geht das ungefähr?

Spezifikation der Zielsetzung

Aufgabenstellung

Performancemessung

Datenanforderung

Ausgangsfrage: Was soll der erlernte Algorithmus tun können?

Aus der konkreten Zielsetzung wird ein **Machine-Learning-Task** abgeleitet:

- z.B. Vorhersage eines Ereignis (Storno, Kapitalwahl) oder eines Werts (Kürzungsvolumen)
 - Geht es dabei um das **Ob** oder **Wann**?
 - Für **welchen** Zeitraum wird prognostiziert?
 - Soll dies einmalig oder laufend geschehen?
- oder aber Clustering (Kunden, Vermittler), Identifikation von Anomalitäten (Betrug), etc.
 - mit jeweils analogen Detailfragen

Beispiel Hochkostenfälle

Machine-Learning-Task: Vorhersage von Hochkostenfällen

- Was ist ein Hochkostenfall: ausgewählte Diagnose(n) oder eine Mindesthöhe für Leistungsausgaben?
- Welche Manifestierung interessiert uns: irgendwann oder zeitnah? Welche Vorlaufzeit ist für die Einleitung von präventiven Maßnahmen notwendig?
- Wollen wir Vorhersagen zu festgelegten Stichtagen oder Dashboard (laufend mit neuen Informationen)?



- Die Zielsetzung muss in eine **konkrete Aufgabe** für die Maschine umformuliert werden!
- Die Maschine kann eine Aufgabe umso **besser** lösen, je **eindeutiger** die Aufgabenstellung (z.B. Definition der vorherzusagenden Zielgröße) ist.

Wie geht das ungefähr?

Spezifikation der Zielsetzung

Aufgabenstellung

Performancemessung

Datenanforderung

Ausgangsfrage: Wie soll die Güte des Algorithmus bewertet werden?

Die Bewertung hängt von der Aufgabenstellung ab:

- Beispiel Vorhersage: Wie bewerten wir den Fehler einer falschen Vorhersage?
 - Wie **wirken** sich Fehler auf die angestrebte Maßnahme aus? Was werden sie uns **kosten**?
 - Sind unterschiedliche Fehler **gleich schlimm** oder **unterschiedlich stark** zu bewerten?

Anschließend **Ableitung eines Performancemaß**:

- im Kern ein zur (methodischen) Aufgabenstellung passendes statistisches Gütemaß
- angereichert um eine zur (fachlichen) Aufgabenstellung geeignete ökonomische Bewertung



- Die Konkretisierung der Performancemessung im Einklang mit der Aufgabenstellung ist von **größter Bedeutung**!
- Bei einem schlechten Lehrer (falsch spezifizierte Aufgabenstellung, ungeeignete Bewertung) wird die Maschine auch nicht richtig lernen.

Beispiel Hochkostenfälle

zwei Fehlerarten möglich

- Hochkostenfall **nicht erkannt**: keine Prävention eingeleitet, d.h. Kostensenkung verpasst
 - Hochkostenfall **fälschlicherweise vorhergesagt**: Prävention eingeleitet obwohl nicht notwendig, d.h. zusätzliche Kosten
- **Tradeoff** muss durch Kosten-Nutzen-Analyse bestimmt werden.

Wie geht das ungefähr?

Spezifikation der Zielsetzung

Aufgabenstellung

Performancemessung

Datenanforderung

Ausgangsfrage: Mit welchen Erfahrungen soll die Maschine lernen die Aufgabe gut zu lösen?

Aus der konkreten Aufgabenstellung ergibt sich eine **Anforderung an die notwendigen Daten**:

- Ist die **Zielgröße** in unseren Daten enthalten?
 - Wie kann sie sonst abgeleitet werden?
- **Welche Daten** sind hinzuzuspielen, sodass die Maschine einen Algorithmus erlernen kann?
 - Gibt es vorab Erkenntnisse oder Hypothesen?
- Liegen **genügend Daten** vor, damit statistische Zusammenhänge erkennbar sind?
- Soll mit einer **vorläufigen Auswahl** an Daten ein erstes Teilergebnis erzielt werden?

Beispiel Hochkostenfälle

Mustererkennung mittels Vergangenheitsdaten

- Bestand zu vergangenem Stichtag **ableiten**
- Definition des Hochkostenfalls **hinzuspielen** (z.B. Diagnosen aus eingereichten Belegen auslesen)
- zum Stichtag bekannte Hochkostenfälle **entfernen**, zukünftige (nach Präventionszeitraum) **markieren**
- zum Stichtag bekannte Daten zur Person, zur Leistungshistorie etc. **anreichern** und dabei
- medizinisches Wissen **berücksichtigen** (z.B. Hochkostenfall auch dann, wenn Insulin ohne Diagnose)



