

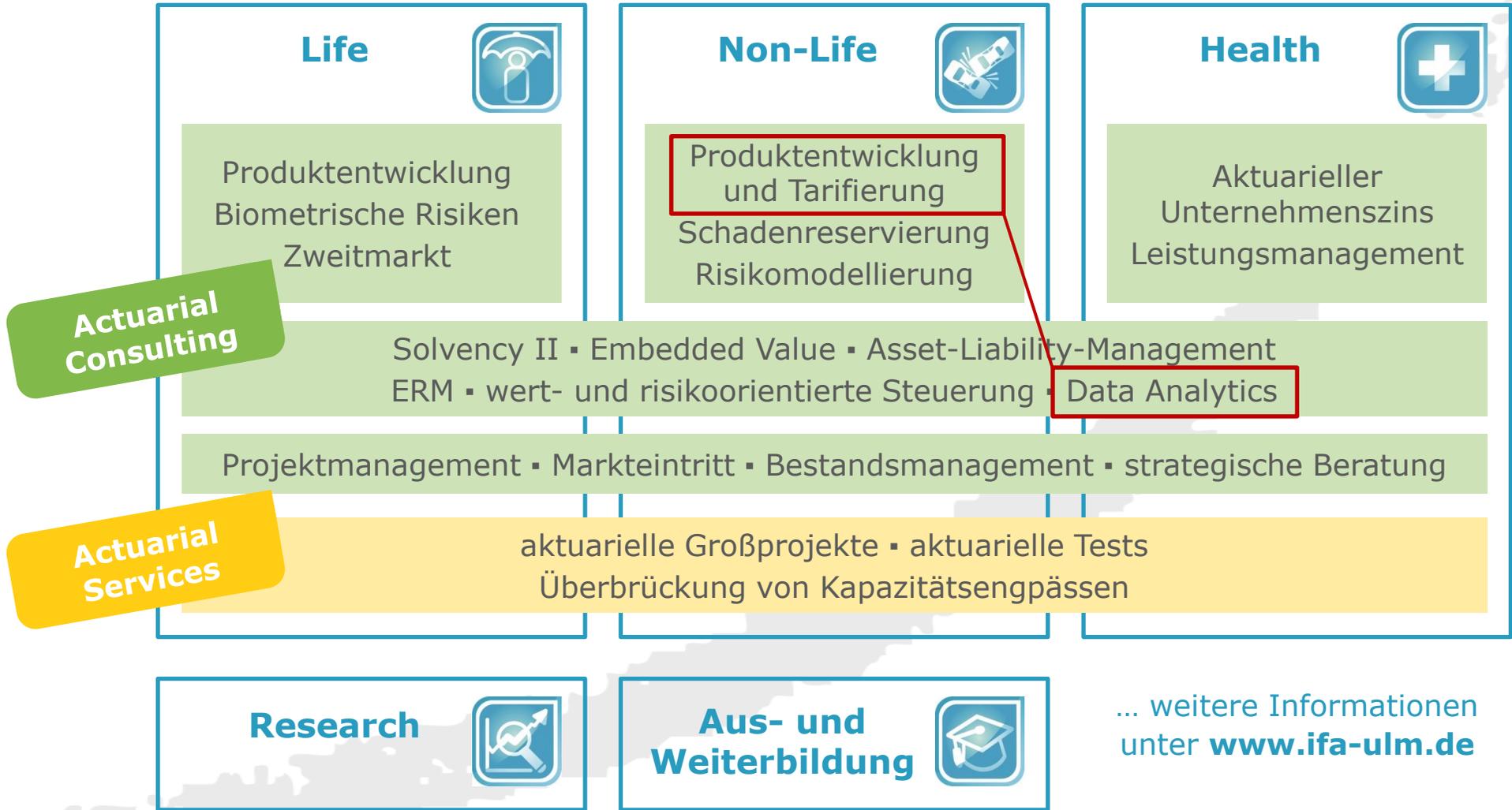
Data Analytics in der Tarifierung

- Dr. Andreas Reuß
- 14. MCC-Fachkonferenz – „Kfz-Versicherung der Zukunft“
- Düsseldorf, 24. September 2018



Institut für Finanz- und Aktuarwissenschaften (ifa)

