Data-Analytics-Verfahren in der PKV

Use Case Stornoprävention

Kinga Böhm, Eugenia Jung, Ella Russer (Hallesche Krankenversicherung a.G.)

Dr. Johannes Schupp (ifa)

11. Mai 2021

Agenda

Einführung

Daten in der PKV

Data-Analytics-Verfahren zur Stornoprognose

Modelltuning und Optimierung

Ergebnis und Anwendung

Fazit & Lessons Learned

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Was ist eigentlich Data Analytics?

- "… Verfahren und Technologien, die interessante Muster in umfangreichen Datenbeständen aufdecken und Prognosen über zukünftige Ereignisse und Gegebenheiten anstellen können…"
  (Quelle: Business Analytics – Grundlagen, Methoden und Einsatzpotenziale von Prof. Dr. Peter Gluchowski, TU Chemnitz, Springer Fachmedien Wiesbaden 2016)

- Ein Vorgehen um verschiedene Datenquellen zu extrahieren und zu untersuchen, und darin Muster und Strukturen zu finden.
- Ziel: Schlussfolgerungen/Handlungen aus den Daten ableiten
- Daten kommen aus unterschiedlichen Quellen und sind grundsätzlich in der strukturierten Form vorhanden

- Wer macht was?
  - Data Engineers, Data Analysts, Data Scientists
  - Know-how ifa

Projektzusammensetzung

- Use Case: Erstellung eines Prognosemodells für die Stornowahrscheinlichkeit jedes einzelnen Versicherungsnehmers, um mit gezielten Maßnahmen eine höhere Bestandsbindung in der Vollversicherung zu erreichen

- Projektteam:
- Projektablauf: zweites Halbjahr 2020
- Interdisziplinär
  - verschiedene Expertisen und Schwerpunkte
  - Erhöhung der Akzeptanz und des Verständnisses von Data-Analytics-Verfahren

Agenda

Einführung

Daten in der PKV

Data-Analytics-Verfahren zur Stornoprognose

Modelltuning und Optimierung

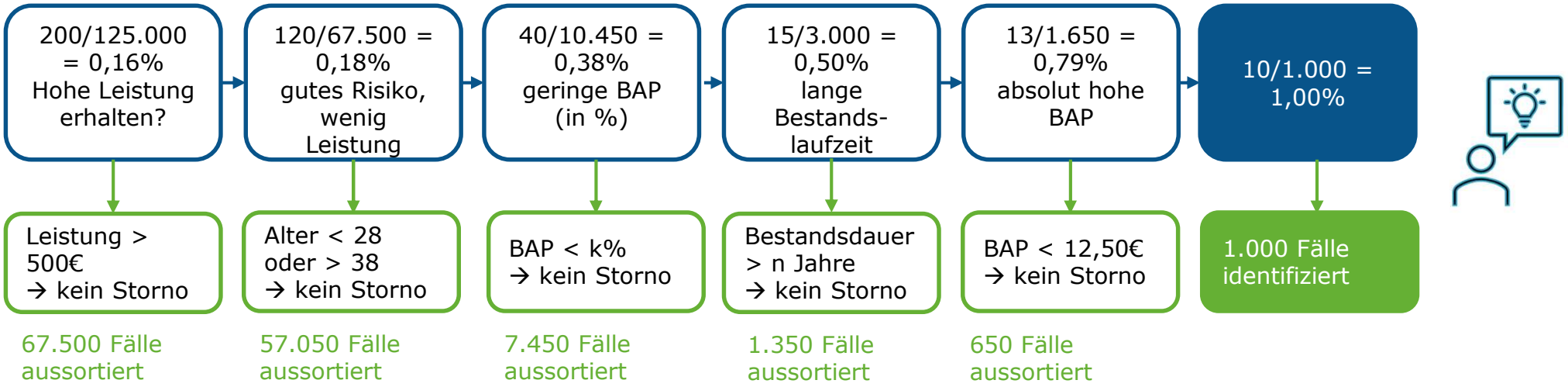
Ergebnis und Anwendung

Fazit & Lessons Learned

Fachexperte

Es ist eine Kundenbindungsaktion mit einem Budget für 1.000 Kunden geplant. Auftrag für Fachexperte: „Bitte selektieren Sie 1.000 Fälle aus 125.000, die stark stornogefährdet sein könnten!“

- „Sukzessives Filtern in einer Bestands-Excel Tabelle“



DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Erste Präzisierung des Stornobegriffs für den Use Case

- Konzentration auf verhinderbares Storno: Abgang in PKV (Storno zum 01.01.2021)
- Personen mit einer Krankheitskostenvollversicherung. Folgende Tarife / Personengruppen werden dabei nicht berücksichtigt:
  - Beihilfe-Tarife, Mediziner
  - Ausbildungstarife, große Anwartschaften
- Es werden nur Personen ab Alter 21 betrachtet.
- Für die Vorhersage der Abgänge in 2021 liegen uns Daten für die Jahre 2016-2020 vor.
  - Die dabei beobachtete Stornoquote betrug x%.

Ampel

- systematische Anzeige eines Scorewertes
- aufgrund geringer Stornozahl insbesondere möglichst gute „rote Ampel" wichtig
- Fachexperte: 1%

Use Case

Die geringe Anzahl der Stornofälle ist sehr gut für die Hallesche, aber schlecht für Data Analytics.

DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Welche Daten spielen eine erklärende Rolle?

versicherte Person

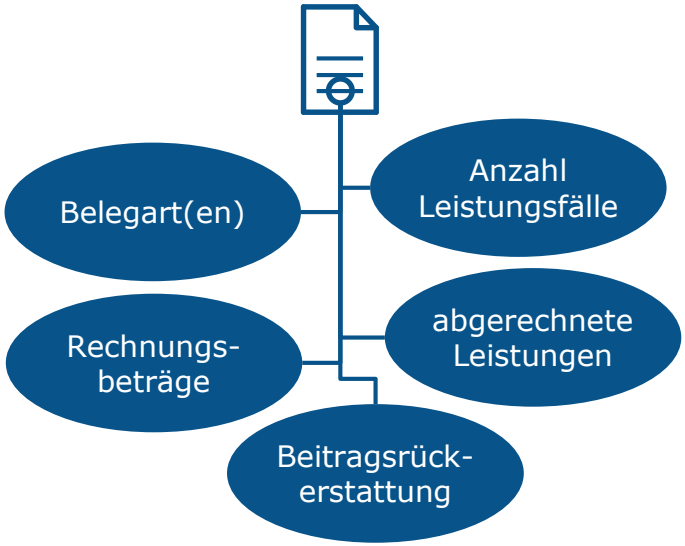
z.B. in der privaten Krankenvollversicherung

Person
Alter
Geschlecht
Wohnort
…

Tarif
Vertragsdauer
Selbstbehalt
Premium/Basic
…

eingereichte Belege

z.B. Leistungshistorie



Vertrieb

z.B. Makler, Ausschließlichkeitsorganisation

ABER:

- Ein Data-Analytics-Verfahren kann nicht sämtliche Daten eines Krankenversicherers verarbeiten.
- **Es gilt nicht**: mehr Daten = bessere Prognose! (Zufallsrauschen stört Kalibrierung robuster Modelle)
- Es ist eine sinnvolle, aber hinreichend große Auswahl zu treffen (sog. **prädiktive Merkmale**).

ifa

Datengrundlage

- Stornostatistik des PKV-Verbands
- Datenquellen für weitere Merkmale: Dispositiver Bestand, operativer Bestand, Statistikdatenbank, Datenbank zur Beitragsrückerstattung, Statistisches Bundesamt
- Nicht alle Daten stehen zur Verfügung (Gründe: nicht erfasst, veraltet, technisch nicht möglich, sehr hoher Aufwand).
- Umfangreiche Ermittlung der Datengrundlage:
  - 814.000 Datensätze und 130 Spalten zu Kunden- und Vertragsmerkmalen

  → 106 Mio. Datenpunkte
- Erstellung eines Merkmals kann bis zu mehreren Tagen dauern.
- Merkmale müssen teilweise überarbeitet / korrigiert werden.
- unklar, ob jedes Merkmal Bedeutung für unsere Analyse hat
- Bevor man mit der Analyse der Daten und der Modellierung beginnt, sollte man die Daten hinsichtlich der Struktur, Fehler und Merkmalskodierung untersuchen.
  - Fehlende Werte mit „NA" umschlüsseln (mit geeigneten Werten wie etwa Mittelwert ersetzen)
  - Merkmale mit hoher Zahl an Ausprägungen gruppieren.
  - Merkmale umkodieren (Einführung der Dummyvariablen)

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Datenanalyse mit Pandas

Korrelationsmatrix

```
corr_matrix ["label"].sort_values(ascending=False)
```

```
label              1.000000
WeltenKZ           0.056506
Unisex             0.051563
Haftung            0.022252
RZ_VJ_pro_PTNK     0.013691
                     ...
PKV_Zug           -0.040221
Alter             -0.041987
Bestandsdauer     -0.042535
AR_VJ             -0.042709
AR_bw_VJ          -0.045571
```

	WeltenKZ	Unisex	Haftung	RZ_VJ_pro_PTNK	PKV_Zug	Alter	Bestandsdauer	AR_VJ	AR_bw_VJ
WeltenKZ	1.000000	0.537771	0.433410	-0.023942	-0.620174	-0.445163	-0.599777	-0.593226	-0.629292
Unisex	0.537771	1.000000	0.062265	-0.009830	-0.442349	-0.347382	-0.402683	-0.384932	-0.409481
Haftung	0.433410	0.062265	1.000000	-0.010174	-0.273199	-0.175374	-0.264329	-0.258029	-0.274467
RZ_VJ_pro_PTNK	-0.023942	-0.009830	-0.010174	1.000000	0.055876	0.131166	0.109650	0.093306	0.096749
PKV_Zug	-0.620174	-0.442349	-0.273199	0.055876	1.000000	0.382007	0.893763	0.676629	0.671671
Alter	-0.445163	-0.347382	-0.175374	0.131166	0.382007	1.000000	0.696896	0.616456	0.648273
Bestandsdauer	-0.599777	-0.402683	-0.264329	0.109650	0.893763	0.696896	1.000000	0.781711	0.783951
AR_VJ	-0.593226	-0.384932	-0.258029	0.093306	0.676629	0.616456	0.781711	1.000000	0.969849
AR_bw_VJ	-0.629292	-0.409481	-0.274467	0.096749	0.671671	0.648273	0.783951	0.969849	1.000000

Merkmale mit hoher Korrelation nicht mehrfach verwenden!