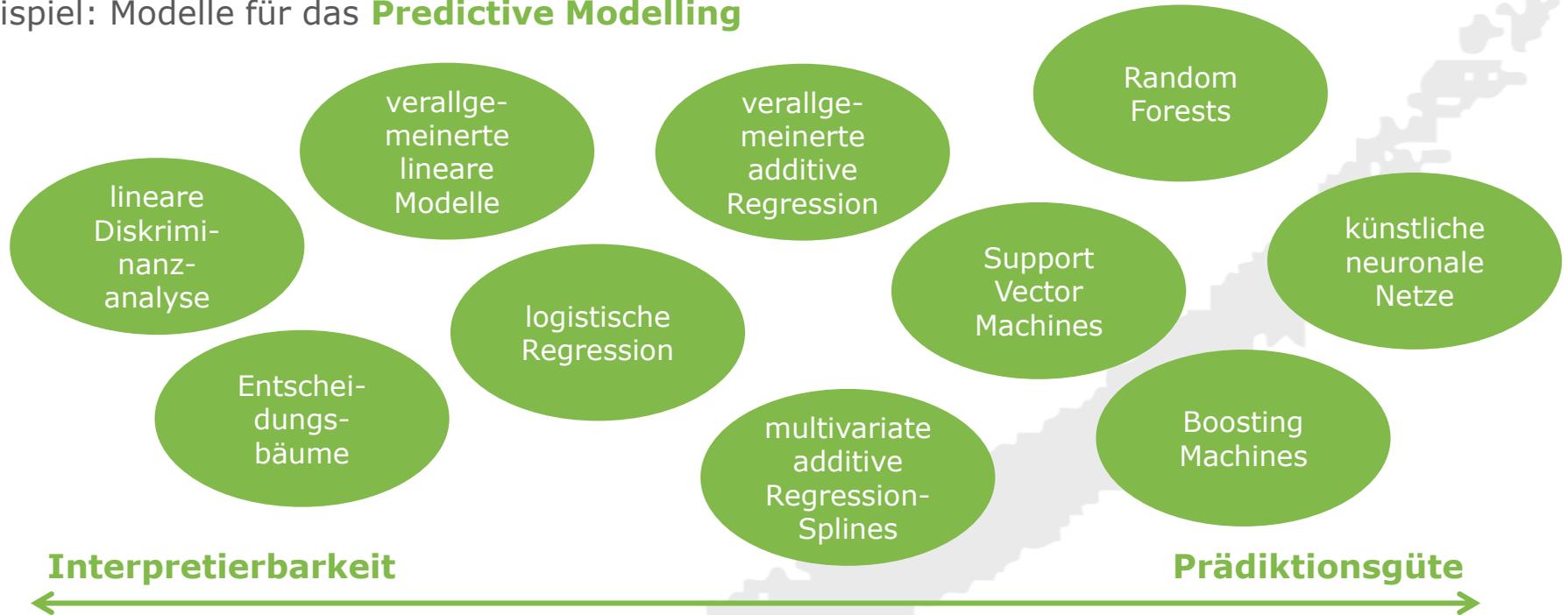
- Die Maschine kann nur Muster identifizieren, die in den Daten vorhanden sind.
- Fehlende Daten bedeutet auch verpasstes Potenzial.
- Zu viele (unnötige) Daten erschweren den Lernprozess.

Wie geht das ungefähr?

Modellierung

Für den ML-Algorithmus stehen je nach Aufgabenstellung verschiedene **Modellklassen** zur Verfügung.

■ Beispiel: Modelle für das **Predictive Modelling**

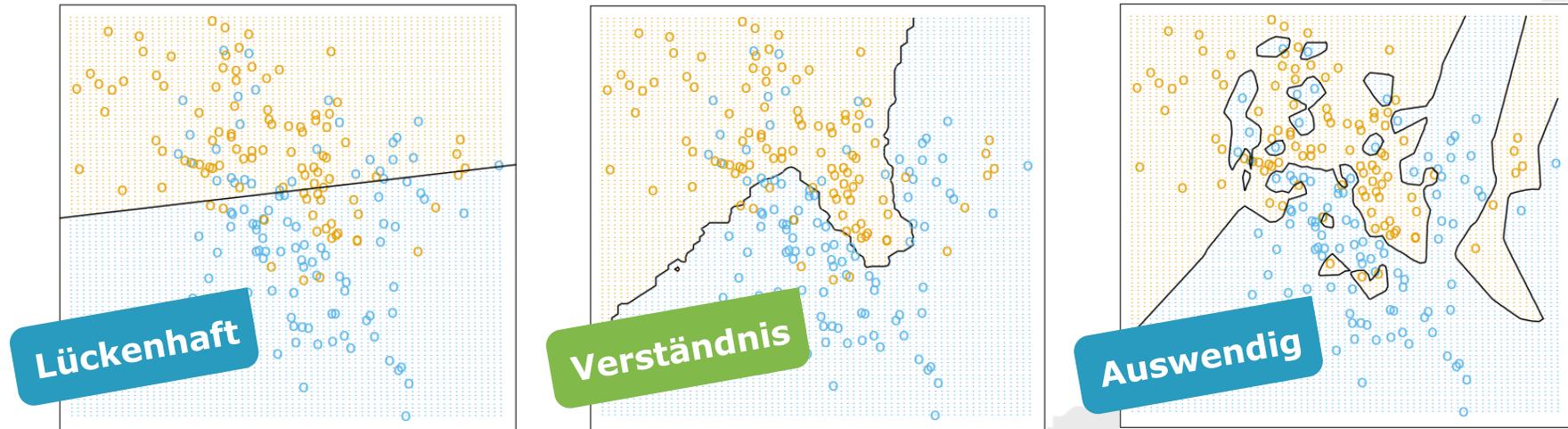


Es steht eine **Vielzahl von Machine-Learning-Methoden** zur Verfügung. Die Entscheidung für bestimmte Verfahren sollte mit Blick auf die Zielsetzung und erst nach Analyse der Daten getroffen werden. **Gemeinsames Verständnis** der **statistischen Methoden** und des **fachlichen Hintergrunds** ist unabdingbar!

