

Big Data, künstliche Intelligenz und Data Analytics

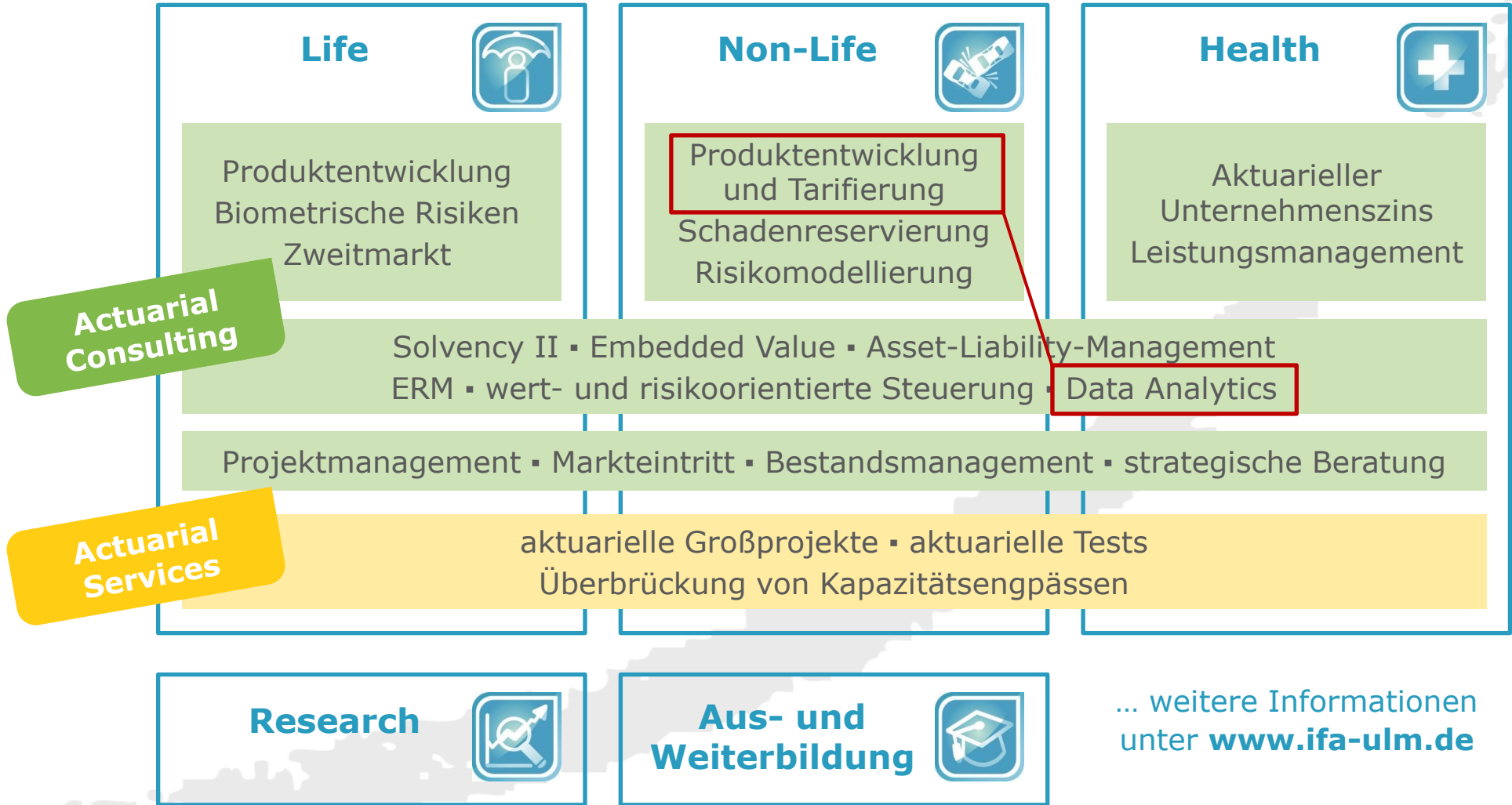
in der Tarifierung, Kundendatenauswertung und Produktentwicklung

- Dr. Andreas Reuß
- 4. Jahrestagung Car Data – Telematik – Mobilität – FAS –
Autonomes Fahren – eCall – Connected Car
- Köln, 11. Oktober 2018



Institut für Finanz- und Aktuarwissenschaften (ifa)

Unser Beratungsangebot



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

Big Data, künstliche Intelligenz und Data Analytics

Agenda

**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was hat das mit Tarifierung zu tun?**

Klassische Tarifierung mit GLM

Erweiterung um Data-Analytics-Methoden

Weitere Anwendungen in Produktentwicklung und Produktmanagement

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

Was hat das mit Tarifierung zu tun?