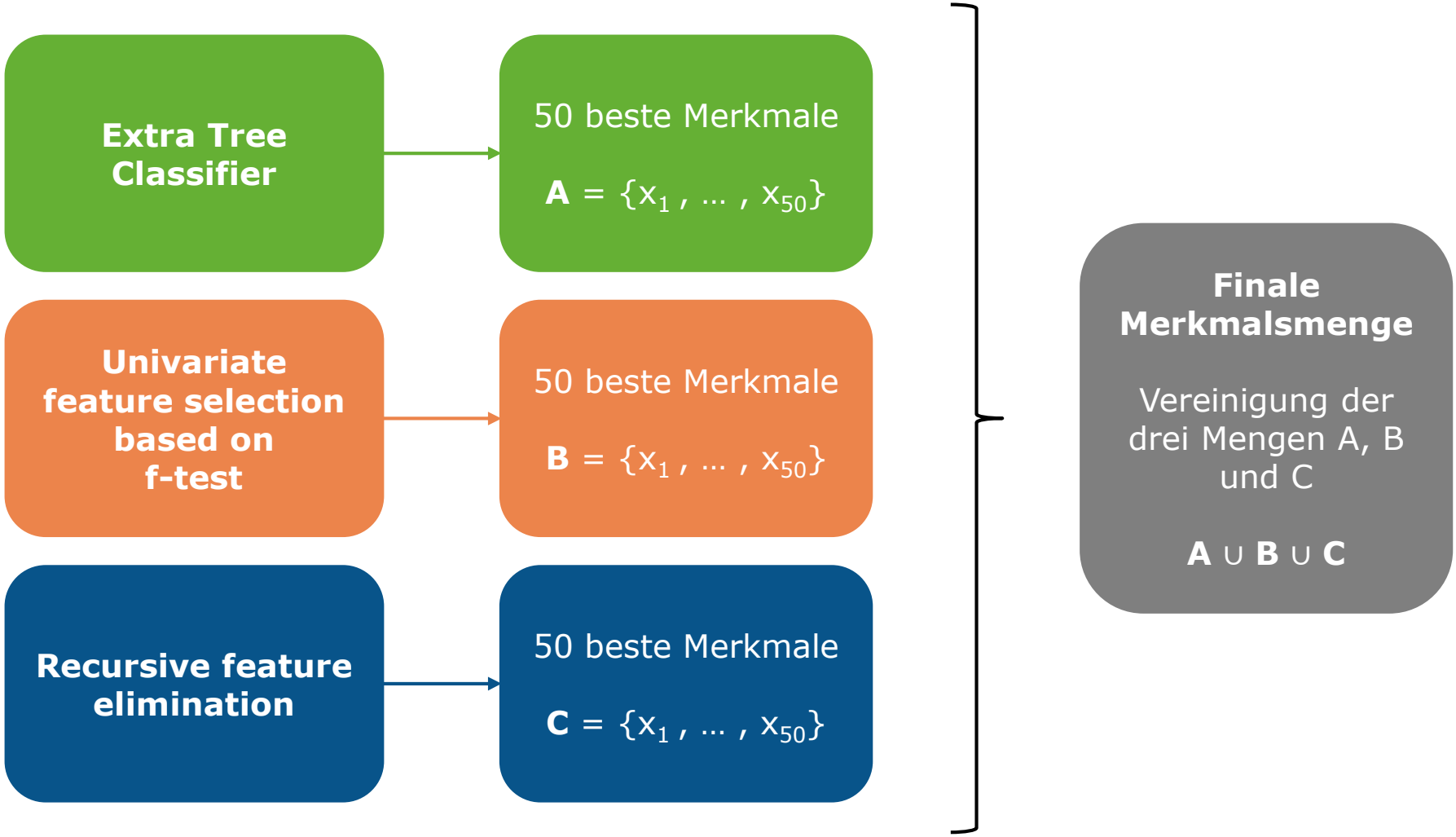
Warum Feature Selection?

- Fortschreitende Technologien ermöglichen es, Daten im großen Stil zu erfassen und zu speichern.
- Große Datenmengen in Form von Quantität, Anzahl der Daten als auch in der Anzahl der charakteristischen Merkmalen je Beobachtung ➔ Daten, die im hochdimensionalen Raum einzelne Punkte abbilden.
- Datenanalysewerkzeuge ermöglichen die Verarbeitung und Reduktion der Dimension.
- Aus M Datenmerkmalen wird eine Teilmenge von m ≤ M Merkmalen selektiert.
  - Ziel: Teilmenge enthält Merkmale, die einen Zusammenhang zur Zielvariablen haben

Vorteile	Nachteile
• Schnellere Bearbeitungsperformance • Leichteres Datenverständnis auf die wesentlichen Merkmale • Irrelevante Merkmale aussortieren	• Mögliche Kollinearität • Überanpassungen • Inkonsistenz

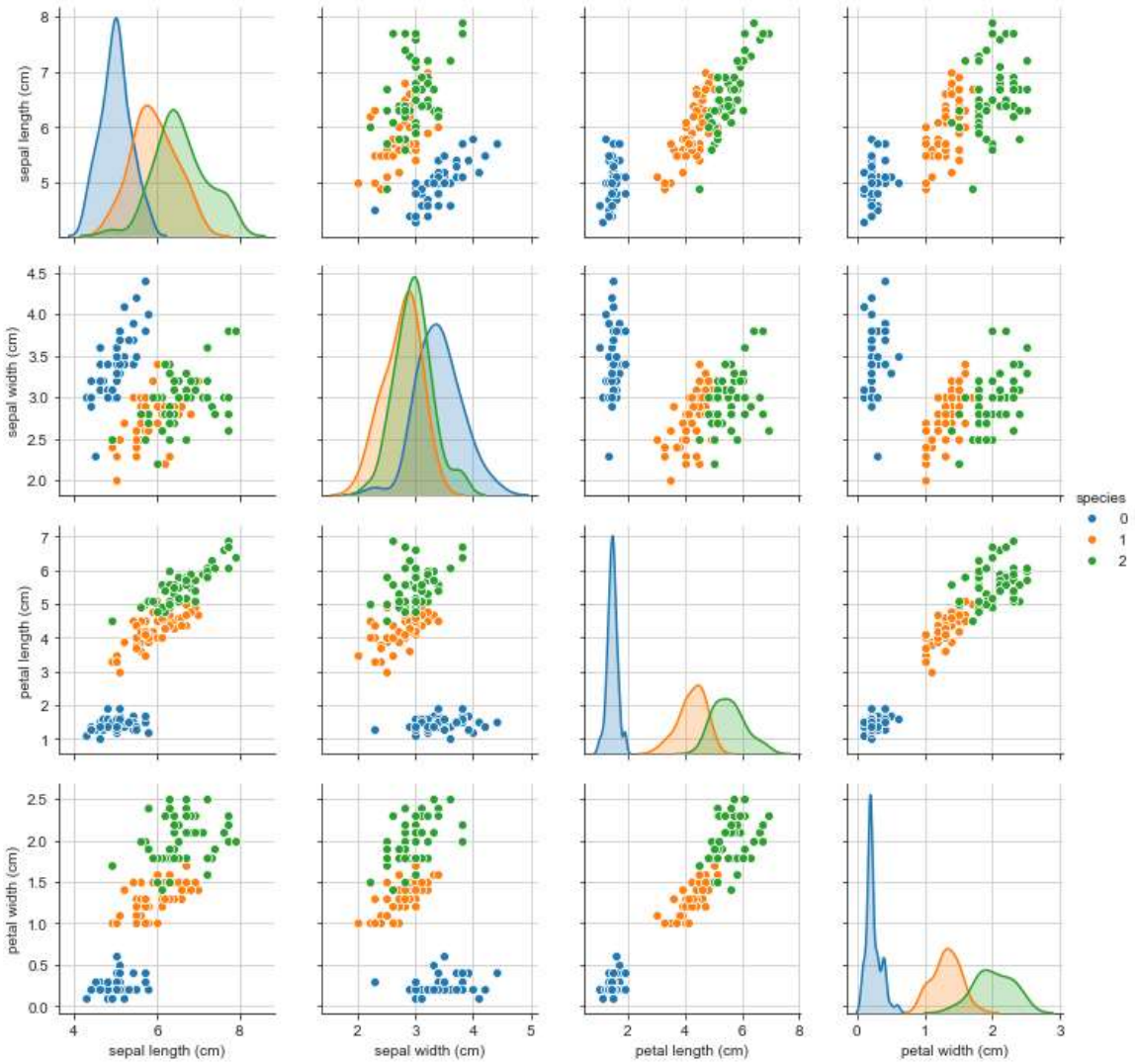
DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Durchführung