Wie geht das ungefähr?

Modellierung

Herausforderung: Was ist ein guter Lernprozess?



Bilderquelle: Hastie et al. (2009). The Elements of Statistical Learning – Data Mining, Inference, and Prediction



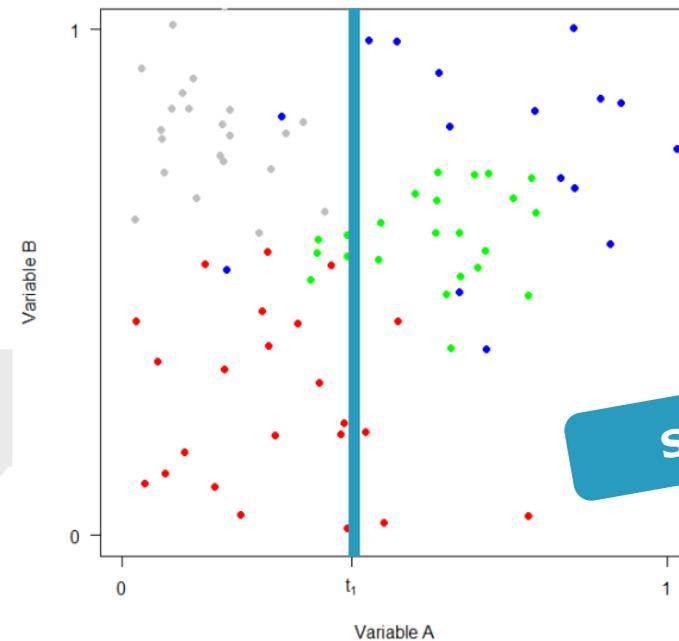
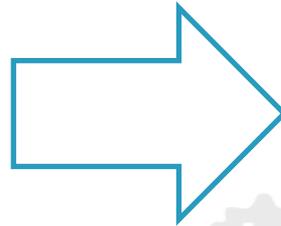
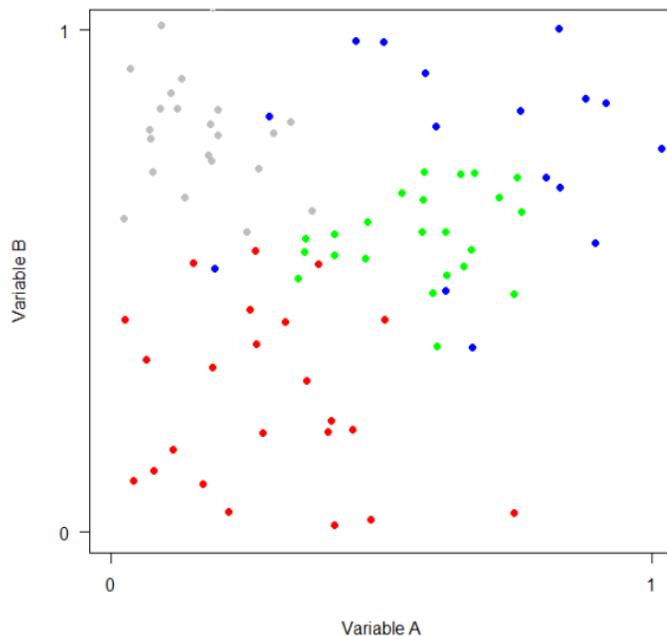
Der Machine-Learning-Algorithmus muss die Maschine befähigen, das gewählte Modell derart zu nutzen, dass es die gelernte Erfahrung **bestmöglich auf neue Situationen anwenden** kann. Diese Fähigkeit muss von uns validiert werden.

Wie geht das ungefähr?

Beispiel: Klassifikationsbaum

- Zwei Merkmale A und B, z. B. Blutwerte
- Zielgröße mit vier Kategorien (grau, rot, grün, blau), z. B. medizinische Risikoklasse
- **Aufgabe:** Risiko anhand von A und B klassifizieren
- **Performance:** möglichst geringe Fehlklassifikation
- **Lerner:** Klassifikationsbaum