Unser Beratungsangebot



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

Data Analytics in der Tarifierung

Agenda

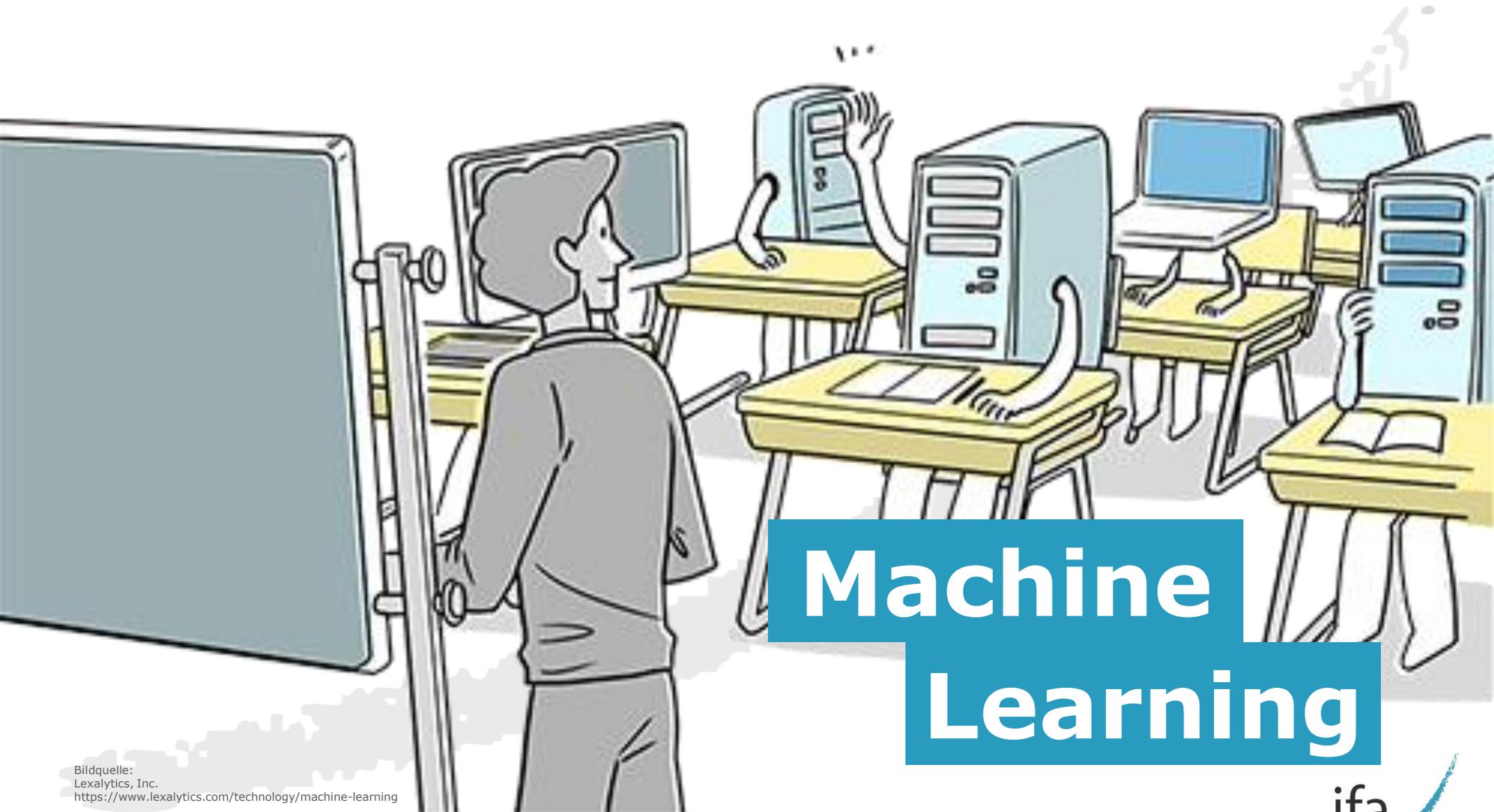
**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was ist das eigentlich?**

Ergänzung klassischer aktuarieller Methoden um Data Analytics

Weitere Anwendungen von Data Analytics

Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.

Was ist das eigentlich?



Machine Learning

Bildquelle:
Lexalytics, Inc.
<https://www.lexalytics.com/technology/machine-learning>

Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.

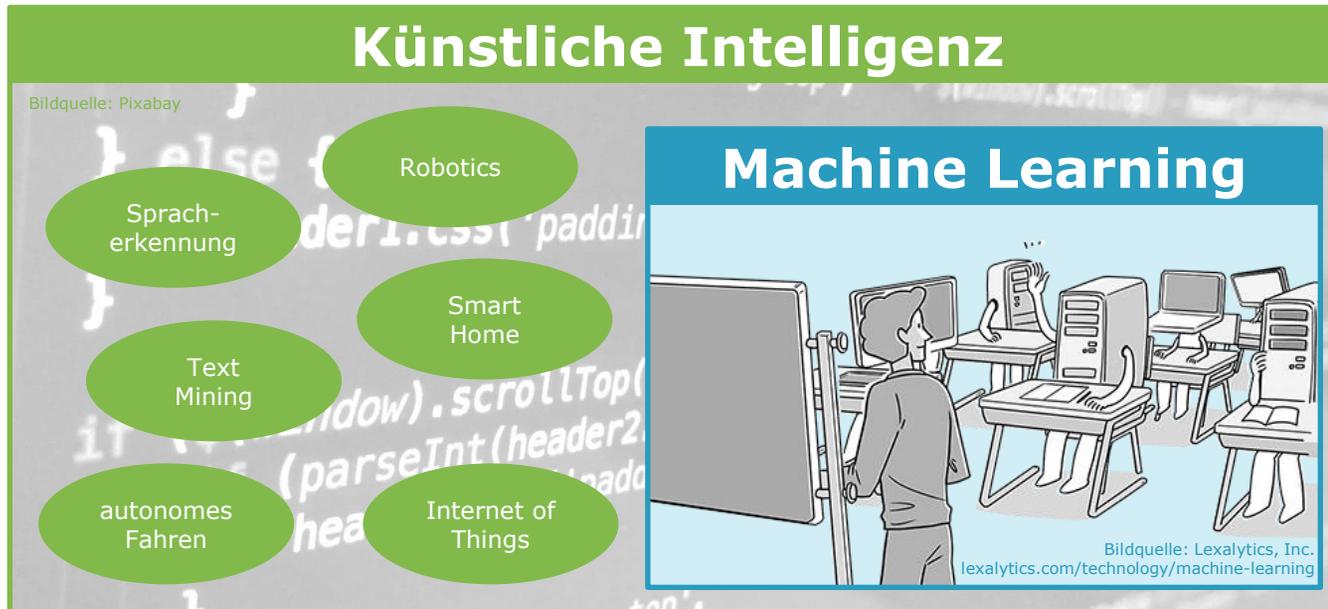
Was ist das eigentlich?



Machine Learning (ML) ist nicht die klassische Konstruktion von regelbasierten Algorithmen zur Lösung einer Aufgabe, sondern die **Konstruktion von Algorithmen**, die durch Daten so **einen Algorithmus selbst konstruieren können**.

Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.

Was ist das eigentlich?