Datenvisualisierung

Textbook example seaborn pairplot – Iris

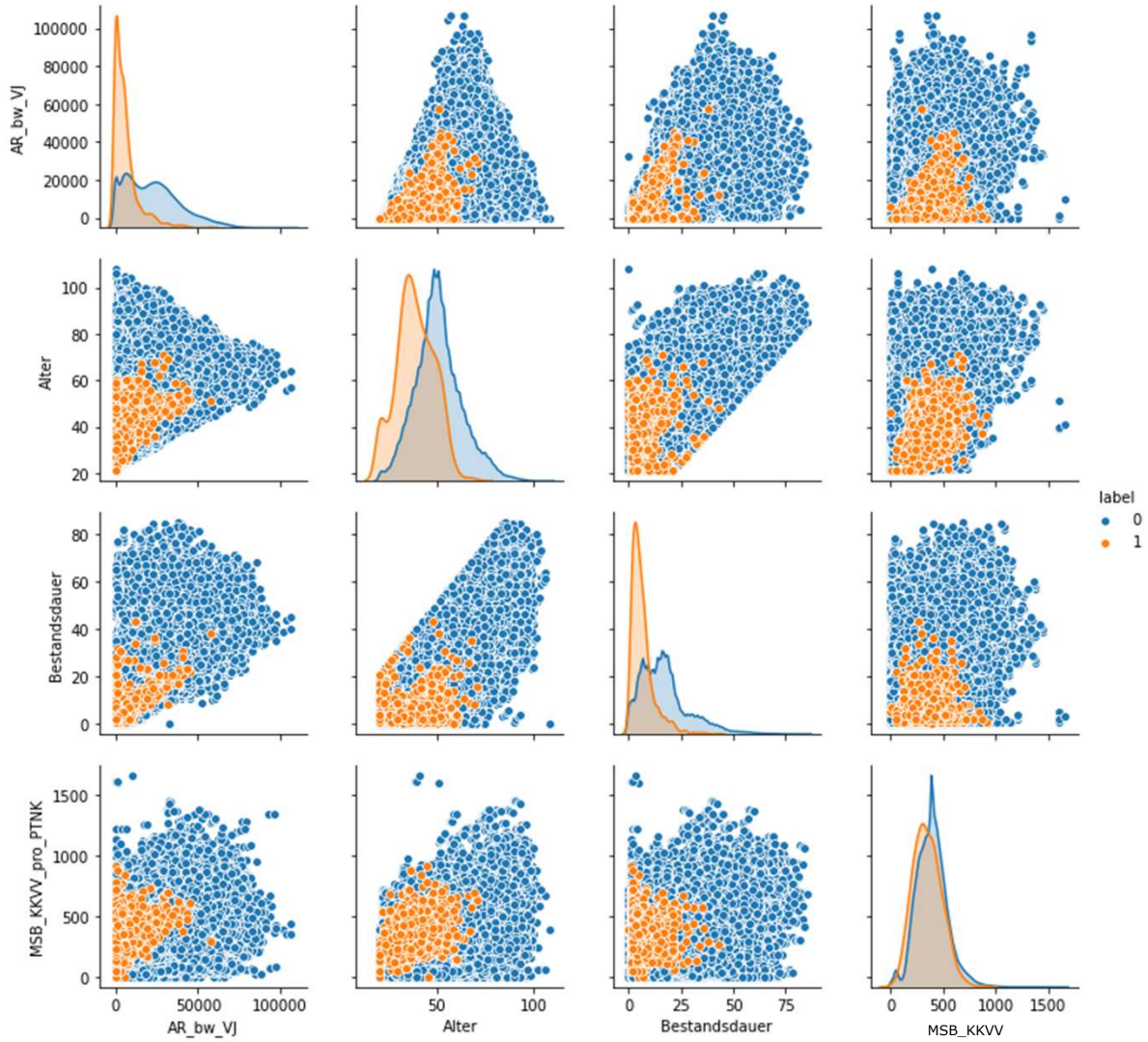


- Drei Arten können bereits anhand der Visualisierung differenziert werden.
- Eigentlich benötigt man mit dieser Erkenntnis keine komplexen Methoden mehr.
  - Plausibilisierung ob das Data-Analytics-Verfahren die Muster erkennt, die das Auge sowieso erkannt hätte!

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Datenvisualisierung

PKV example seaborn pairplot – Storno



- sehr viel mehr Datenpunkte
  - weniger übersichtlich
- trotzdem scheint ein gewisses Muster vorzulegen

Können Data-Analytics-Verfahren diese Muster besser erkennen als unsere Experten?

ifa

Agenda

Einführung

Daten in der PKV

Data-Analytics-Verfahren zur Stornoprognose

Modelltuning und Optimierung

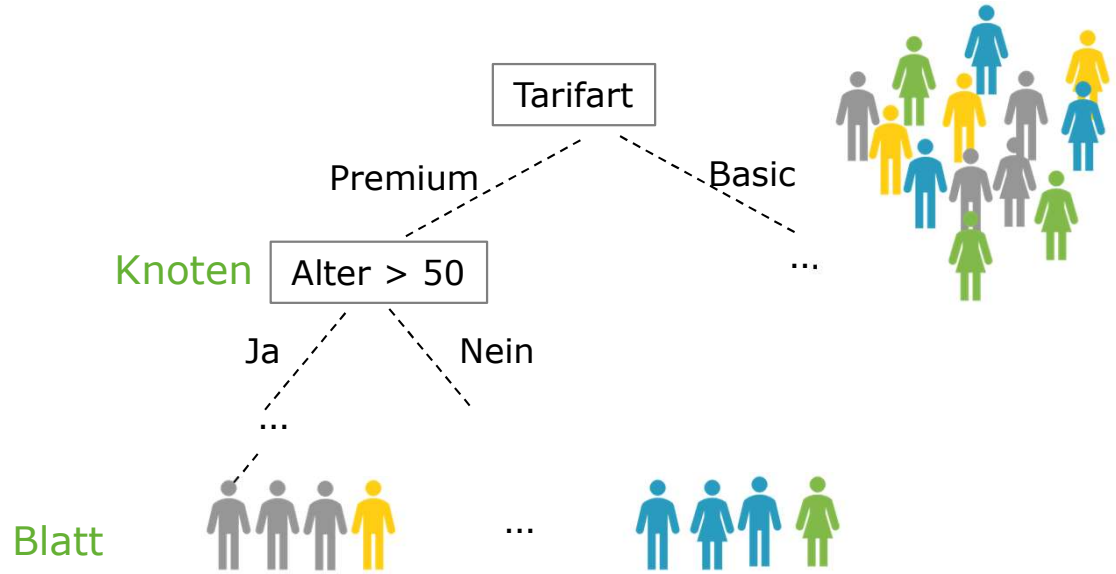
Ergebnis und Anwendung

Fazit & Lessons Learned

Einführung in Data Analytics

Ein einfaches Beispiel: Bestandsanalyse

- Entscheidungsbäume unterteilen den Bestand gemäß der Merkmale in Gruppen
- Beispiel: unterschiedliche Leistungskosten der Verträge um Erkenntnisse für das aktuarielle Risikocontrolling zu liefern.



- **Der Algorithmus ermittelt, welche Merkmale in die Klassifizierung einfließen sollten, und zwar automatisiert und datengetrieben.**
  - Ziel des Algorithmus (vereinfacht): möglichst einheitliche Blätter erzeugen.

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Entscheidungsbäume

Wie kommt so ein Baum zustande?

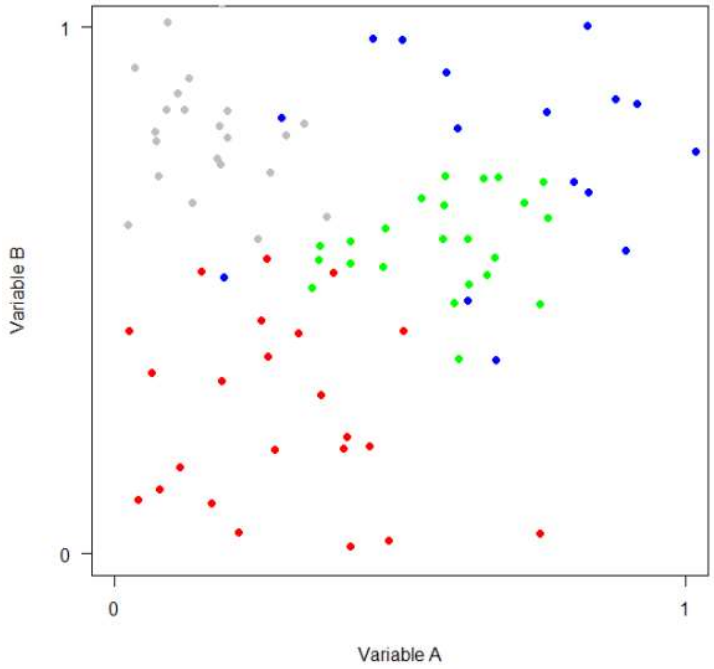
Konstruktion eines Klassifikationsbaums

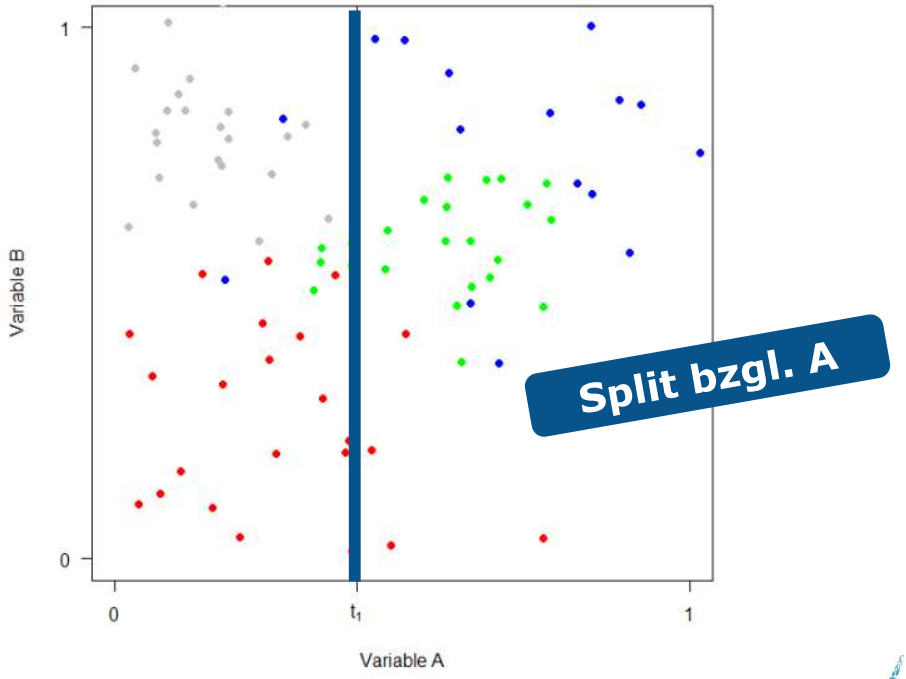
- Zielgröße: z.B. **vier Kategorien** (sehr hohe Kosten, hohe Kosten, durchschnittliche Kosten, geringe Kosten)
- Attribute (A, B) zwischen 0 und 1 (skaliert), z. B. Alter, Tarifart, usw.
- Frage im **Risikocontrolling**: Für welche Kunden sind höhere Kosten zu erwarten?