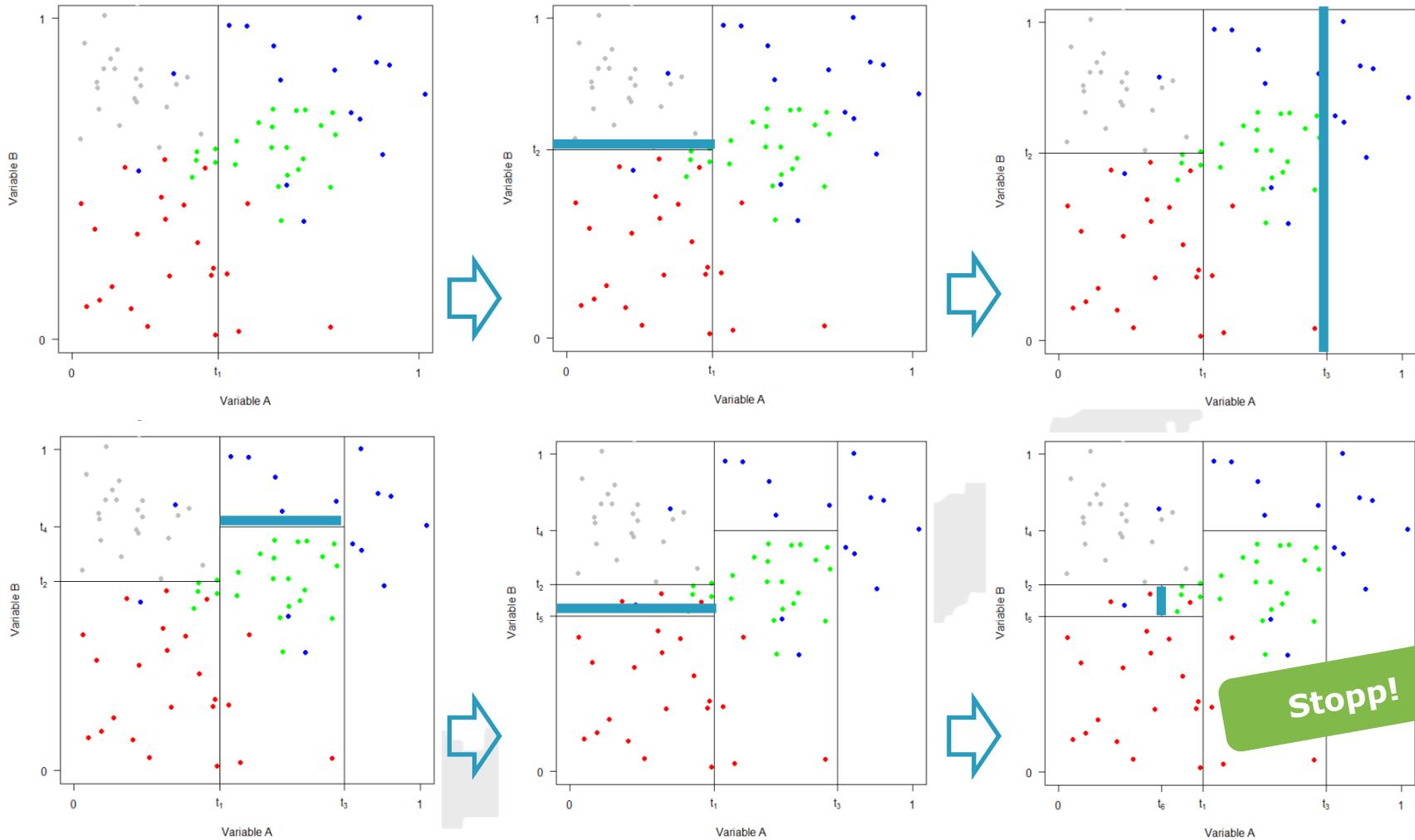
Bilderquelle: Weinhold (2014), Analyse und Anwendung von Entscheidungsbäumen zur Fehlererkennung im Gebäudebetrieb.



Wie geht das ungefähr?

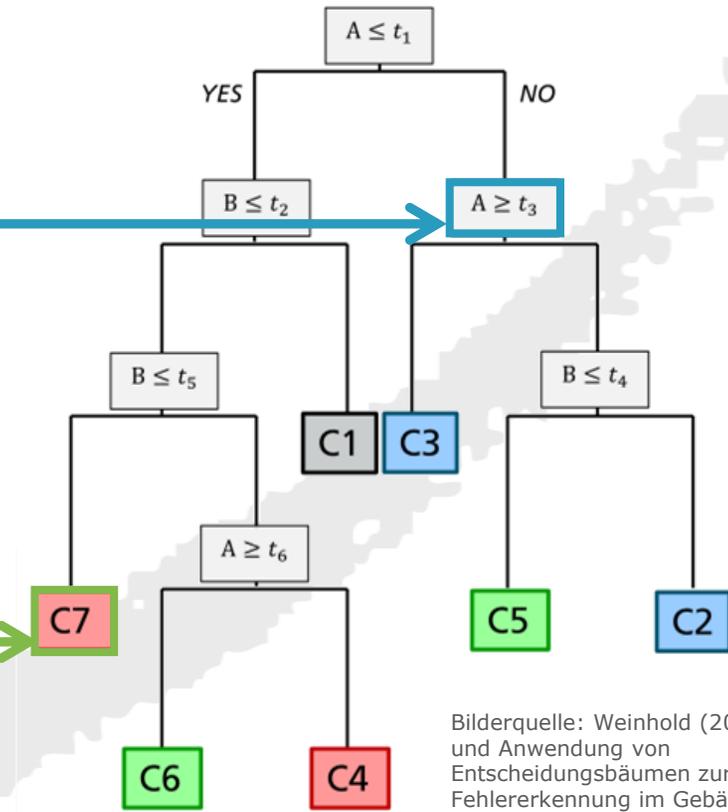
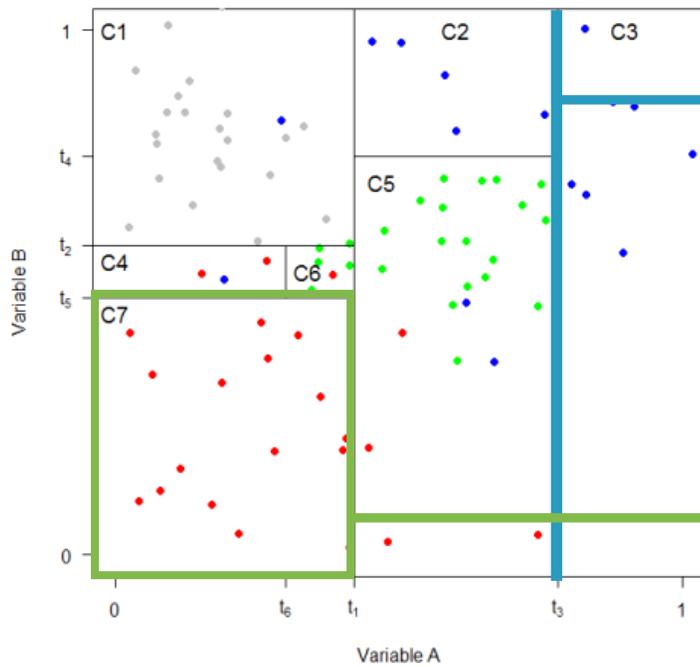
Beispiel: Klassifikationsbaum

Bilderquelle: Weinhold (2014), Analyse und Anwendung von Entscheidungsbäumen zur Fehlererkennung im Gebäudebetrieb.