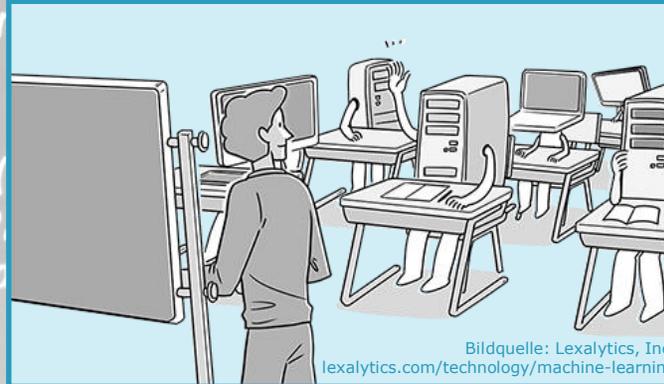
Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.

Was ist das eigentlich?

Künstliche Intelligenz

Bildquelle: Pixabay

Machine Learning



Data Mining: Identifikation komplexer Muster zur Wissensgenerierung (z.B. Clustering)

Vor Clustering



Nach Clustering



Beispiel
Telematikdaten



Predictive Modelling: bestmögliche individuelle Vorhersage für optimale Entscheidungsfindung

Verfügbare Daten



Neue Daten



Beispiel
Tarifierung

Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.

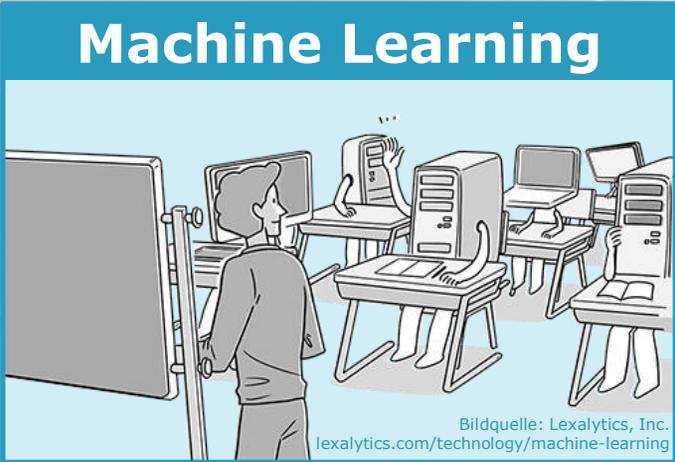
Was ist das eigentlich?

Künstliche Intelligenz

Bildquelle: Pixabay

```
} else {  
  header1.css('padding'  
}  
  
if ($(window).scrollTop()  
  if (parseInt(header2  
    header2.css('padding'  
  }  
}
```

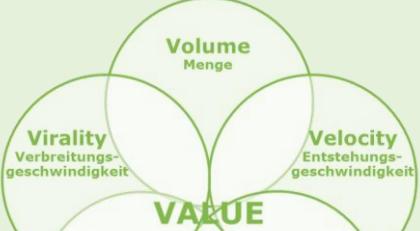
Machine Learning



Bildquelle: Lexalytics, Inc.
lexalytics.com/technology/machine-learning

Big Data

Datenmengen, welche für manuelle Auswertungsmethoden zu groß, komplex, schnelllebig oder schwach strukturiert sind

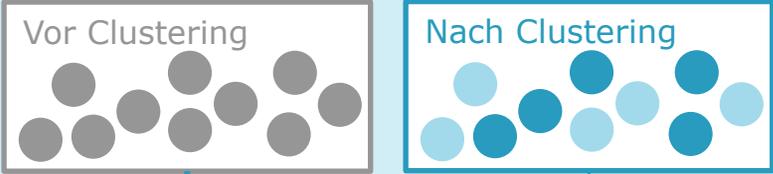


Bsp.: Telematik, eCall, Dashcams, Pools



Data Mining:

Identifikation komplexer Muster zur Wissensgenerierung (z.B. Clustering)

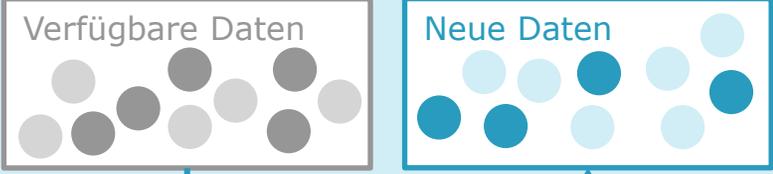


Vor Clustering Nach Clustering



Predictive Modelling:

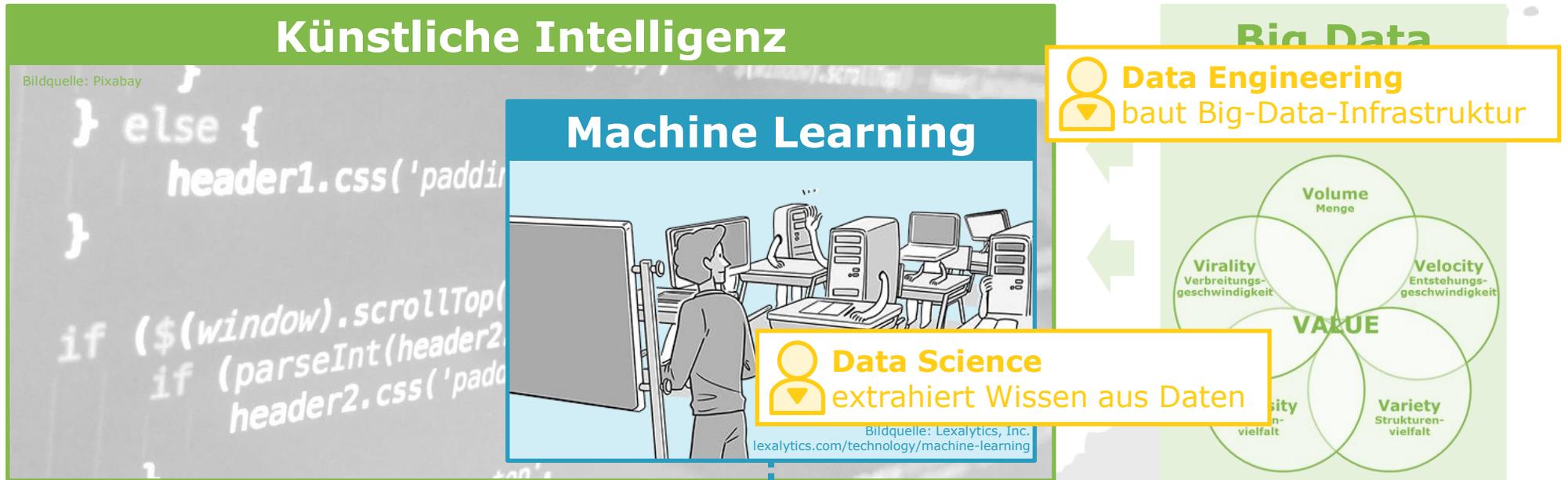
bestmögliche individuelle Vorhersage für optimale Entscheidungsfindung