Vorgehensweise hier:

1. In welchem bereits bestehenden Ast?
2. Welches Attribut (A oder B)?
3. Welcher Schwellenwert (0 bis 1)?

Ziel: möglichst effektive Trennung nach Kategorie



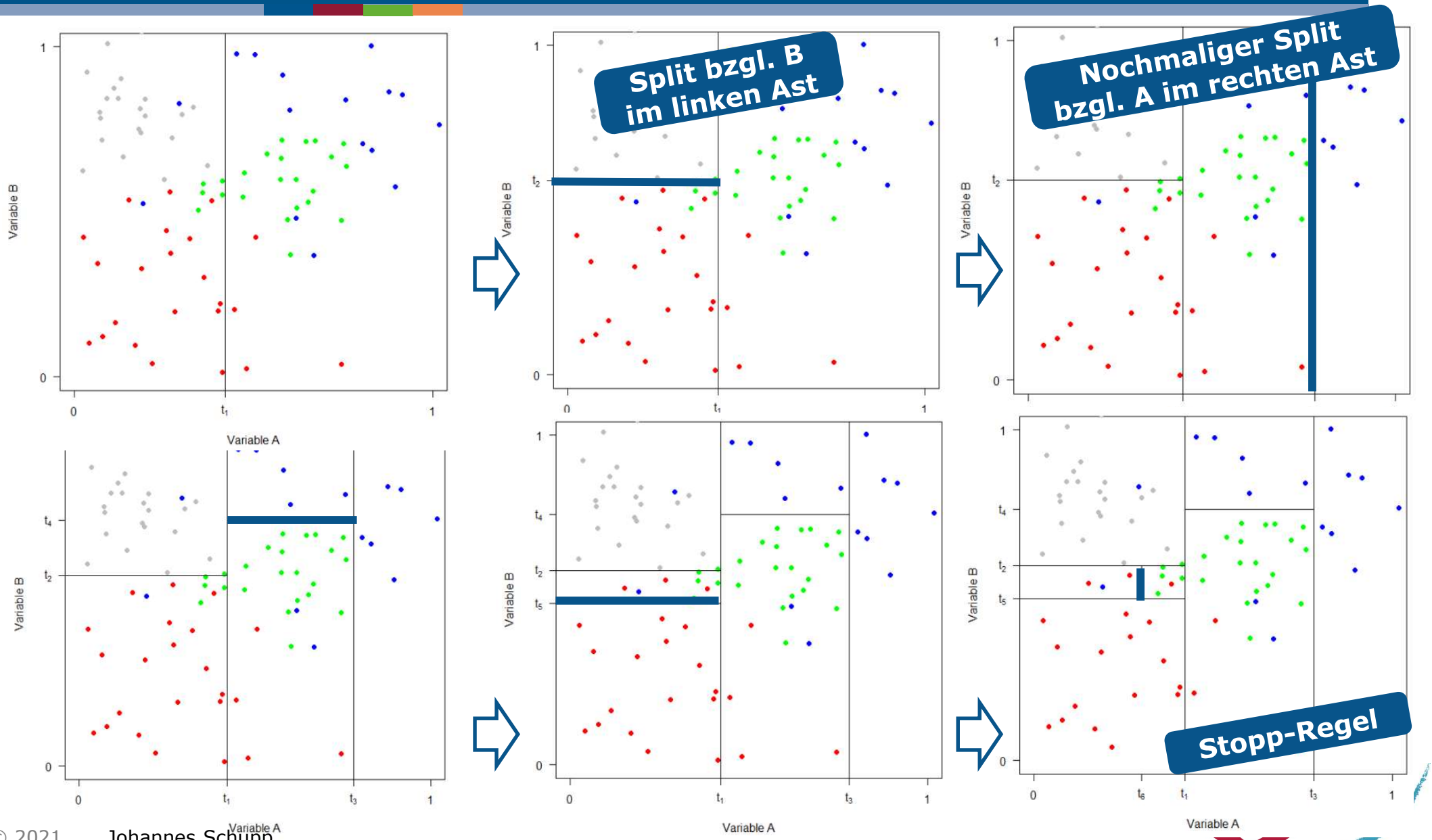


ifa

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Entscheidungsbäume

Der Baum entsteht

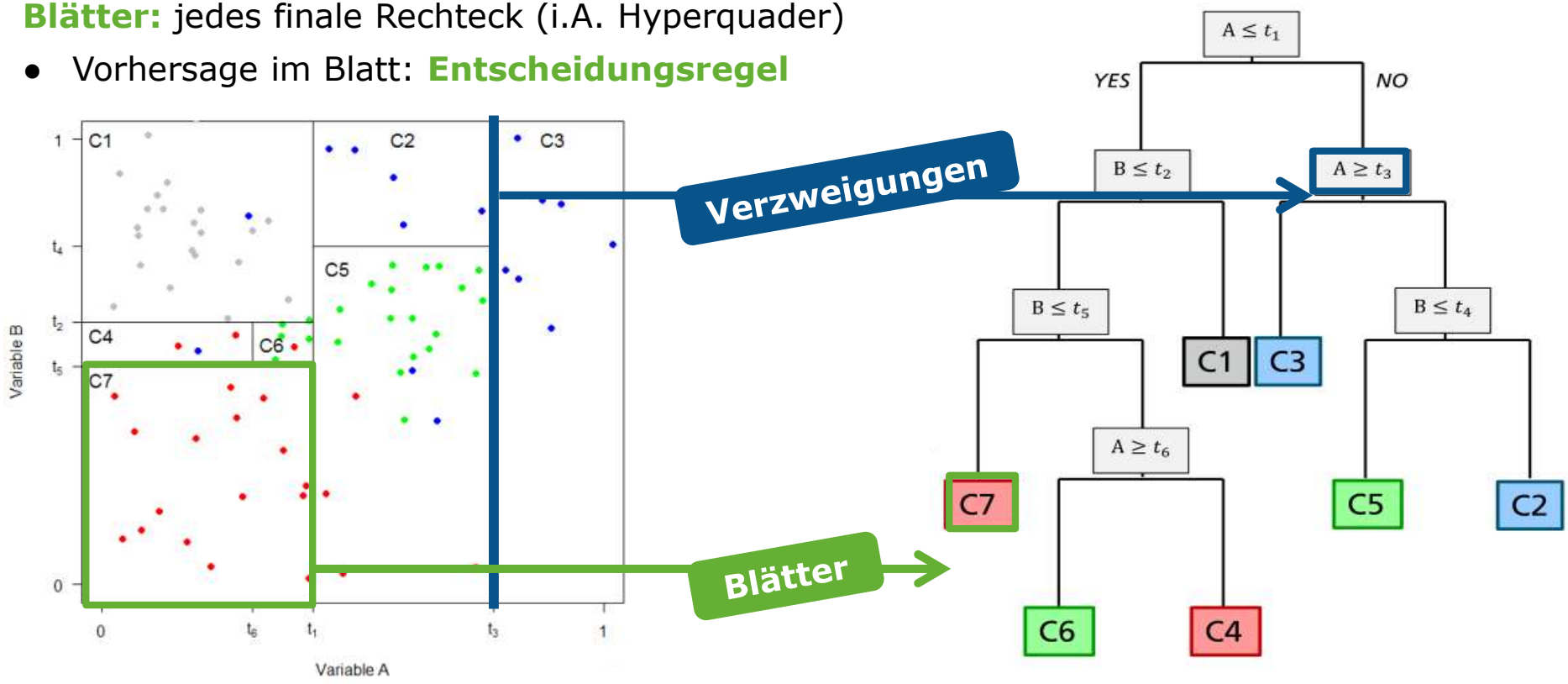


Entscheidungsbäume

Darstellung des Baums

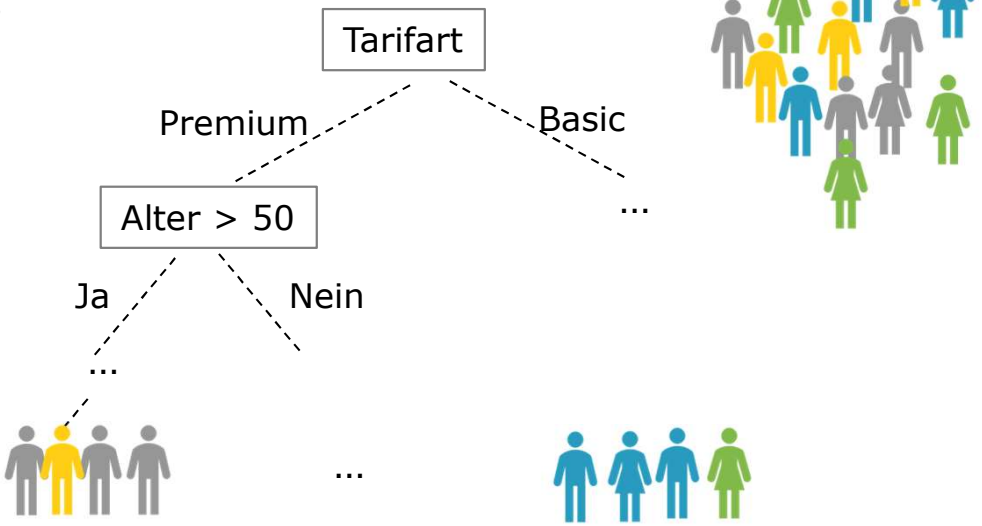
Formulierung des finalen Klassifikationsbaums

- Exakte Formulierung als Baum möglich
  - **Verzweigungen:** jede eingezogene Linie (i.A. Hyperebene)
    - Entscheidungsregel im Knoten: ausgewähltes Merkmal, ausgewählter Schwellenwert
  - **Blätter:** jedes finale Rechteck (i.A. Hyperquader)
    - Vorhersage im Blatt: **Entscheidungsregel**



Entscheidungsbäume
Vorteile

- Entscheidungsbäume ermöglichen eine **interpretierbare** Modellprognose
- Zahlreiche Anwendungen in allen Sparten, z.B.
  - vgl. aktuarielles Risikocontrolling in allen Sparten (→ s.a. ifa informiert live)
  - Analyse und Optimierung des Churn Managements
  - Kapitalwahl in der Leben
  - Tarifwechsel in der Kranken
  - Cross- und Upselling
  - vgl. vorliegender Use Case in der PKV



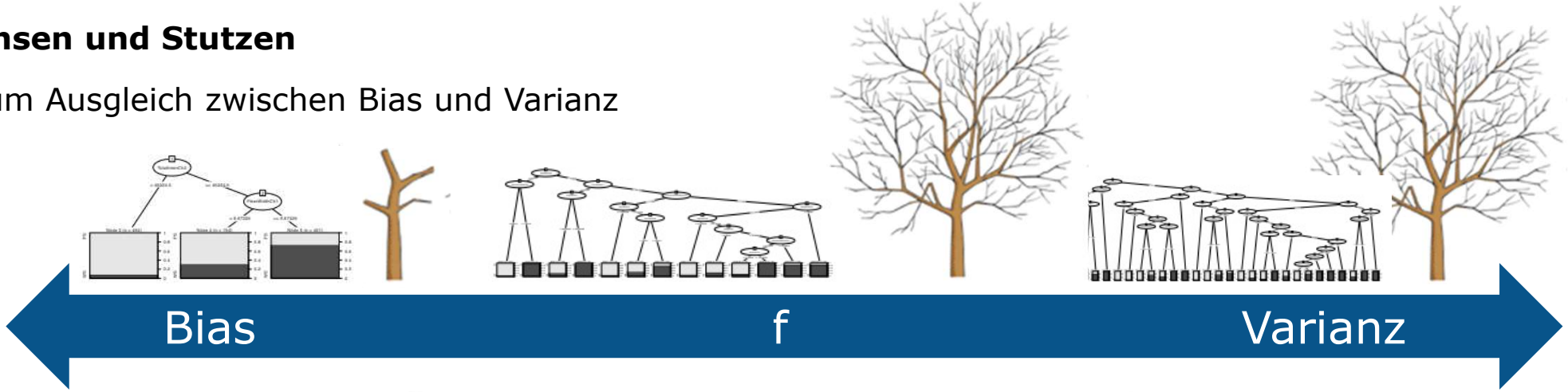
Modell	Vorteile	Nachteile	Einsatz
Entscheidungsbaum	• Interpretierbar • Flexibel • Robust	• Restriktion bei Mustererkennung • anfällig gegen Overfitting	Einfache Regeln im Fokus