Wie geht das ungefähr?

Beispiel: Klassifikationsbaum



Bilderquelle: Weinhold (2014), Analyse und Anwendung von Entscheidungsbäumen zur Fehlererkennung im Gebäudebetrieb.

Wie geht das ungefähr?

Deployment

Wie wird ein (erfolgreicher) Algorithmus in die Geschäftsprozesse integriert?

Zunächst muss das Modell **final bewertet** werden und über die **Umsetzung entschieden** werden.

- Kann die angestrebte Maßnahme (technisch) tatsächlich umgesetzt werden?
- Hat die Maschine **Muster erkannt**? Wollen und können wir die Muster interpretieren?
- Ist die Mustererkennung **gut genug**, sodass sich die Maßnahme wirklich lohnt?
- Welchen Aufwand und Effekt erwarten wir?

Wie wird **nach dem Deployment** fortgefahren?

- Wie **controllen** wir das Modell?
- Wie kann das Modell **aktualisiert** werden?

Beispiel Hochkostenfälle

Bewertung: hilft der Output für eine Prävention zukünftiger Hochkostenfälle

- Sagt der Algorithmus **zukünftige** Hochkostenfälle hinreichend gut vorher?
- Kann die Vorhersage technisch für den aktuellen Bestand erfolgen?
- Wie viele Hochkostenfälle werden erkannt, wie viele fälschlicherweise vorhergesagt und was kostet uns die Durchführung?
- Controlling: Versuchsgruppe mit und Kontrollgruppe ohne Präventionsmaßnahme



- Mit dem finalen Algorithmus erfolgt die genaue Abschätzung der **Zielerreichung**.
- Schon beim Deployment muss zwingend das **zukünftige Controlling** und die **Weiterentwicklung des Modells** vorgesehen werden!

Machine Learning, Data Analytics und Co.

Agenda

Was ist das eigentlich?

Wie geht das ungefähr?

Und viel wichtiger: Was kann man damit anfangen?

Ein potenzielles Zitat

Nutzen von Machine Learning

Anwendungsbeispiele

Fazit

Institut für Finanz- und Aktuarwissenschaften

Und viel wichtiger: Was kann man damit anfangen?

Ein potenzielles Zitat

”