Verfügbare Daten Neue Daten

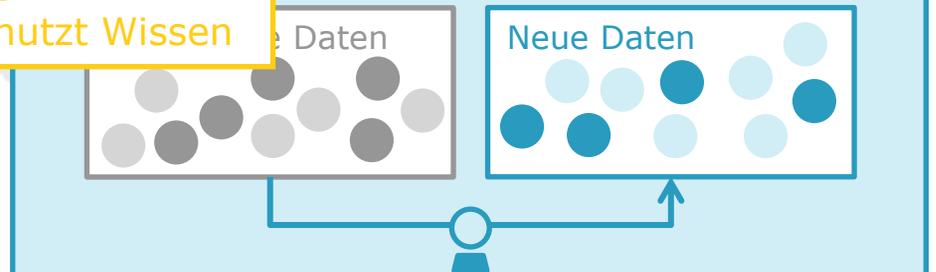
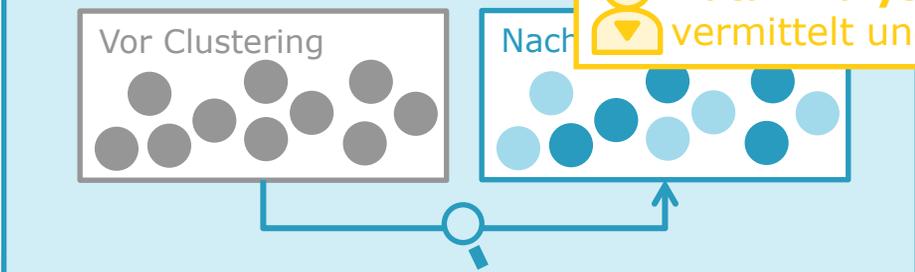
Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.

Was ist das eigentlich?



Data Mining: Identifikation komplexer Muster zur Wissensgenerierung (z.B. Clustering)

Predictive Modelling: bestmögliche individuelle Vorhersage für optimale Entscheidungsfindung



Data Analytics vermittelt und nutzt Wissen

Data Analytics in der Tarifierung

Agenda

**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was ist das eigentlich?**

Ergänzung klassischer aktuarieller Methoden um Data Analytics

Weitere Anwendungen von Data Analytics

Tarifierung

Die Revolution des klassischen Ansatz?

Klassischer Ansatz: verallgemeinerte lineare Modelle (Generalized Linear Models, GLM)

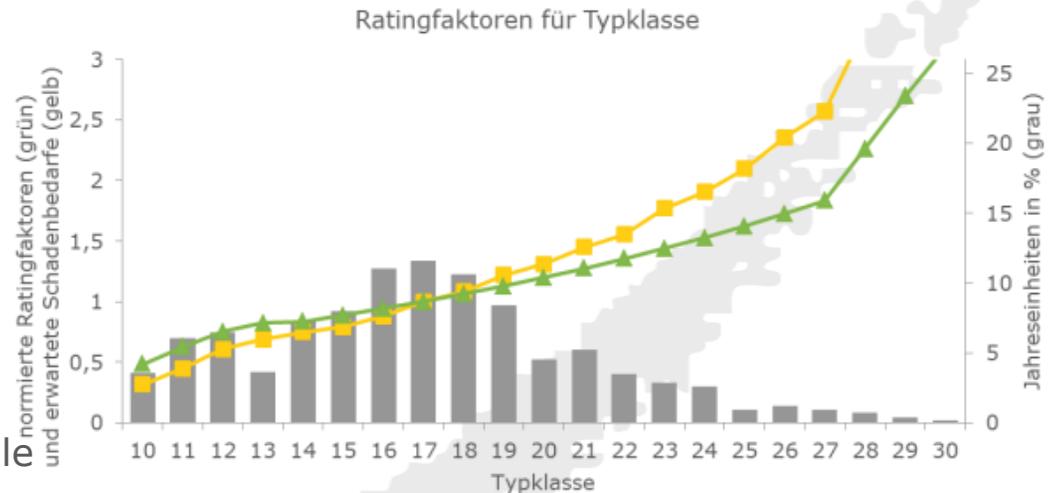
■ viele **Vorteile:**

- robuste Modelle, d.h. stabil gegenüber Ausreißern in den Daten
- trotz Linearitätsannahme sehr vielfältig (Interaktionen, Transformationen)
- Abbildung aller für die Tarifmodellierung relevanten Besonderheiten möglich
- Verständnis über Einfluss der Tarifmerkmale auf Schadenkennzahlen pro Schadenart
- leichte Visualisierung, Interpretation und Kommunikation der Ergebnisse

■ aber auch **Nachteile:**

- Anpassungsgüte wegen linearer Struktur eingeschränkt, insbesondere an den Rändern
- manuelles Modellieren anfällig für Overfitting
- fehlende komplexe Interaktionen