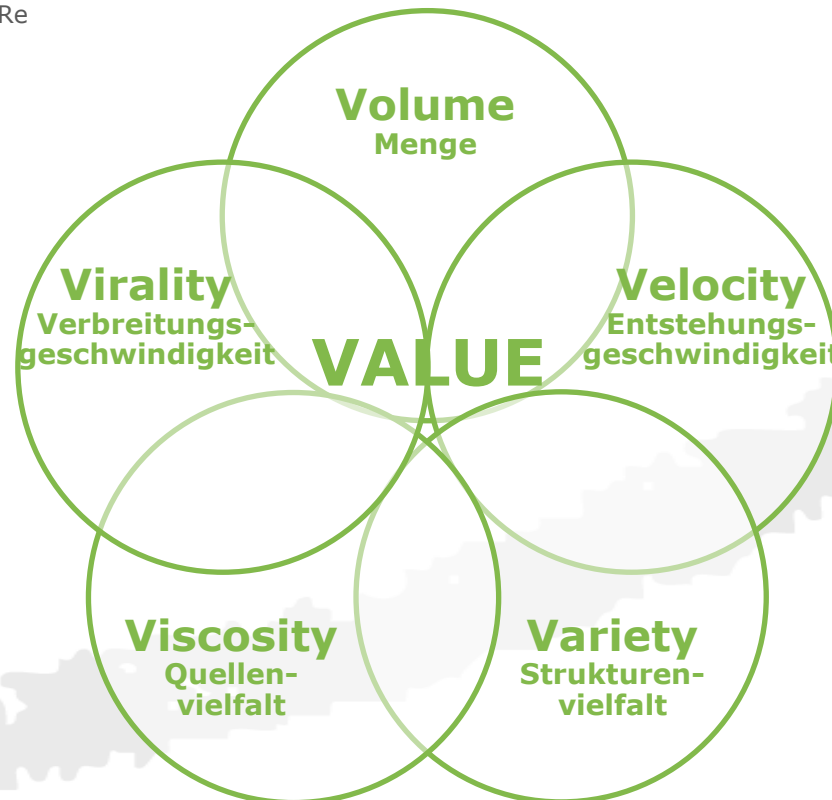
Big Data

- Definition: Datenmengen, welche zu **groß**, zu **komplex**, zu **schnelllebig** oder zu **schwach strukturiert** sind, um sie mit manuellen und herkömmlichen Methoden der Datenverarbeitung auszuwerten.

Angelehnt an: Studie zur kommerziellen digitalen Überwachung im Alltag. Österr. Bundesarbeitskammer, 2014.

- 3-V-Modell: Big Data erfüllt mindestens drei der fünf V-Kriterien

Angelehnt an: Munich Re



Big Data, künstliche

How big is big?

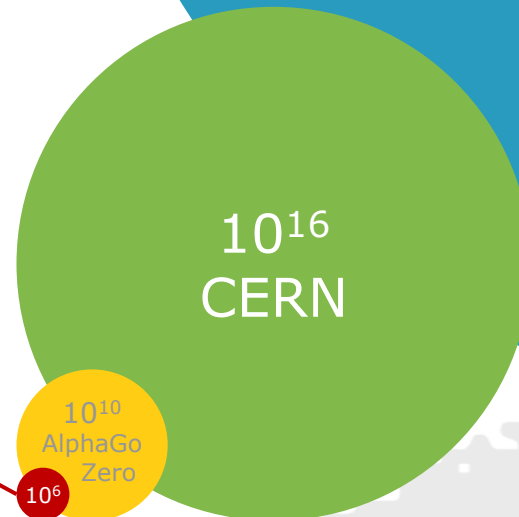
10^{19}
Google

Google hält schätzungsweise 10-15 Exabytes (10^{19} Bytes) weltweit.

CERN sammelt jährlich etwa 10^{16} Bytes aus Kollisionsexperimenten.

AlphaGo Zero hat aus 4,9 Millionen Spielen gegen sich selbst etwa 10^{10} Bytes erzeugt.

Die Polizei hat 2017 $2,6 \cdot 10^6$ **Unfälle** erfasst.