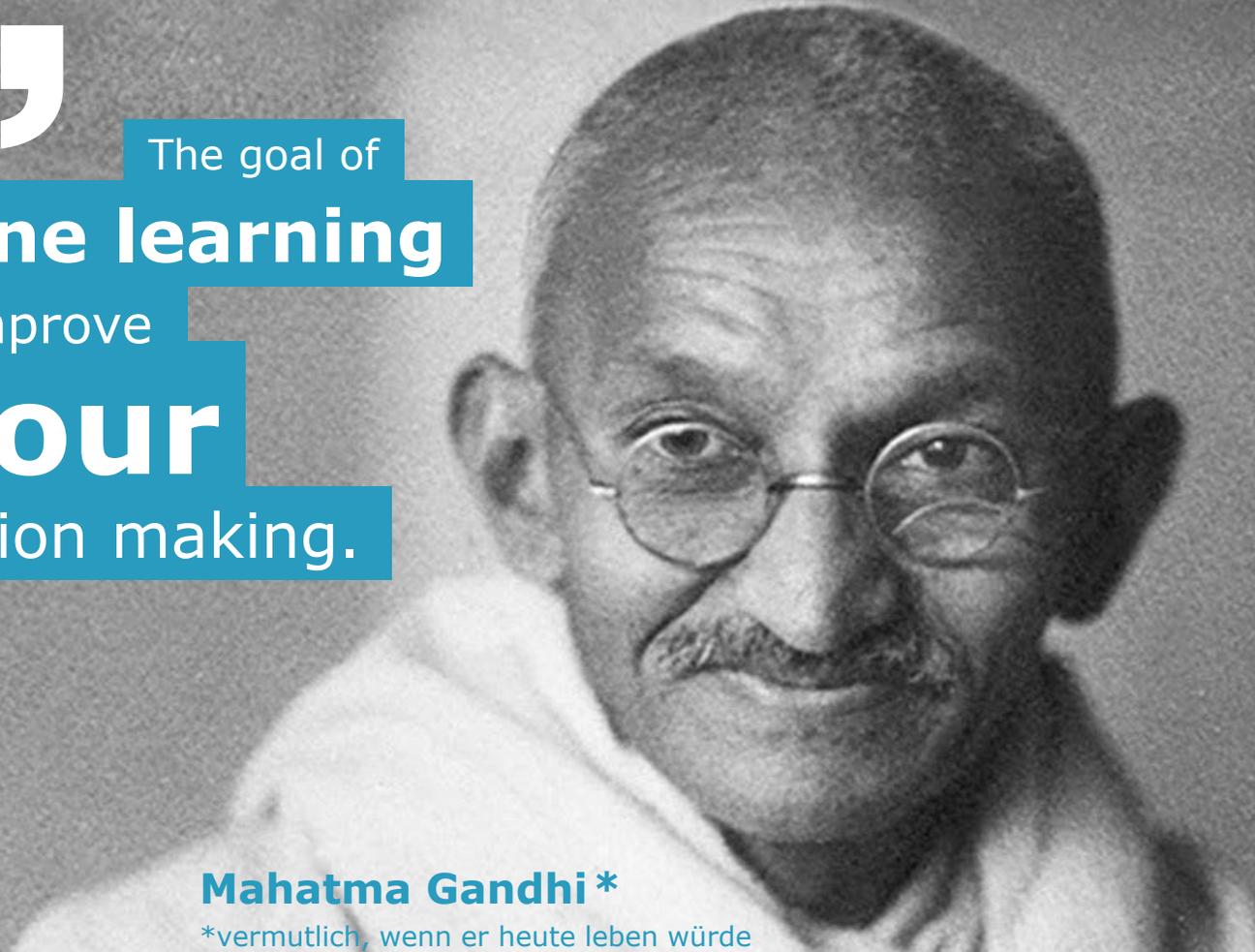
The goal of

machine learning

is to improve

our

decision making.



Mahatma Gandhi*

*vermutlich, wenn er heute leben würde

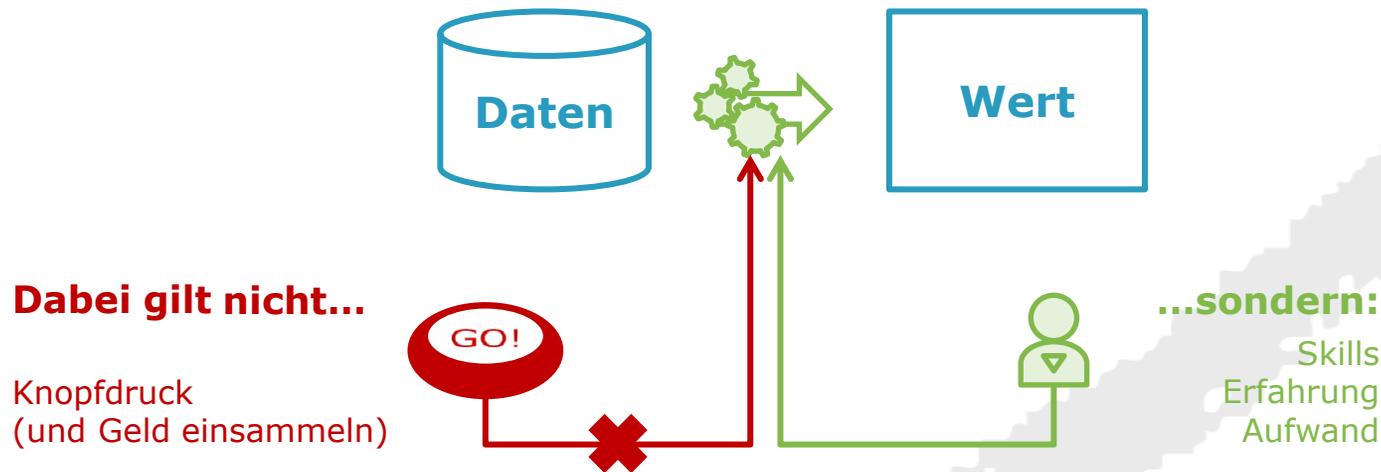
Bildquelle: Daily Express

<https://www.express.co.uk/life-style/top10facts/911836/mahatma-ghandi-facts-indian-independence-activist-top-ten-facts>

Und viel wichtiger: Was kann man damit anfangen?

Nutzen von Machine Learning

Aus der Natur der Algorithmen für Machine Learning (und Data Mining, Predictive Modelling, etc.) ergibt sich als zentraler **unternehmerischer Nutzen**, aus **Daten** einen **(Mehr-)Wert** für unsere Entscheidungsfindungen zu generieren.



Grundvoraussetzungen für eine **Wertschöpfung** sind neben fachlicher Expertise

- ein klar definiertes Ziel inkl. angestrebter Maßnahmen und
- deren fachliche Berücksichtigung innerhalb des ML-Algorithmus.

Und viel wichtiger: Was kann man damit anfangen?

Anwendungsbeispiele

Kundenzentrierter Vertrieb

■ **Cross- und Upselling** im eigenen Bestand

- Handlungsempfehlungen im Vertrieb und automatisierte Recommender-Systeme
- auf Basis interner Bestandsdaten (Assoziationen) oder externen Daten (Identifikation von verändertem Versicherungsbedarf)
- Identifikation typischer BU-Kunden
- Potenzial für Zuzahlungen

■ **Stornoprophylaxe** und Kundenrückgewinnung zur Bestandssicherung

- Identifizierung von Verträgen oder Vertragskonstellationen mit erhöhter Stornoquote
- Möglichkeiten der Kundenansprache in Verbindung mit einem Kundenwertmodell

■ **Vertriebscontrolling** und -kanaloptimierung

- Vermittler nach Gütekriterien wie Storno und Kundenwert des Geschäfts clustern
- mögliche Maßnahmen zur Incentivierung
- Identifikation von Betrug oder schlechtem Geschäft durch Vermittler
- Frühwarnsystem von atypischem Geschäft einzelner Vermittler

■ bessere **Annahmeregeln** und Risikoprüfung

- Erkenntnisse aus historischen, sich langfristig realisierenden Leistungsdaten
- langfristiges Tarifwechsel-/Stornoverhalten bei Risikozuschlägen bei ähnlichen Anträgen
- dynamische Risikoprüfung durch intelligente Steuerung der Folgefragen auf Basis bereits beantworteter Fragen

Und viel wichtiger: Was kann man damit anfangen?