→ **Ausweg: GLM-Risikomodell durch komplexere Data-Analytics-Methode ersetzen?**



Tarifierung

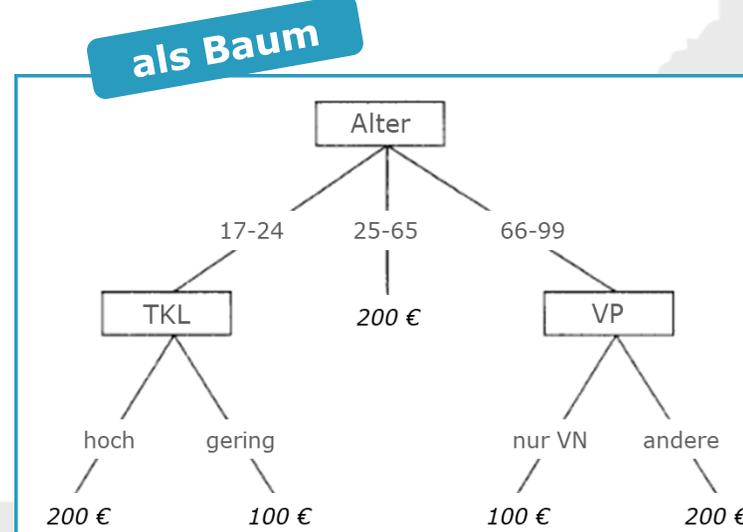
Ansatz: Risikomodellierung mit Entscheidungsbäumen

Alternativer Ansatz: Entscheidungsbäume

- Hierbei differenzieren intelligente Suchalgorithmen den erwarteten Schadenbedarf der individuellen Risiken möglichst effizient mittels sukzessiver **Entscheidungsregeln** auf Basis der Risikomerkmale.
- vereinfachtes Beispiel: welche Risiken haben geringe (100 €) bzw. hohe (200 €) Kaskoschäden?

als Daten

Alter	TKL	VP	Schaden
25-65	hoch	nur VN	200 €
25-65	gering	nur VN	200 €
25-65	gering	andere	200 €
17-24	hoch	nur VN	200 €
17-24	hoch	andere	200 €
17-24	gering	nur VN	100 €
17-24	gering	andere	100 €
66-99	hoch	nur VN	100 €
66-99	gering	nur VN	100 €
66-99	hoch	andere	200 €
66-99	gering	andere	200 €



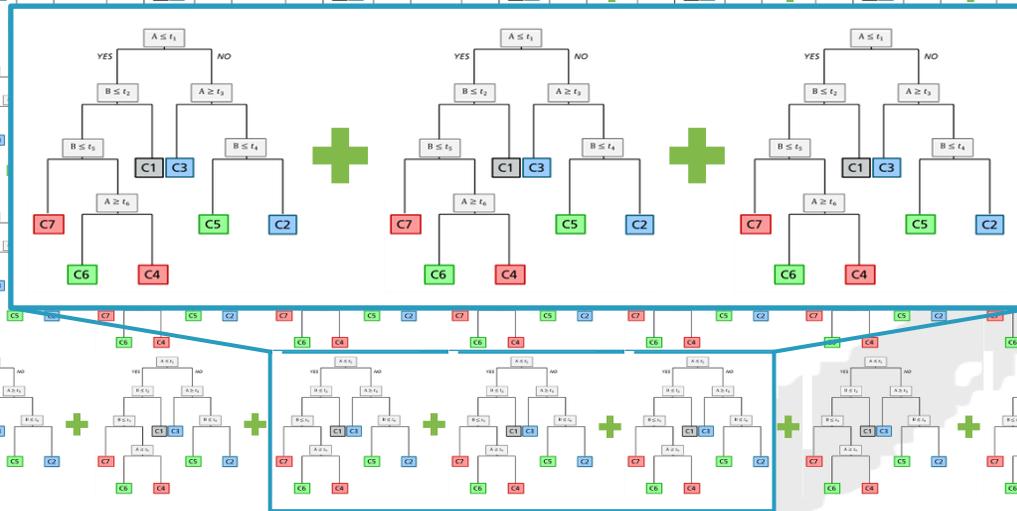
- (Kleine) Entscheidungsbäume sind zwar weiterhin **interpretierbar**, aber...
- **instabiler** (starke Modelländerung bei kleiner Datenänderung), anfällig für **Overfitting** und i.d.R. **nicht performanter** als ein (gut modelliertes) GLM.

Tarifierung

Ansatz: Risikomodellierung mit Wäldern und Boosting und Netzen und ...

Erfolgreiche Data-Analytics-Methoden sind deshalb komplexe ML-Verfahren.

■ ein Ansatz: viele Bäume parallel wachsen lassen: Entscheidungswälder („Random Forest“)



Durch wiederholtes zufälliges Ziehen von Teilmengen der Daten und der Risikomerkmalen werden viele unterschiedliche Entscheidungsbäume erzeugt. → „Random Forest“ als Durchschnitt