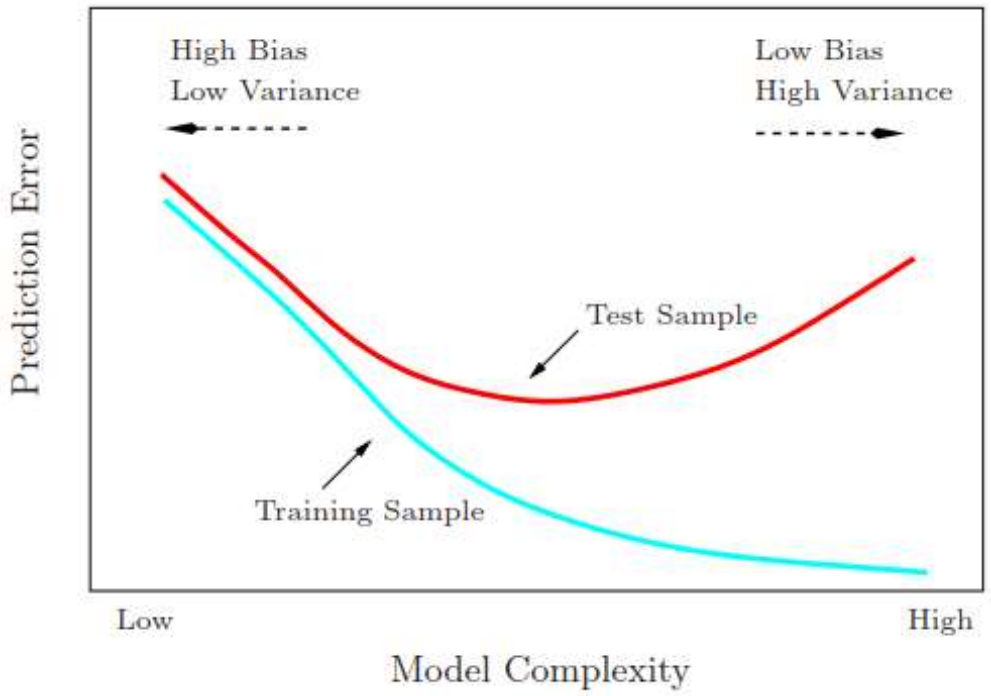
Entscheidungsbäume

Wie findet man einen guten Baum?

Wachsen und Stutzen

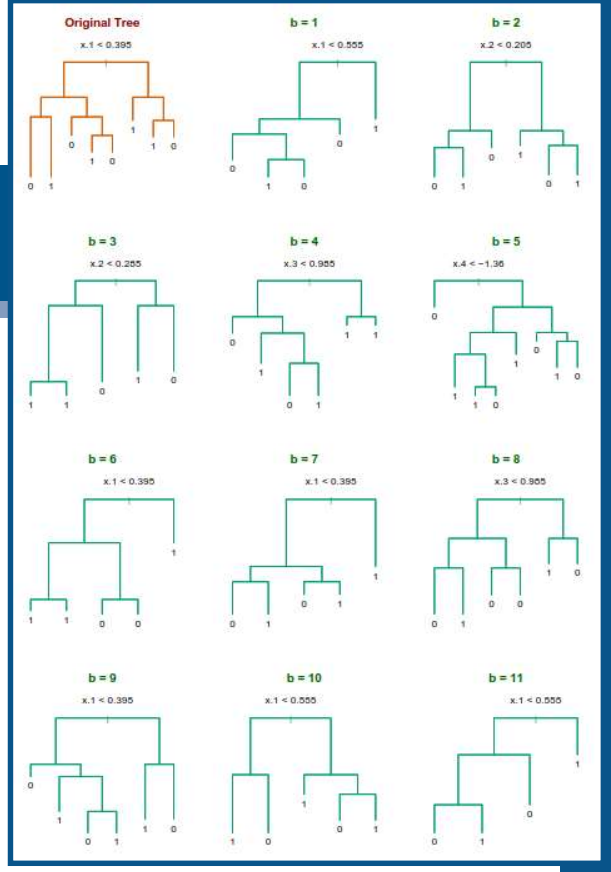
- Zum Ausgleich zwischen Bias und Varianz





Entscheidungsbäume

Erweiterungen



Methoden des Ensemble-Learning

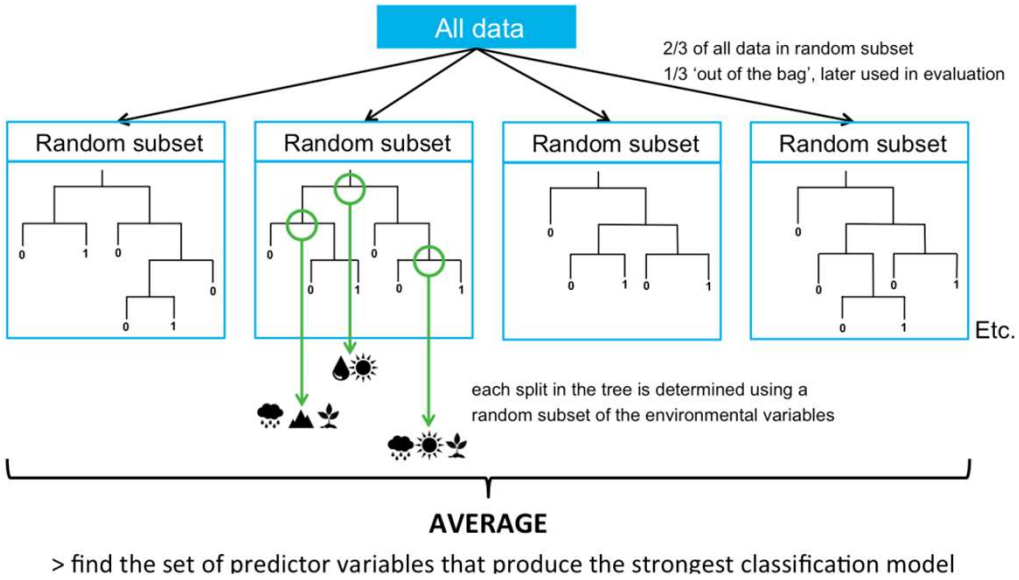
- Grundidee:
  - Einen bekannten Lernalgorithmus (z.B. CART) **mehrfach anwenden**
  - **„Durchschnittliche" Vorhersage** als finales Modell verwenden
- Motivation:
  - Beobachtung: Einzelne Modellinstanz tendiert zu Overfitting
    - Beispiel Bäume:
      - Alle Splits werden auf denselben Trainingsdaten ermittelt
      - Für eine Vorhersage kommen mehrere dieser Splits sukzessive zur Anwendung
  - Ansatz: Durch Mittelung mehrerer Modellinstanzen die **Varianz senken** (bei konstantem Bias)
    - Beispiel Bäume: Mittelung vieler verschieden trainierter Bäume mit geringem Bias
  - Ziel: Bessere Vorhersagegüte des Ensemble der Modellinstanzen im Vergleich zur Einzelinstanz

Entscheidungsbäume
Erweiterungen

Ensemble-Learning: Random Forests (Entscheidungswälder)

im Use Case

- Breiman (2001)
- Ziel: Varianzreduktion durch **gleichzeitige Variabilität in Datengrundlage und Merkmalen**
  - Insbesondere Verbesserung von Bagging durch zusätzliche Dekorrelation der Bäume
- Idee: **Kombination von Bagging mit Random Subspace Method** pro Knoten
  - Bagging: Resampling der Trainingsdaten mittels Bootstrap
  - Random Subspace Method:
    - zufällige Merkmalsauswahl für jeden einzelnen Knoten
  - Vorhersage durch Mittelung



Agenda

Einführung

Daten in der PKV

Data-Analytics-Verfahren zur Stornoprognose

Modelltuning und Optimierung

Ergebnis und Anwendung

Fazit & Lessons Learned

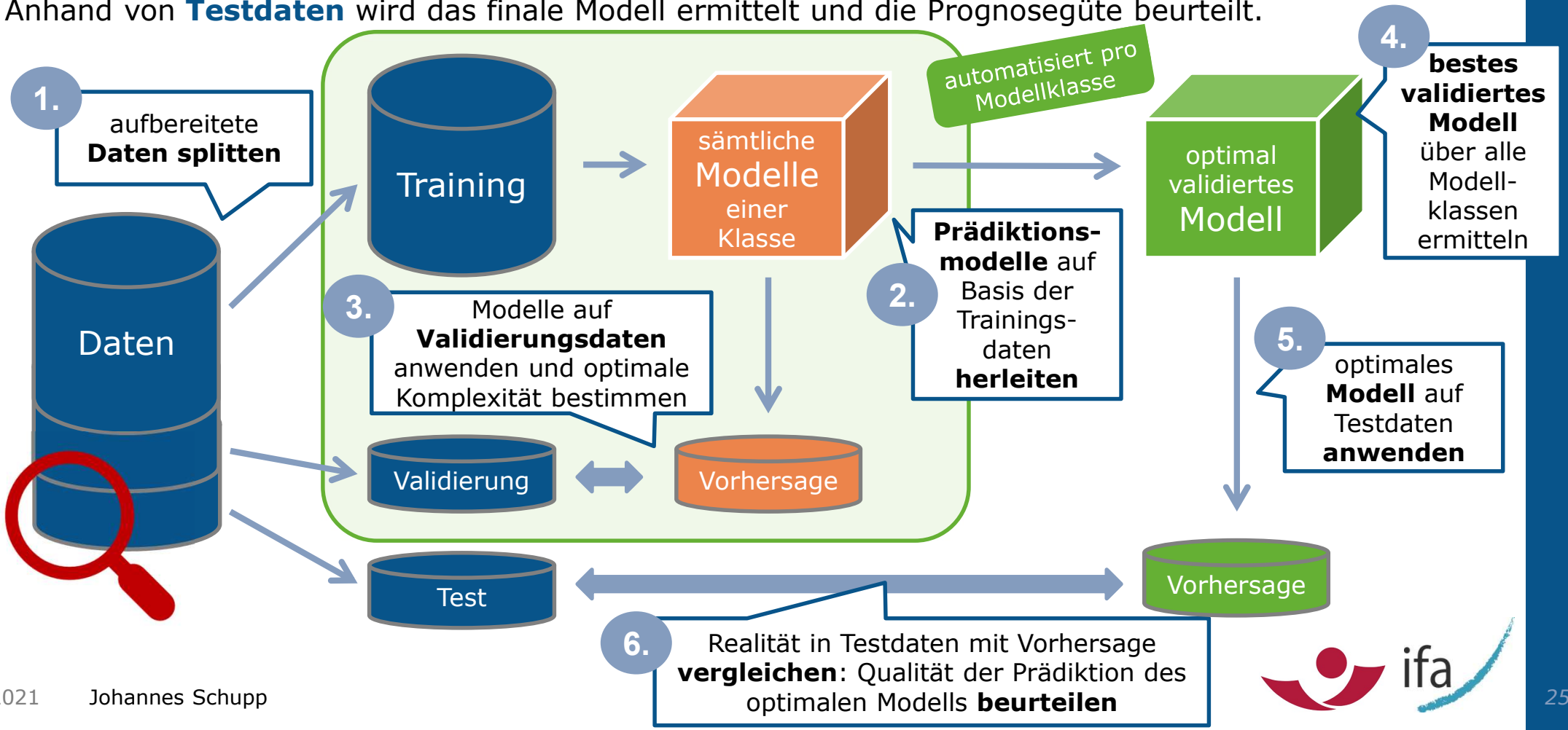
DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Optimierungsprozess