Quellen:
Schiller (2018), Big Data meets Disability Insurance, ICA 2018
Statistisches Bundesamt, Fachserie 8, Reihe 7, 05/2018



- Big Data in der Versicherung sind relativ betrachtet Tiny Data.
- Dennoch werden auch in der Versicherungsbranche durch zunehmende Digitalisierung, Telematik- und Internetdaten sowie Datenpools die Daten **umfangreicher, schnelllebiger, vielfältiger und instabiler**.

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

Was ist das eigentlich?

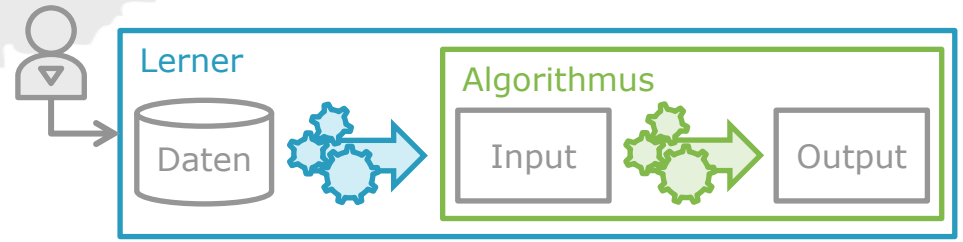
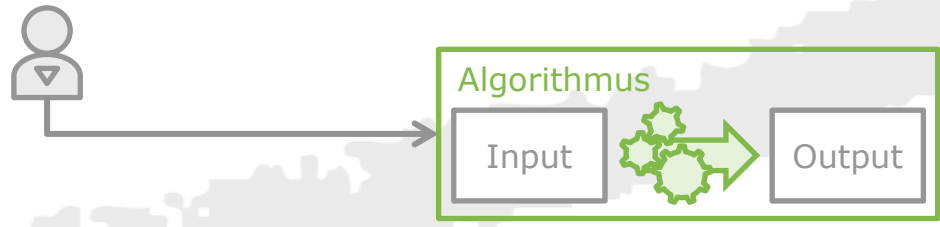


intelligent
aber doch kein ML

intelligent
Machine Learning!

- **Deep Blue:** Schachcomputer von IBM
- Sieg über Weltmeister Kasparov im Jahr 1997
- ➔ **Designer** Algorithmus (Experten und Historie)

- **AlphaGo:** Go-Computer von Google DeepMind
- Sieg über „9-dan“ Lee Sedol im Jahr 2016
- ➔ **Erlerner** Algorithmus (Self-Play-Reinforcement)



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

- Versicherungsunternehmen besitzen **große Datenmengen**, die zahlreiche Informationen z.B. zu Kunden und Schäden enthalten. Data Analytics beinhaltet die intelligente Informationsgewinnung aus solchen Daten und die praktische Umsetzung der daraus gewonnenen Erkenntnisse.
- Wesentliche Prozessschritte sind die **Konkretisierung** der Zielsetzung und Datenanforderung, die **technische Datenanalyse**, die kontextbasierte **Auswertung**, die **Interpretation und Kommunikation** von gewonnenen Erkenntnissen sowie die daraus abgeleitete **Entscheidungsfindung** und **Umsetzung**.
- Neben klassischen Ansätzen der Datenanalyse kommen dabei vermehrt **Methoden des Machine Learnings** zur Anwendung (**Advanced Analytics**).