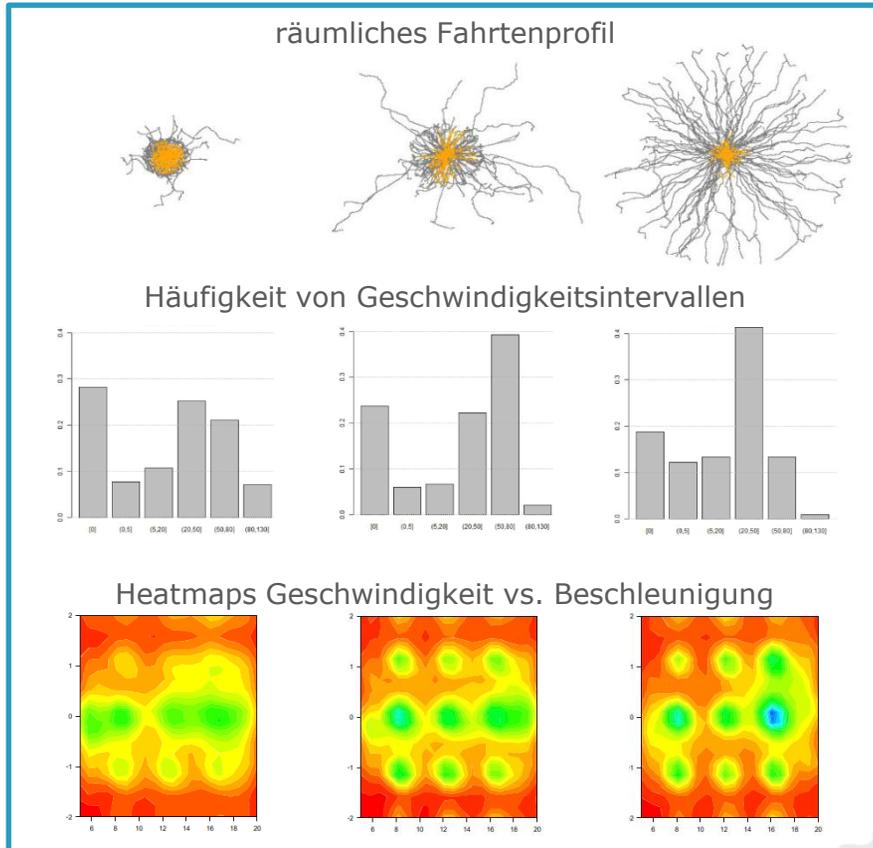


- deutlich bessere Anpassung zu erwarten (nicht-lineare Muster, Interaktionen, ...)
- aber zu Lasten der Interpretation (Effekte der Merkmale nicht nachvollziehbar)
- ➔ Tarifmodellierung nicht wie bisher möglich – aber auch regulatorische Fragen

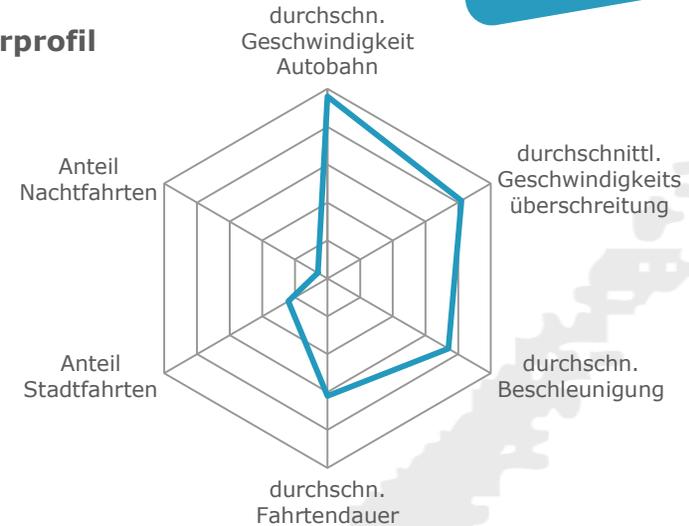
Tarifierung

Big Data modellieren – Beispiel: Telematikdaten

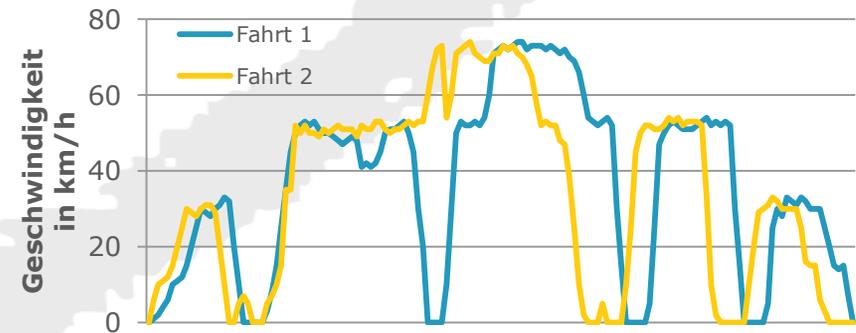
illustrativ



Fahrerprofil



Identifikation von Routinefahrten



- Mit Data Analytics lassen sich große unstrukturierte Telematikdaten analysieren.
- So lassen sich bspw. Fahrerprofile ableiten oder Routinefahrten identifizieren.
- Gewonnene Erkenntnisse (z.B. Straßentypnutzung) lassen sich in Verbindung mit Fachwissen und externen Daten (z.B. statistisches Bundesamt) verwerten.

ifa

Tarifierung

GLMs unterstützen – Beispiel: Interaktionseffekte und nicht-lineare Zusammenhänge

Interaktionseffekte

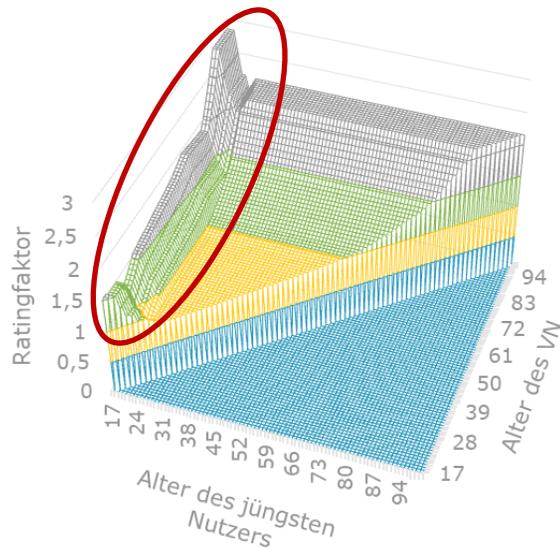
Beispiel: Analyse des Alters des Versicherungsnehmers und des jüngsten Nutzers