Training, Validierung und Test

Die Optimierung des Lernprozess erfolgt mit **Aufteilung der Daten** für Training, Validierung und Test:

- Auf den Erfahrungen in den **Trainingsdaten** lernt jedes Modell (verschiedene Komplexitäten).
- Mittels Erfahrungen in den **Validierungsdaten** wird die optimale Komplexität pro Modell ermittelt.
- Anhand von **Testdaten** wird das finale Modell ermittelt und die Prognosegüte beurteilt.



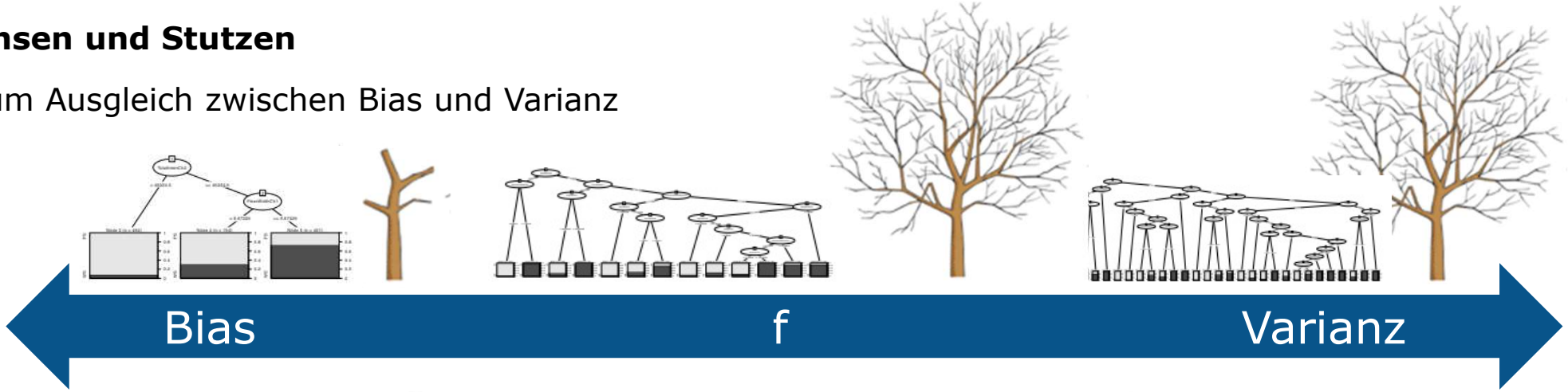
ifa

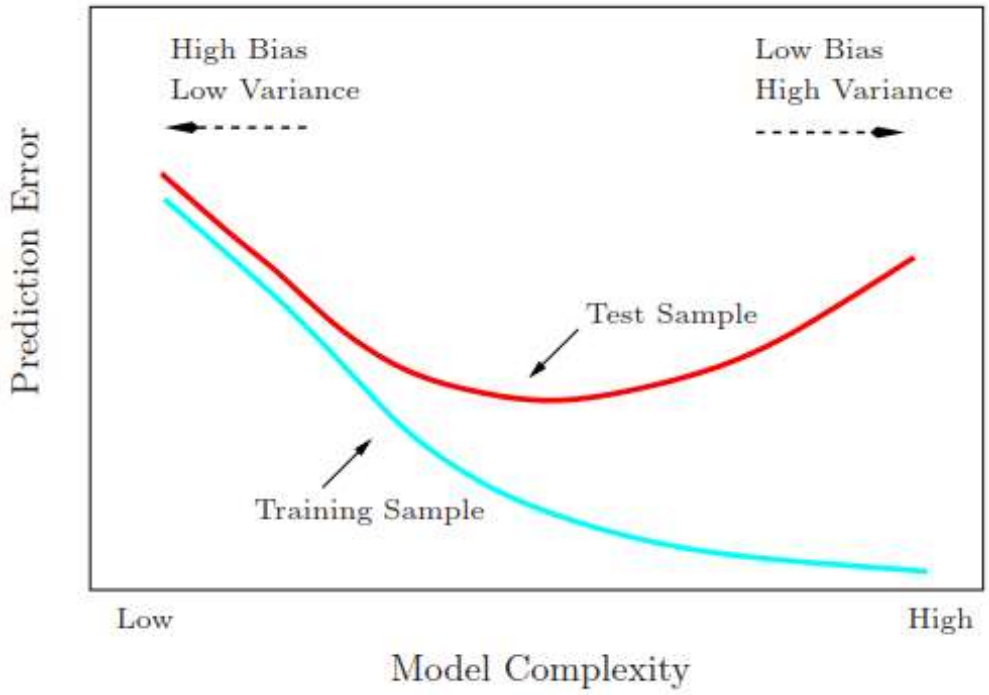
Optimierungsprozess

Wie findet man einen bestmöglichen Baum?

Wachsen und Stutzen

- Zum Ausgleich zwischen Bias und Varianz



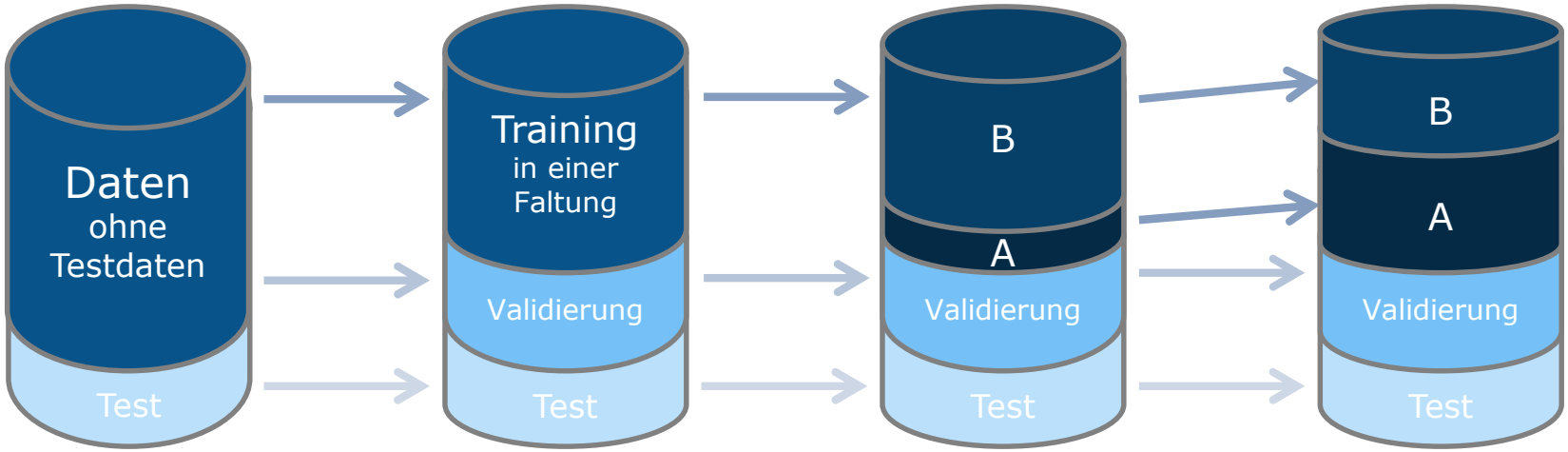


Optimierungsprozess

Vorgehen Imbalanced Data

Umgang mit einem ungleichgewichteten (imbalanced) Datensatz

- Beispielsweise liegt in einem Datensatz die Zielgröße in 5% mit A und in 95% mit B vor.
- Zum Umgang mit Ungleichgewicht in den Daten kann der Datensatz für das Training (ggf. in jeder Kreuzvalidierungsfaltung) **gewichtet** oder **verändert** werden (direkt oder indirekt).
- **Möglichkeit 1**: Gewichtung der Beobachtungen
  - Gewichtung von A verbleibt bei 1. Beobachtungen mit B erhalten Gewicht x/(1-x)=1/19, wobei x = 1/20, dem Anteil von A am Trainingsdatensatz.
  - In der Optimierungsfunktion werden die Gewichte x_i pro Beobachtung i ergänzt, d.h. implizit wird eine 50:50-Datenbasis als Ausgangslage genutzt.