Bildquelle:
Pixabay

Big Data, künstliche Intelligenz und Data Analytics

Agenda

**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was hat das mit Tarifierung zu tun?**

Klassische Tarifierung mit GLM

Erweiterung um Data-Analytics-Methoden

Weitere Anwendungen in Produktentwicklung und Produktmanagement

Tarifierung

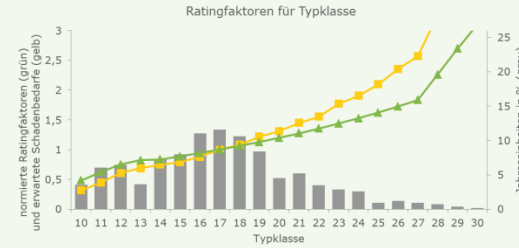
Klassische Tarifierung mit

Überblick über einen typischen Tarifierungsprozess eines Kfz-Versicherers

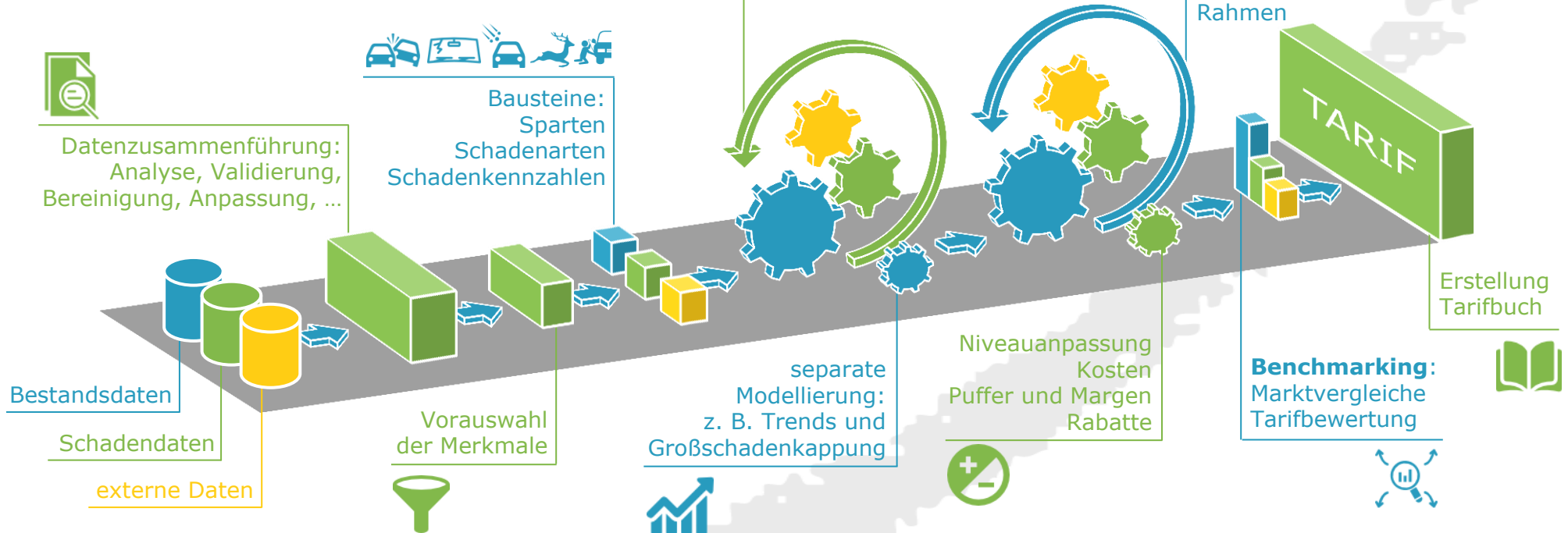


$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



Tarifmodellierung:
Spreizung / Dämpfung
Organik / Beitragssätze
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



- Kern des Tarifierungsprozesses ist die **Risikomodellierung** zur Differenzierung des erwarteten Schadenbedarfs nach individuellen Risiken.
- Typischerweise werden **verallgemeinerte lineare Modelle** (Generalized Linear Models, GLM) verwendet. Sie sind leicht interpretierbar und adjustierbar.

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

Tarifierung

Klassische Tarifierung mit GLM

Erfolgreiche Data-Analytics-Ansätze basieren auf komplexen Machine-Learning-Verfahren: Diese können das klassische GLM ersetzen.