illustrativ

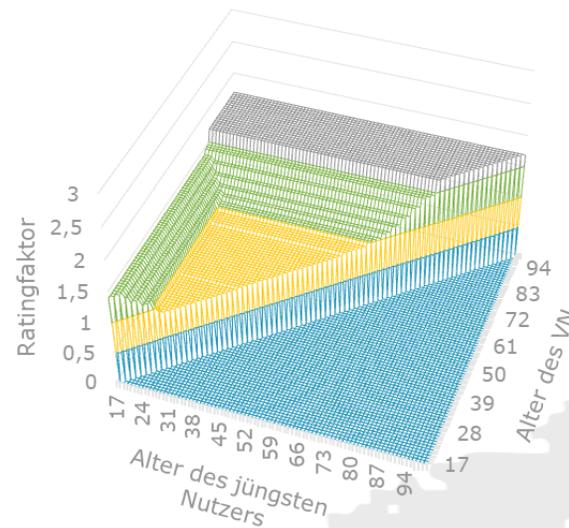
Nicht-lineare Effekte

Beispiel: Alters des Versicherungsnehmers

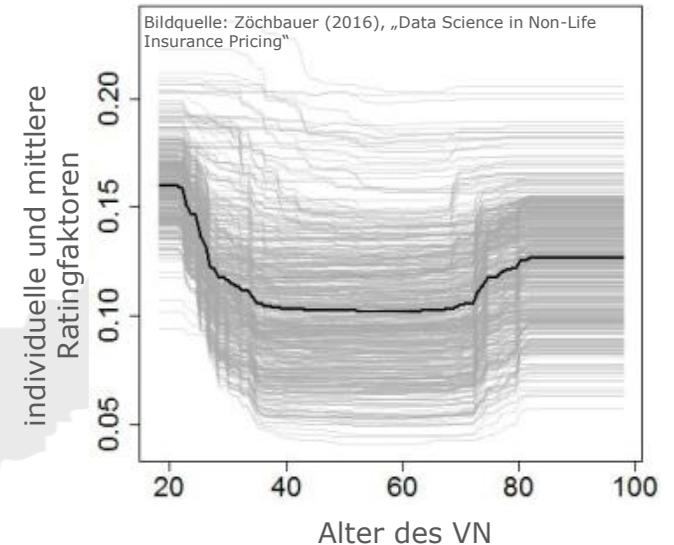
datengetriebene Interaktion



Interaktion im GDV-Tarif 2017



PDP/ICE-Plot

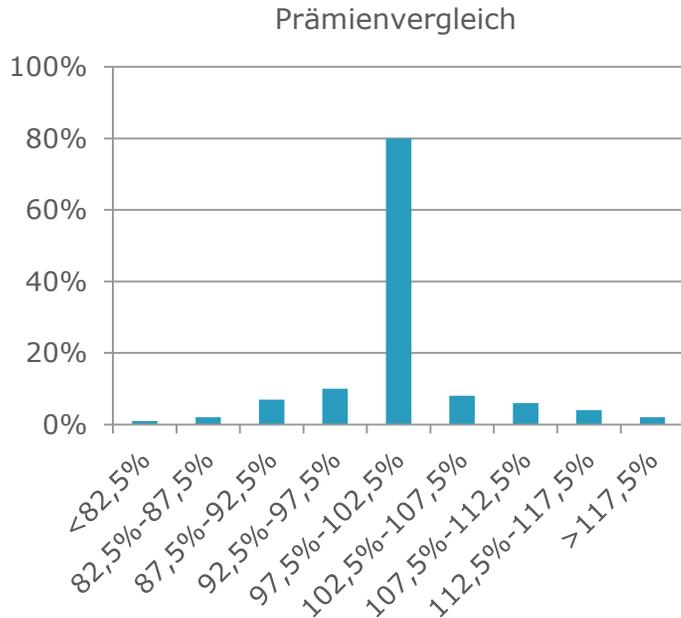


- Mit Data Analytics lassen sich in den umfangreichen Daten **komplexe Interaktionen und Zusammenhänge identifizieren**, die anschließend im GLM explizit modelliert werden können.

Tarifierung

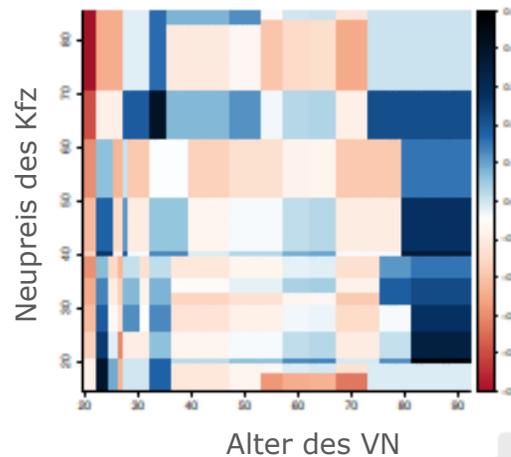
GLMs benchmarken – Beispiel: Abweichungsanalyse und Einzelfalluntersuchung

Abweichungsanalyse



illustrativ

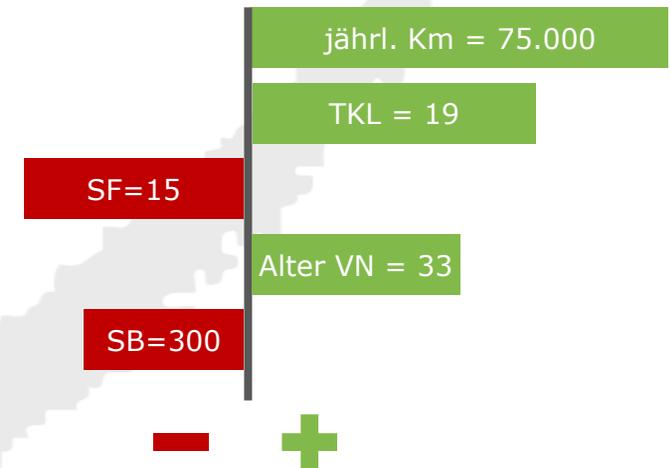
Relative Differenz in Abhängigkeit der Tarifmerkmale