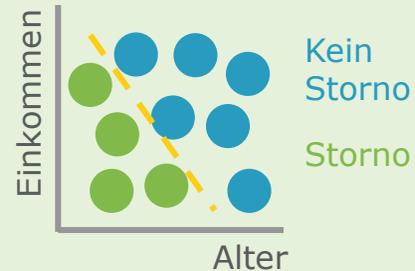
Anwendungsbeispiele

Übersicht gängiger **überwachter** Machine-Learning-Tasks

Überwacht: Für jeden Input gibt es einen Output.

Klassifikation

- Einflussgrößen → Klasse (Kategorie) der Zielgröße



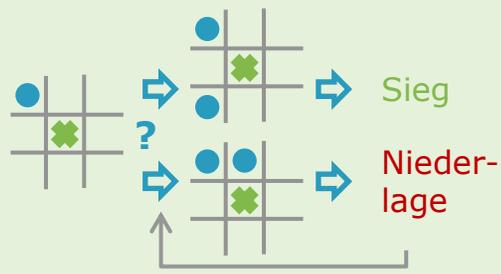
Regression

- Einflussgrößen → reellwertige Ausprägung der Zielgröße



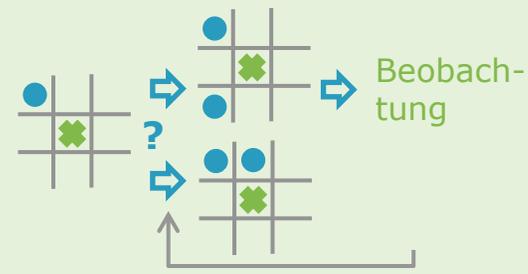
Bestärkung

- Situation → Aktion → Erfolg



Imitierung

- Situation → Aktion → beobachtete Aktion



Und viel wichtiger: Was kann man damit anfangen?

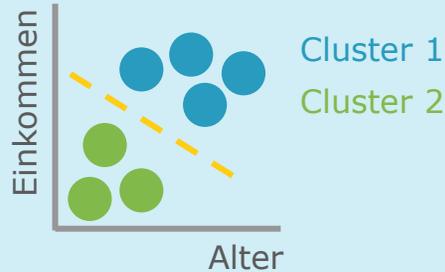
Anwendungsbeispiele

Übersicht gängiger unüberwachter Machine-Learning-Tasks

Unüberwacht: Es gibt keinen klar definierten Output.

Clustering

- Merkmale → Gruppen



Assoziation

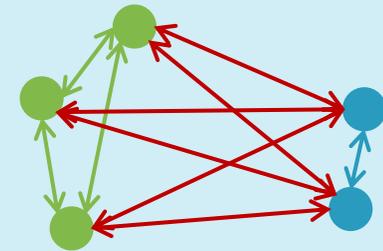
- Merkmale → Kombinationen

Bier, Windeln, Milch
Bier, Windeln, Eier
Brot, Zeitung, Mehl
Bier, Windeln, Mehl
Bier, Mehl
Bier, Windeln, Milch

Bier + Windeln

Ähnlichkeit

- Paare → Ähnlichkeitsmaß



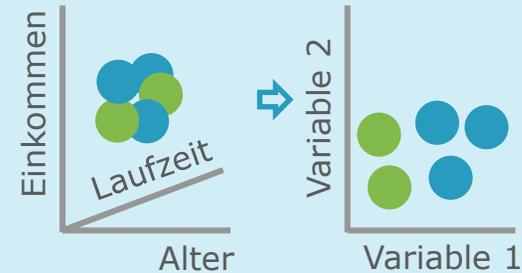
Anomalitäten

- Merkmale → Ausreißer



Dimensionsreduktion

- viele Merkmale → wenige Merkmale



Und viel wichtiger: Was kann man damit anfangen?

Anwendungsbeispiele

Zusammenfassung

überwachter ML-Task

- Vorhersage (Klassifikation und Regression)
- Bestärkung und Imitierung

Anwendung **mit** vorgegebener Zielgröße

- Stornoprophylaxe; Steuerung von Kapitalwahl und Fondswechsel; bessere Annahmeregeln; individuelle Customer Journey
- Standardvorgänge in der Sachbearbeitung automatisieren

unüberwachter ML-Task

- Clustering
- Assoziation und Ähnlichkeit
- Anomalitäten
- Dimensionsreduktion

Anwendung **ohne** vorgegebene Zielgröße

- Vertriebscontrolling; Bestandssegmentierung; Bestandsverdichtung
- Cross- und Upselling
- Betrugserkennung; Verletzung vorvertraglicher Anzeigepflichten
- Kommunikation von Ergebnissen; Wissensgenerierung

Und viel wichtiger: Was kann man damit anfangen?

Anwendungsbeispiele



Machine Learning, Data Analytics und Co.

Agenda

Was ist das eigentlich?

Wie geht das ungefähr?