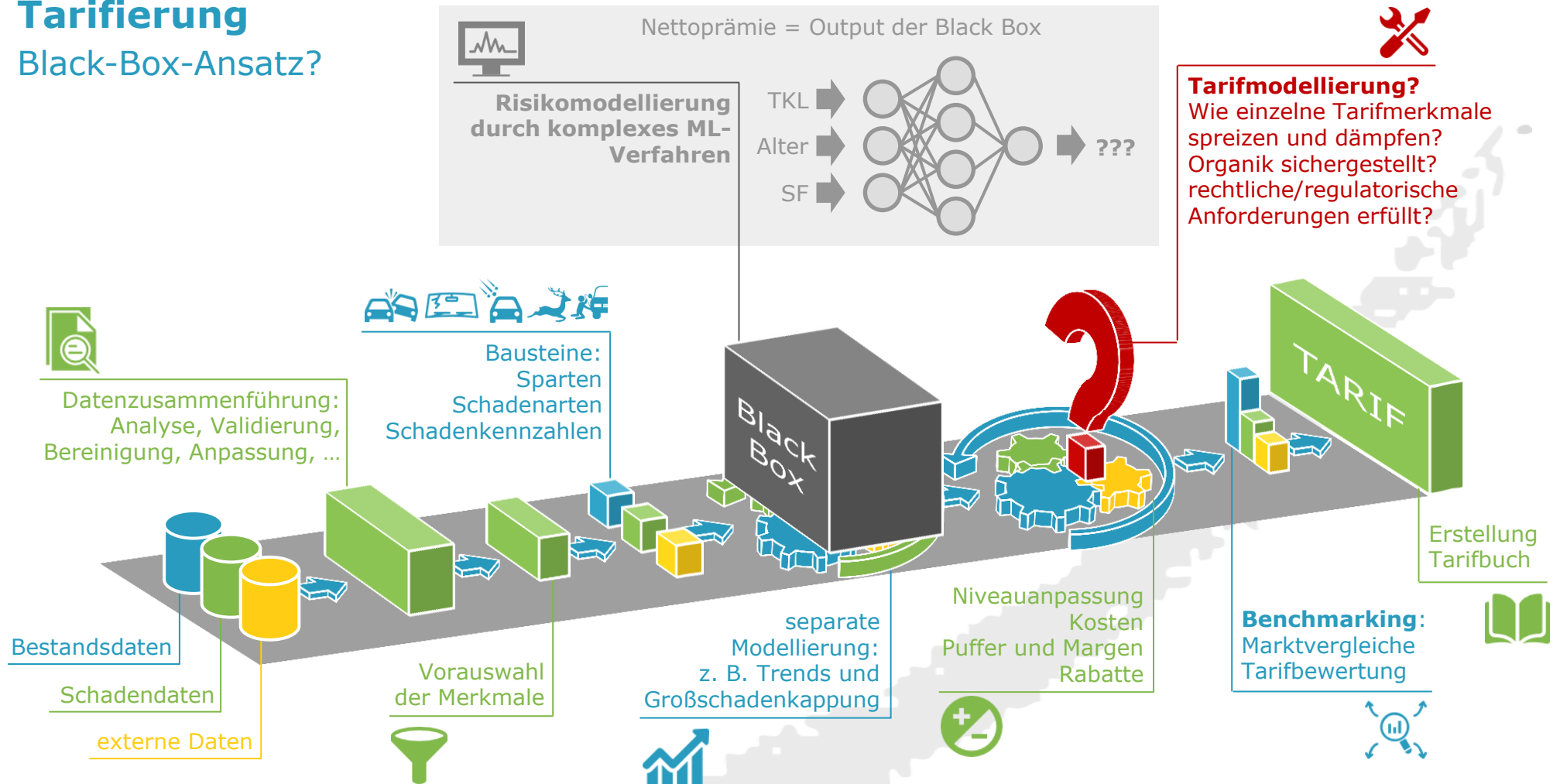
- Beispiel: viele GLMs miteinander verschachteln → neuronale Netze



- Durch Parallel- und Hintereinanderschaltung von GLMs entsteht ein **neuronales Netz**, das beliebig umfangreich werden kann.
- Theoretisch lassen sich beliebige Funktionen erlernen. Durch die Verschachtelung der GLMs ist der Einfluss der Tarifmerkmale aber **nicht allgemein interpretierbar**.

Tarifierung

Black-Box-Ansatz?



- Wird die **Risikomodellierung** durch ein Black-Box-Verfahren ersetzt, ist nicht mehr klar, wie ein einzelner Preis zustande kommt.
- Es besteht Gefahr von **Strukturbrüchen** und **unerwünschten Seiteneffekten**.

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

Big Data, künstliche Intelligenz und Data Analytics

Agenda

**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was hat das mit Tarifierung zu tun?**

Klassische Tarifierung mit GLM

Erweiterung um Data-Analytics-Methoden

Weitere Anwendungen in Produktentwicklung und Produktmanagement

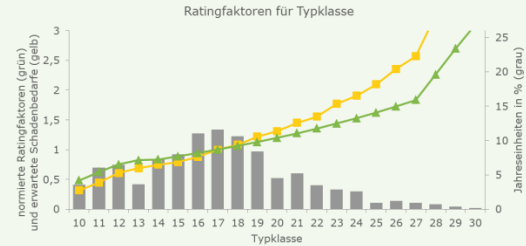
Tarifierung

I. Big Data modellieren



$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



Tarifmodellierung:
Spreizung/Dämpfung
Organik
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



Datenzusammenführung:
Analyse, Validierung, Bereinigung, Anpassung, ...