Bildquelle: Zöchbauer (2016), „Data Science in Non-Life Insurance Pricing“

Einzelfalluntersuchung

Ursachenforschung mittels „Local Interpretable Model-Agnostic Explanations“ (LIME)



Zusammensetzung der Prämie für Max Mustermann anhand seiner Merkmalsausprägungen



- Die Ergebnisse des GLMs lassen sich gegen komplexere ML-Verfahren benchmarken, um defizitäre Teilbestände/Segmente und Verbesserungspotenzial des GLMs zu erkennen.
- Stark abweichende Einzelfälle können genauer untersucht werden.

Tarifierung

White-Box statt Black-Box

White-Box-Ansätze: Data-Analytics-Methoden, die sich zur verbesserten Risikomodellierung **nahtlos in die klassische Anwendung von GLM** integrieren lassen.

- Der Ansatz ist beispielsweise dazu geeignet,
 - eine in die Modellanpassung **integrierte Variablen- und Interaktionsselektion** automatisiert und datengetrieben durchzuführen,
 - neue oder granularere Tarifmerkmale beim Fitting **implizit zu dämpfen** oder
 - im GLM bisher nicht erkannte **nicht-lineare Effekte zu identifizieren** und geeignet zu berücksichtigen.
- Zugehörige methodische Ansätze sind u.a.
 - sogenannte **Regularisierungsverfahren** zur impliziten Bestrafung von Overfitting (LASSO, Ridge-Regression, Elastic Nets),
 - Generalized Linear **Mixed** Models (GLMM) zur impliziten Dämpfung von Ratingfaktoren mittels Credibility-Ansätzen,
 - (teilweiser) Übergang zu Generalized **Additive** Models (GAM) für die nicht-lineare Modellierung einzelner Merkmale (z.B. Alter),
 - Automatisierung der **Knotenpunktsuche** für stückweise stetige Funktionen und Splines.

Data Analytics in der Tarifierung

Agenda

**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was ist das eigentlich?**

Ergänzung klassischer aktuarieller Methoden um Data Analytics

Weitere Anwendungen von Data Analytics

Weitere Anwendungen von Data Analytics

Neben der Risikomodellierung bietet auch die **Ausgestaltung des Tarifs** Potenzial für **Data-Analytics-Methoden**.

- Modellierung der Wettbewerberpreise („Reverse Engineering“)
 - Machine-Learning-Verfahren sind hierfür geeignet, da nur systematische Muster vorhanden und zu identifizieren sind.
 - Verständnis über die Preisstruktur der Wettbewerber bei der Tarifausgestaltung nutzen
- Modellierung von Markteinflüssen
 - Beispielsweise lässt sich die Risikoprämie anpassen, abhängig davon
 - wie preiselastisch der Kunde eingeschätzt wird,
 - welche Abschluss- oder Verlängerungswahrscheinlichkeit erwartet wird und
 - welcher Kundenwert (z.B. erwartete Verweildauer im Bestand) geschätzt wird.
 - Die entsprechenden Modelle können in die Tarifausgestaltung integriert und so für die Preispolitik genutzt werden.

Weitere Anwendungen von Data Analytics

Data Analytics kann auch bei anderen Fragestellungen messbaren Mehrwert liefern.

Kunde

- **Neue Verträge gewinnen:** profitables Cross- und Upselling im eigenen Bestand
- **Den Kunden halten:** Stornoprophylaxe zur Bestandssicherung
- **Den Vertriebspartner kennen:** Vertriebscontrolling und -kanaloptimierung
- **Den Kunden kennen:** verbesserte Annahmeregeln und Risikoprüfung

Prozesse

- **Datengetriebene Entscheidungshilfen geben:** Nutzung externer Dienstleister, gezielte Qualitätskontrollen
- **Den Regulierungsprozess steuern:** Potenzialrouting, Beschleunigung, Nutzung von Expertenkapazitäten
- **Die Schadenregulierung automatisieren:** automatisierte Prüfung und Kürzung, Dunkerverarbeitung

Fazit

Data Analytics in der Tarifierung bedeutet **Evolution statt Revolution!**

- Es bietet eine **Erweiterung der Toolbox** des Aktuars zur verbesserten Tarifierung unter Beibehaltung von Modellen zur Interpretation, Validierung, Erklärung, Anpassung etc.
- Ansatzpunkte zur **intelligenten Ergänzung des Tarifierungsprozess** um Data-Analytics-Methoden reichen von der Auswertung von Big-Data-Datenquellen über den unterstützenden Einsatz von Black-Box-Modellen bis hin zur Erweiterung des GLM-Frameworks durch White-Box-Verfahren.
- Data Analytics erlaubt eine **datengetriebene Mustererkennung**, die neben der Tarifierung auch in anderen Anwendungsfeldern genutzt werden kann.
- Grundvoraussetzungen für eine **Wertschöpfung aus Daten** sind
 - frühzeitige Schaffung der notwendigen Datengrundlage,
 - fachliche Expertise (klar definierte Ziele inkl. Maßnahmen, Datenkenntnis, Expertenwissen) und
 - statistisches Knowhow und Erfahrung für deren Transfer in ein modernes Data-Analytics-Modell.

Institut für Finanz- und Aktuarwissenschaften

Kontaktdaten

Dr. Andreas Reuß

Director, Partner

+49 (731) 20 644-251

a.reuss@ifa-ulm.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.