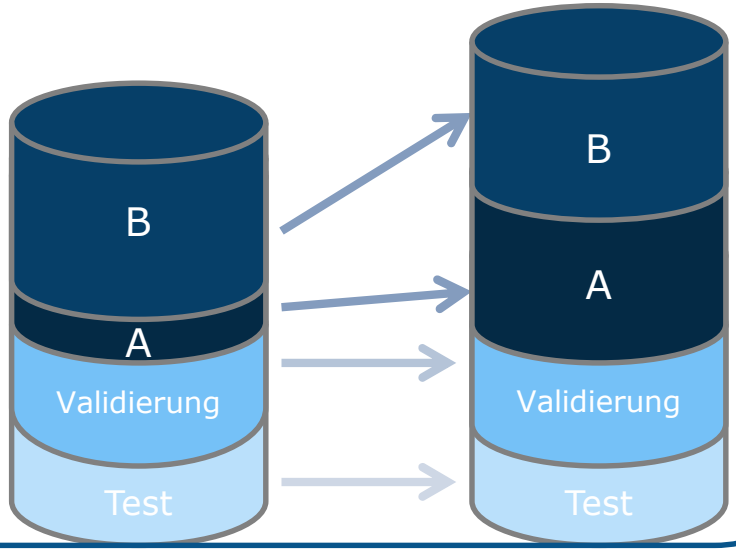


Optimierungsprozess

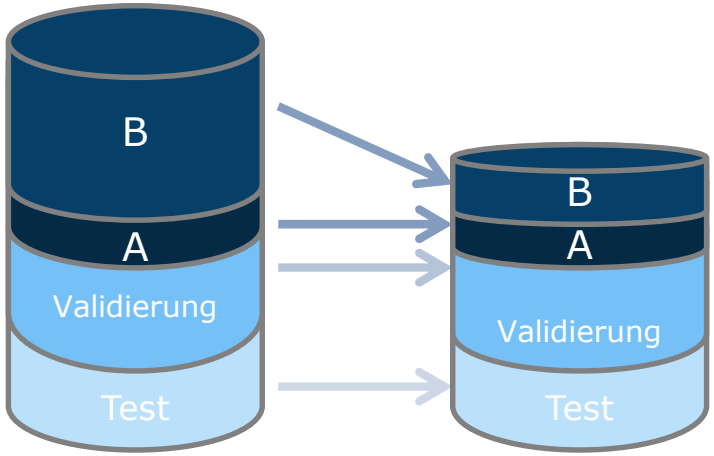
Vorgehen Imbalanced Data

Umgang mit einem ungleichgewichteten (imbalanced) Datensatz

- **Möglichkeit 2**: Upsampling (auch Oversampling)
  - Ziehe mit Zurücklegen aus Minderheit.



- **Möglichkeit 4**: Downsampling (auch Undersampling)
  - Ziehe ohne Zurücklegen aus Mehrheit.



im Use Case

- **Möglichkeit 3**: Wahl geeigneter Gütemaße
  - accuracy oder MSE ungeeignet
  - Statistische Gütekriterien (logloss) oder AUC schaffen einen Ausgleich zwischen ungleichgewichteten Klassen.

- **Möglichkeit 5:** Ensemble Downsampling
  - Bilde ein Ensemble aus (1-x)/x Modellen, sodass alle Daten mit Ausprägung B verwendet werden können.

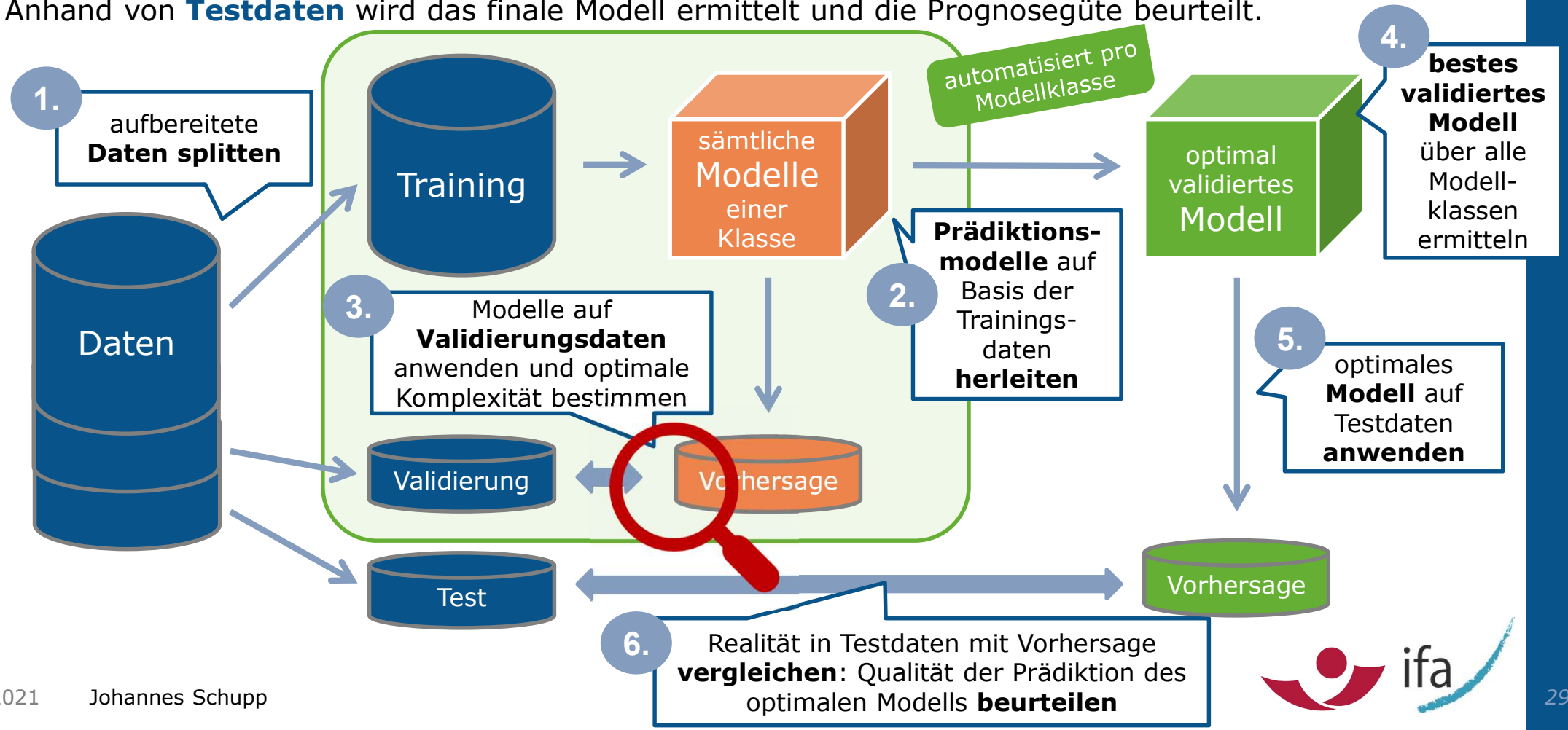
DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Optimierungsprozess

Training, Validierung und Test

Die Optimierung des Lernprozess erfolgt mit **Aufteilung der Daten** für Training, Validierung und Test:

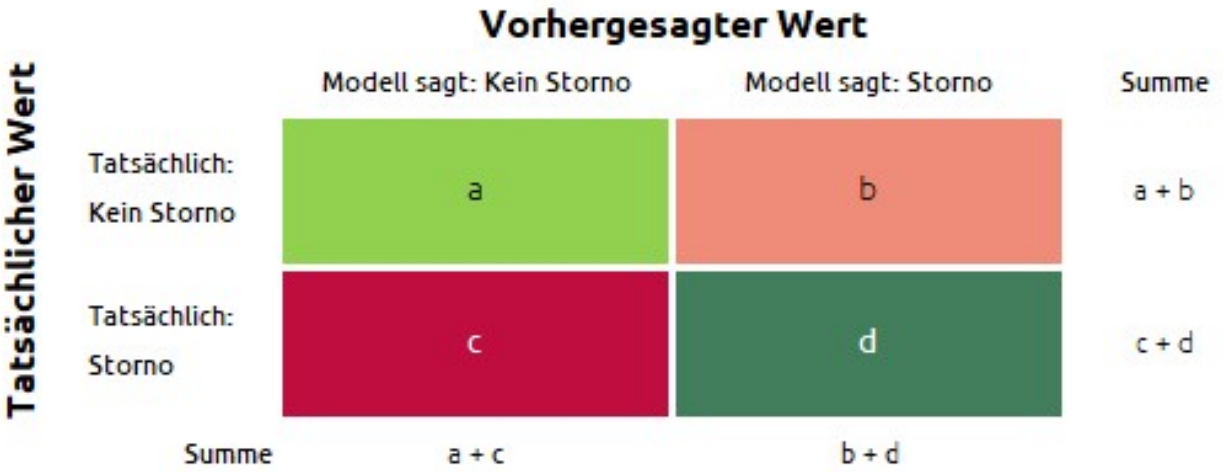
- Auf den Erfahrungen in den **Trainingsdaten** lernt jedes Modell (verschiedene Komplexitäten).
- Mittels Erfahrungen in den **Validierungsdaten** wird die optimale Komplexität pro Modell ermittelt.
- Anhand von **Testdaten** wird das finale Modell ermittelt und die Prognosegüte beurteilt.



Evalutionsmetriken: Konfusionsmatrix

Wie kann man Modelle miteinander vergleichen?

- Bewertung anhand der Konfusionsmatrix:



- **Precision** $\frac{d}{d+b}$

Anteil tatsächlicher Stornierer an Personen, für die Storno vorhergesagt wird

→ *zu niedrig:* (zu) viele Fälle als Stornierer klassifiziert

- **Sensitivität** $\frac{d}{d+c}$

Anteil der korrekt vorhergesagten Abgänge an allen Stornierern

→ *zu niedrig:* (zu) wenige Abgänge als Stornierer klassifiziert

- **Spezifität** $\frac{a}{a+c}$

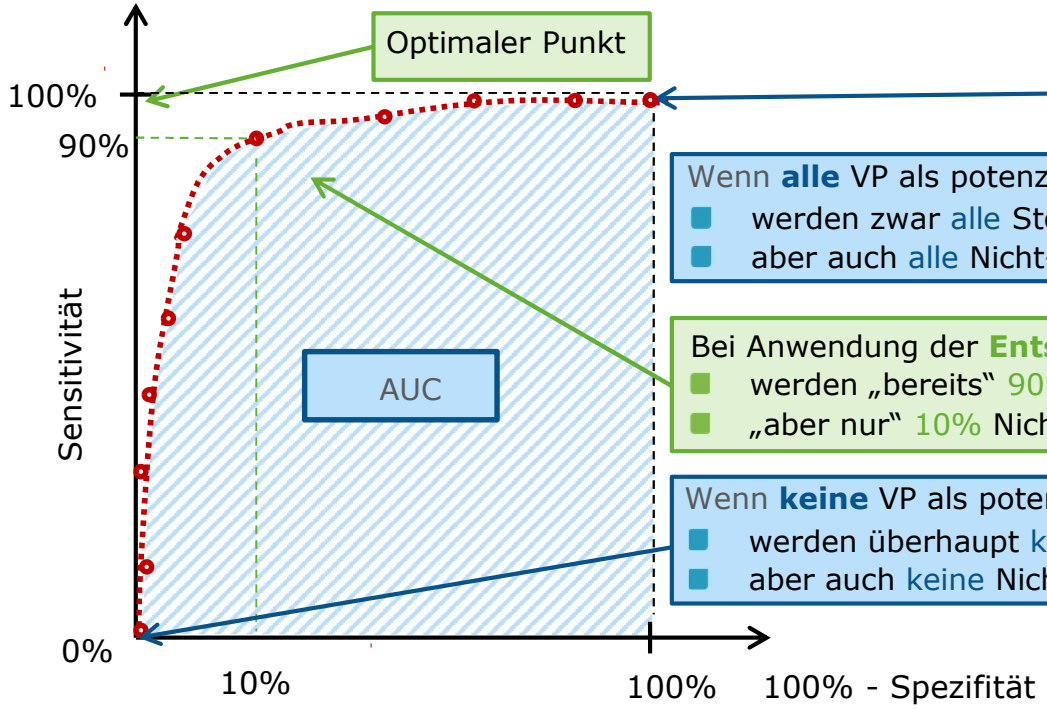
Anteil tatsächlicher Nicht-Stornierer an Personen, für die Nicht-Storno vorhergesagt wird