Bausteine:
Sparten
Schadenarten
Schadenkennzahlen



separate Modellierung:
z. B. Trends und Großschadenkappung

Niveaueinpassung
Kosten
Puffer und Margen
Rabatte

Benchmarking:
Marktvergleiche
Tarifbewertung

Erstellung
Tarifbuch



Bestandsdaten

Schadendaten

externe Daten

Big Data (z.B. Telematik)

- Herleitung **neuer Merkmale** für das GLM („Scores“)
- **Verfeinerung** vorhandener Merkmale (z.B. eigene Regionalklassen, PLZ-Faktoren)

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

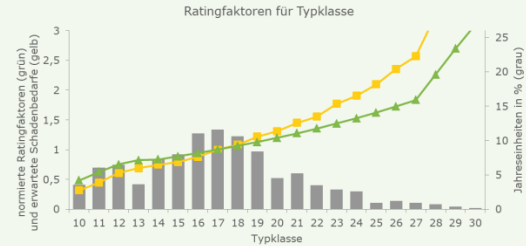
Tarifierung

II. Daten vorbereiten



$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



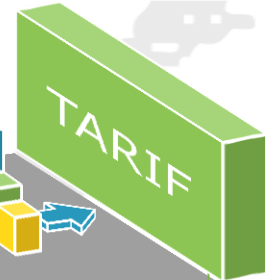
Tarifmodellierung:
Spreizung/Dämpfung
Organik
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



Datenzusammenführung:
Analyse, Validierung, Bereinigung, Anpassung, ...



Bausteine:
Sparten
Schadenarten
Schadenkennzahlen



Erstellung Tarifbuch



Benchmarking:
Marktvergleiche
Tarifbewertung

Bestandsdaten

Schadendaten

externe Daten

Vorteil der Merkmale



separate Modellierung:
z. B. Trends und
Großschadenkappung

Niveaueinpassung
Kosten
Puffer und Margen
Rabatte



■ datengetriebene **Merkmalsauswahl und -verdichtung** zur Verbesserung und Vereinfachung der anschließenden GLM-Modellierung

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

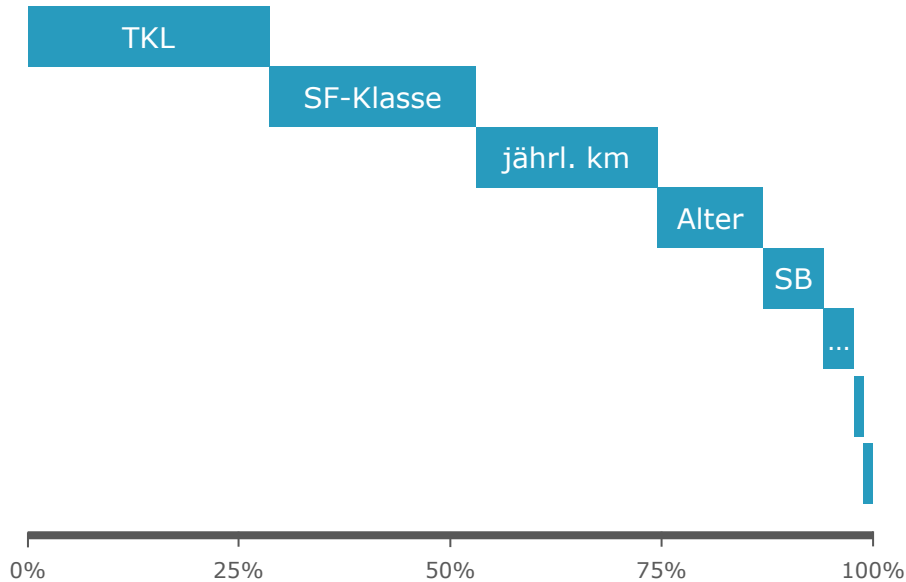
Tarifierung

II. Daten vorbereiten – Beispiel: Merkmalsauswahl und -verdichtung

illustrativ

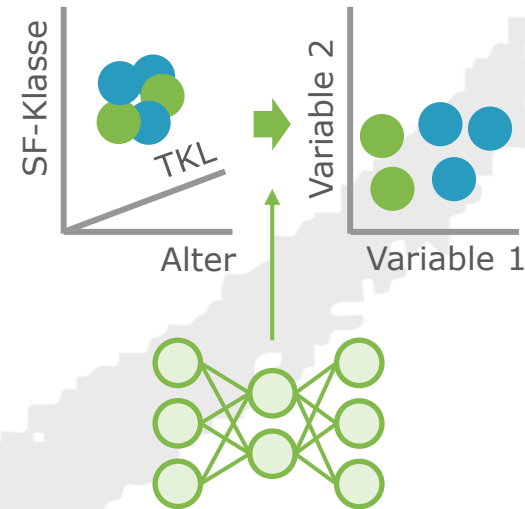
Merkmalsauswahl

- anhand der „Variable Importance“ der komplexen Machine-Learning-Modelle



Verdichtung der Merkmale

- Dimensionsreduktion mittels Hauptkomponentenanalyse oder Autoencoder



■ Mit Data Analytics lassen sich große und vielfältige **Datenmengen selektieren und zusammenfassen**, um die GLM-Modellierung zu verbessern und zu vereinfachen.