Und viel wichtiger: Was kann man damit anfangen?

Fazit

Institut für Finanz- und Aktuarwissenschaften

Themen identifizieren

- Ziele, Maßnahmen und Bewertungskriterien definieren
- zielgerichtete Daten erfassen
- **Wichtig:** Grundvoraussetzung für erfolgreiche Machine-Learning-Projekte

Machine Learning einsetzen

- „State-of-the-art“-Methoden zur bestmöglichen Mustererkennung verwenden
- dabei betriebswirtschaftlichen Nutzen zur Wertschöpfung optimieren
- **Wichtig:** hohes Maß an Erfahrung mit Daten und Methoden erforderlich

In Geschäftsprozessen umsetzen

- Potenzial zur optimierten Prozessgestaltung ausnutzen
- laufendes Lernen von Mustern zur Anpassung der Modelle vorsehen
- **Wichtig:** Controlling der Modelle und Kommunikation im Unternehmen erforderlich

Machine Learning, Data Analytics, etc.

Agenda

Was ist das eigentlich?

Wie geht das ungefähr?

Und viel wichtiger: Was kann man damit anfangen?

Institut für Finanz- und Aktuarwissenschaften

Kontakt

Beratungsangebot

Formale Hinweise

Kontakt

Lukas Hahn

+49 731 20644-239

l.hahn@ifa-ulm.de



Beratungsangebot

Life



Produktentwicklung
Biometrische Risiken
Zweitmarkt

Non-Life



Produktentwicklung
und Tarifierung
Schadenreservierung
Risikomodellierung

Health



Aktuarieller
Unternehmenszins
Leistungsmanagement

**Actuarial
Consulting**

Solvency II ▪ Embedded Value ▪ Asset-Liability-Management
ERM ▪ wert- und risikoorientierte Steuerung ▪ Data Analytics

Projektmanagement ▪ Markteintritt ▪ Bestandsmanagement ▪ strategische Beratung

**Actuarial
Services**

aktuarielle Großprojekte ▪ aktuarielle Tests
Überbrückung von Kapazitätsengpässen

Research



Aus- und Weiterbildung



... weitere Informationen
unter www.ifa-uhl.de

Formale Hinweise

- Dieses Dokument ist in seiner Gesamtheit zu betrachten, da die isolierte Betrachtung einzelner Abschnitte möglicherweise missverständlich sein kann. Entscheidungen sollten stets nur auf Basis schriftlicher Auskünfte gefällt werden. Es sollten grundsätzlich keine Entscheidungen auf Basis von Versionen dieses Dokuments getroffen werden, welche mit „Draft“ oder „Entwurf“ gekennzeichnet sind. Für Entscheidungen, welche diesen Grundsätzen nicht entsprechen, lehnen wir jede Art der Haftung ab.
- Dieses Dokument basiert auf unseren Marktanalysen und Einschätzungen. Wir haben diese Informationen vor dem Hintergrund unserer Branchenkenntnis und Erfahrung auf Konsistenz hin überprüft. Eine unabhängige Beurteilung bzgl. Vollständigkeit und Korrektheit dieser Information ist jedoch nicht erfolgt. Eine Überprüfung statistischer bzw. Marktdaten sowie mit Quellenangabe gekennzeichnete Informationen erfolgt grundsätzlich nicht. Bitte beachten Sie auch, dass dieses Dokument auf Grundlage derjenigen Informationen erstellt wurde, welche uns zum Zeitpunkt seiner Erstellung zur Verfügung standen. Entwicklungen und Unkorrektheiten, welche erst nach diesem Zeitpunkt eintreten oder offenkundig werden, können nicht berücksichtigt werden. Dies gilt insbesondere auch für Auswirkungen einer möglichen neuen Aufsichtspraxis.
- Unsere Aussagen basieren auf unserer Erfahrung als Aktuare. Soweit wir bei der Erbringung unserer Leistungen im Rahmen Ihrer Beratung Dokumente, Urkunden, Sachverhalte der Rechnungslegung oder steuerrechtliche Regelungen oder medizinische Sachverhalte auslegen müssen, wird dies mit der angemessenen Sorgfalt, die von uns als professionellen Beratern erwartet werden kann, erfolgen. Wenn Sie einen verbindlichen Rat, zum Beispiel für die richtige Auslegung von Dokumenten, Urkunden, Sachverhalten der Rechnungslegung, steuerrechtlichen Regelungen oder medizinischer Sachverhalte wünschen, sollten Sie Ihre Rechtsanwälte, Steuerberater, Wirtschaftsprüfer oder medizinische Experten konsultieren.
- Dieses Dokument wird Ihnen vereinbarungsgemäß nur für die innerbetriebliche Verwendung zur Verfügung gestellt. Die Weitergabe – auch in Auszügen – an Dritte außerhalb Ihrer Organisation sowie jede Form der Veröffentlichung bedarf unserer vorherigen schriftlichen Zustimmung. Wir übernehmen keine Verantwortung für irgendwelche Konsequenzen daraus, dass Dritte auf diese Berichte, Ratschläge, Meinungen, Schreiben oder anderen Informationen vertrauen.
- Jeglicher Verweis auf ifa in Zusammenhang mit diesem Dokument in jeglicher Veröffentlichung oder in verbaler Form bedarf unserer ausdrücklichen schriftlichen Zustimmung. Dies gilt auch für jegliche verbale Informationen oder Ratschläge von uns in Verbindung mit der Präsentation dieses Dokumentes.