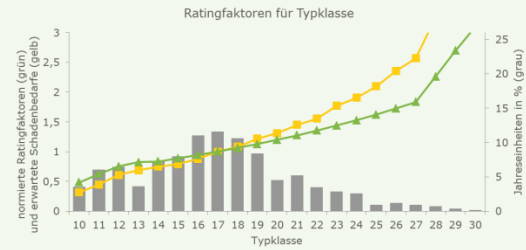
Tarifierung

III. GLMs unterstützen



$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



Tarifmodellierung:
Spreizung/Dämpfung
Organik
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



Datenzusammenführung:
Analyse, Validierung, Bereinigung, Anpassung, ...



Bausteine:
Sparten
Schadenarten
Schadenkennzahlen

Bestandsdaten

Schadendaten

externe Daten

Vorauswahl der Merkmale



Black Box:
z. B. Trends und Großschadenkappung



Niveaueinstellung:
Kosten
Puffer und Margen
Rabatte



Benchmarking:
Marktvergleiche
Tarifbewertung



Erstellung Tarifbuch



- Identifikation **nicht-linearer Effekte** oder **Interaktionen** in komplexen Machine-Learning-Modellen
- anschließend explizite Berücksichtigung im GLM

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

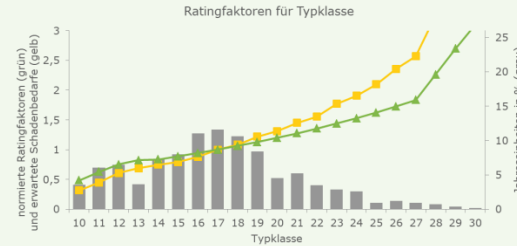
Tarifierung

IV. GLMs benchmarken



$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



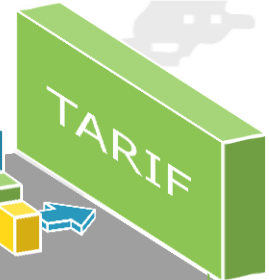
Tarifmodellierung:
Spreizung/Dämpfung
Organik
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



Datenzusammenführung:
Analyse, Validierung, Bereinigung, Anpassung, ...



Bausteine:
Sparten
Schadenarten
Schadenkennzahlen



Erstellung
Tarifbuch

Benchmarking:
Marktvergleiche
Tarifbewertung



Bestandsdaten

Schadendaten

externe Daten

Vorauswahl
der Merkmale



separate
Modellierung:
z. B. Trends und
Großschadenkappung



Kosten
Puffer und Margen
Rabatte



- separate Modellierung mittels komplexem Machine-Learning-Verfahren
- Abweichungsanalyse gegenüber dem GLM zur Erkennung defizitärer Teilbestände/Segmente, der Stabilität der Schadenbedarfsschätzung und möglichen Verbesserungspotenzials des GLMs

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

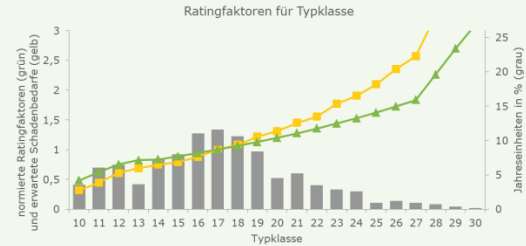
Tarifierung

V. White-Box-Ansatz



$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



Tarifmodellierung:
Spreizung/Dämpfung
Organik
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



Datenzusammenführung:
Analyse, Validierung, Bereinigung, Anpassung, ...



Bausteine:
Sparten
Schadenarten
Schadenkennzahlen

Bestandsdaten

Schadendaten

externe Daten

Vorauswahl der Merkmale



separate Modellierung:
z. B. Trends und Großschadenkappung

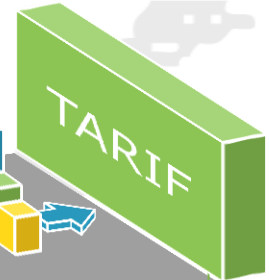
Niveaueinstellung
Kosten
Puffer und Margen
Rabatte



Benchmarking:
Marktvergleiche
Tarifbewertung



Erstellung
Tarifbuch



- Verbesserung des GLM-Frameworks mit White-Box- statt Black-Box-Modellen
- Ein GLM lässt sich um datengetriebene Merkmalsauswahl, Faktorendämpfung oder Erkennung nicht-linearer Zusammenhänge erweitern. Es stellt somit ein ML-Verfahren dar, ohne seine interpretierbare Grundstruktur zu verlieren.

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

Big Data, künstliche Intelligenz und Data Analytics

Agenda

**Big Data, künstliche Intelligenz, Machine Learning, Data Analytics & Co.:
Was hat das mit Tarifierung zu tun?**

Klassische Tarifierung mit GLM

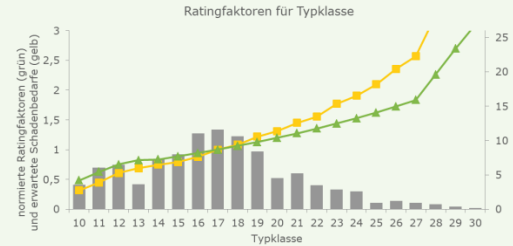
Erweiterung um Data-Analytics-Methoden

Weitere Anwendungen in Produktentwicklung und Produktmanagement

Weitere Anwendungen

$$\text{Nettoprämie} = \text{Basisprämie} * \text{Ratingfaktoren}$$

Risikomodellierung:
Anpassung multivariater Modelle (i.d.R. GLMs) pro Baustein als iterativer Prozess



Tarifmodellierung:
Spreizung/Dämpfung
Organik
Marktpositionierung
Quersubventionierung
Zielgruppengewinnung
rechtlicher/regulatorischer Rahmen



Datenzusammenführung:
Analyse, Validierung, Bereinigung, Anpassung, ...



Bausteine:
Sparten
Schadenarten
Schadenkennzahlen

Bestandsdaten

Schadendaten

externe Daten

Vorauswahl der Merkmale

separate Modellierung:
z. B. Trends und Großschadenkappung

Niveaueinstellung
Kosten
Puffer und Margen
Rabatte

Benchmarking:
Marktvergleiche
Tarifbewertung

Erstellung
Tarifbuch



- Data Analytics bietet sich zusätzlich zur verbesserten Modellierung u.a. von Wettbewerberpreisen („Reverse Engineering“), Preiselastizitäten, Conversion Rates, Rabattvergabe, Kundenwert und Stornoquoten an.
- Die Ergebnisse lassen sich ins Pricing integrieren und für die Preispolitik nutzen.

Quelle der Icons:
All icons made by Freepik from
www.flaticon.com.

Weitere Anwendungen

- **Neue Policen gewinnen:** profitables Cross- und Upselling im eigenen Bestand
 - Identifikation von Kunden, die bereit sind einen weiteren Vertrag abzuschließen oder einen bestehenden Vertrag zu erhöhen
 - passende Angebote durch Handlungsempfehlungen im Vertrieb und automatisierte Recommender-Systeme
 - auf Basis interner Bestandsdaten (Assoziationen) oder externer Daten (Identifikation von verändertem Bedarf)
- **Den Kunden halten:** Stornoprophylaxe zur Bestandssicherung
 - Identifikation von Verträgen oder Vertragskonstellationen mit erhöhter Stornowahrscheinlichkeit
 - aktive Kundenansprache in Verbindung mit einem Kundenwertmodell und Frühwarnsystemen

- **Den Vertriebspartner kennen:** Vertriebscontrolling und -kanaloptimierung
 - Vermittler nach Gütekriterien wie Storno und Kundenwert des Geschäfts clustern
 - mögliche Maßnahmen zur Incentivierung
 - Identifikation von Betrug oder schwarzen Schafen mit schlechtem Geschäft
 - Frühwarnsystem bzgl. atypischem Geschäft einzelner Vermittler
- **Den Kunden kennen:** verbesserte Annahmeregeln und Risikoprüfung
 - durch Erkenntnisse aus historischen, sich langfristig realisierenden Leistungsdaten
 - langfristiges Tarifwechsel-/Stornoverhalten bei ähnlichen Anträgen
 - dynamische Risikoprüfung durch intelligente Steuerung der Fragen

Erfolgsfaktoren und Fazit



Für die erfolgreiche Anwendung von Data-Analytics-Methoden müssen

- die notwendigen Datengrundlagen geschaffen,
- die erforderlichen Skills und Erfahrungen erworben und
- das entsprechende Kontextwissen genutzt werden.

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.