- **Accuracy, F1-Score, …**

Modellvergleich, Modellgüte, Ergebnisformate

ROC-Kurve – Stornomodell

- Die **ROC-Kurve** ergibt sich durch Anwendung verschiedener Entscheidungsregeln auf die Testdaten (jeder Punkt auf der Kurve steht für eine Entscheidungsregel des Modells).
  - Visualisierung der Abhängigkeit von Effizienz und Fehlerrate verschiedener Regeln durch
    - Y-Achse: Anteil der richtigerweise erkannten Stornofälle (im Verhältnis zu allen tatsächlichen Stornofällen) versus
    - X-Achse: den Anteil der fälschlicherweise als Stornofälle identifizierten versicherten Personen VP (im Verhältnis zu allen Nicht-Stornofällen)



Beispiel

Wenn **alle** VP als potenzielle Stornofälle eingestuft werden,
- werden zwar alle Stornofälle rechtzeitig erkannt (sensitiv)
- aber auch alle Nicht-Stornofälle falsch als Stornofälle eingestuft (unspezifisch).

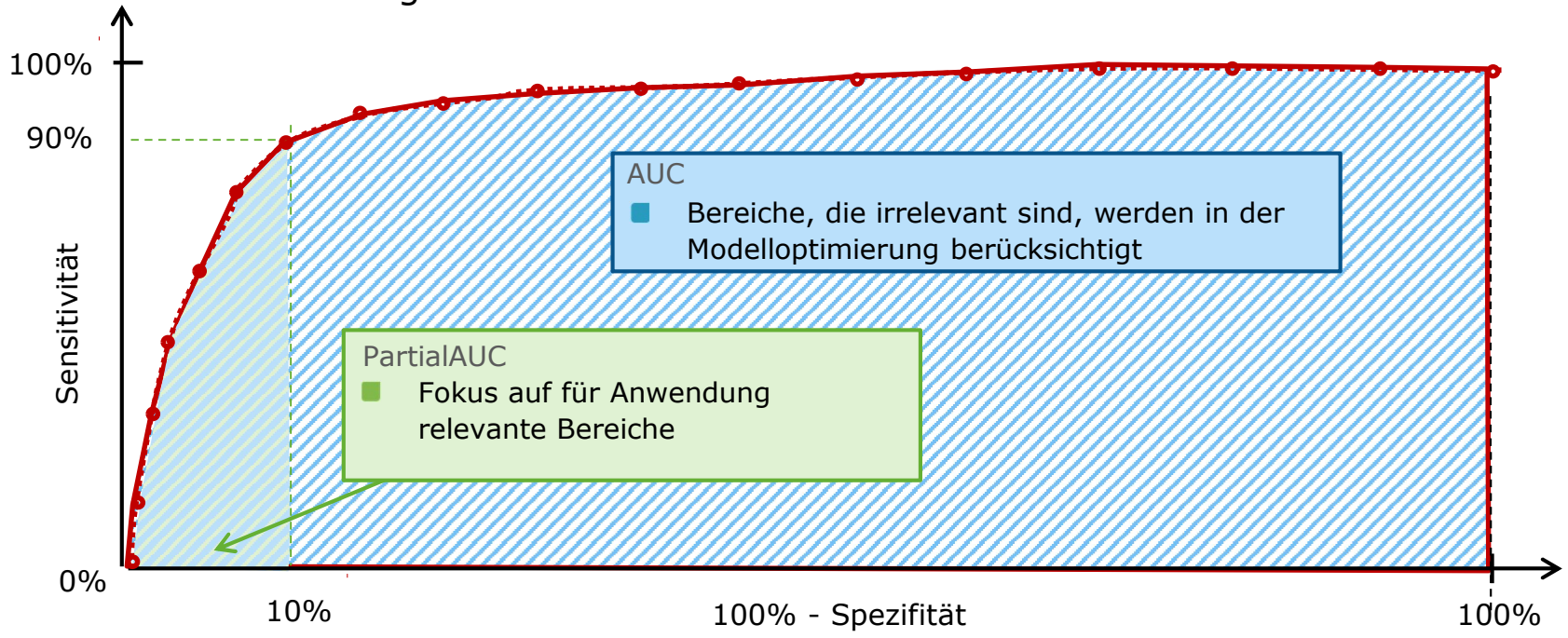
Bei Anwendung der **Entscheidungsregel in diesem Punkt**
- werden „bereits" 90% aller Stornofälle rechtzeitig erkannt („sensitiv")
- „aber nur" 10% Nicht-Stornofälle falsch als Stornofälle eingestuft („spezifisch").

Wenn **keine** VP als potenzielle Stornofälle eingestuft werden,
- werden überhaupt keine Stornofälle rechtzeitig erkannt (nicht sensitiv)
- aber auch keine Nicht-Stornofälle falsch als Stornofälle eingestuft (spezifisch).

Modellvergleich, Modellgüte, Ergebnisformate

ROC-Kurve – Imbalanced Data

- Im Datensatz liegen sehr viel mehr Stornierer vor als Nicht-Stornierer. Deshalb ist der Einfluss der Nicht-Stornierer in der ursprünglichen AUC-Betrachtung größer (X-Achse).
- Anstatt der gesamten ROC-Kurve (via AUC) kann auch nur ein Teil der ROC-Kurve (via PartialAUC) optimiert werden.
  - Dabei wird der Höchstwert der False-Positive Rate (100% - Spezifität) eingeschränkt, z.B. auf 10%.
  - Die Modelloptimierung kann sich auf den späteren Anwendungsbereich mit einer konkreten Schwellenwertregel fokussieren.



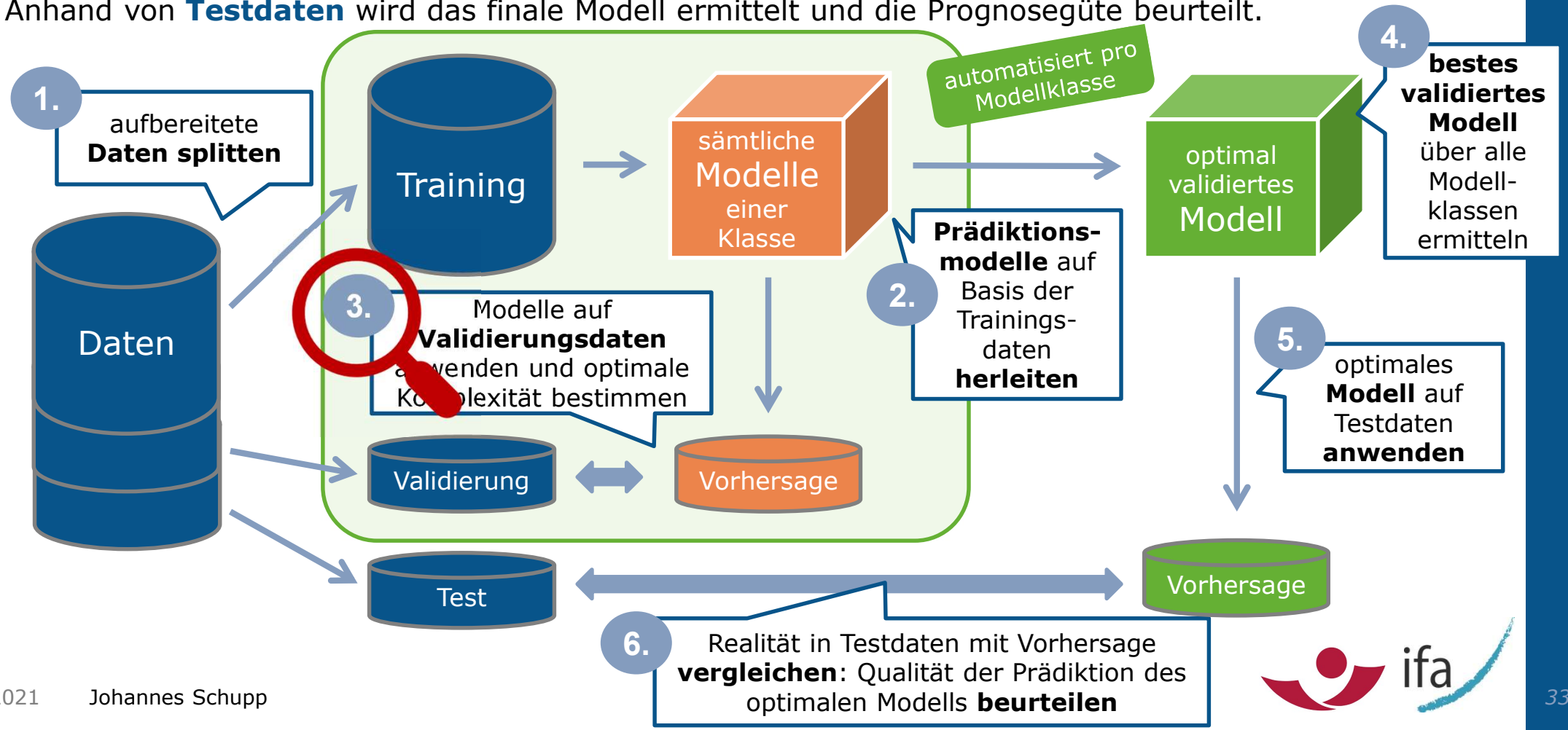
DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Optimierungsprozess

Training, Validierung und Test

Die Optimierung des Lernprozess erfolgt mit **Aufteilung der Daten** für Training, Validierung und Test:

- Auf den Erfahrungen in den **Trainingsdaten** lernt jedes Modell (verschiedene Komplexitäten).
- Mittels Erfahrungen in den **Validierungsdaten** wird die optimale Komplexität pro Modell ermittelt.
- Anhand von **Testdaten** wird das finale Modell ermittelt und die Prognosegüte beurteilt.

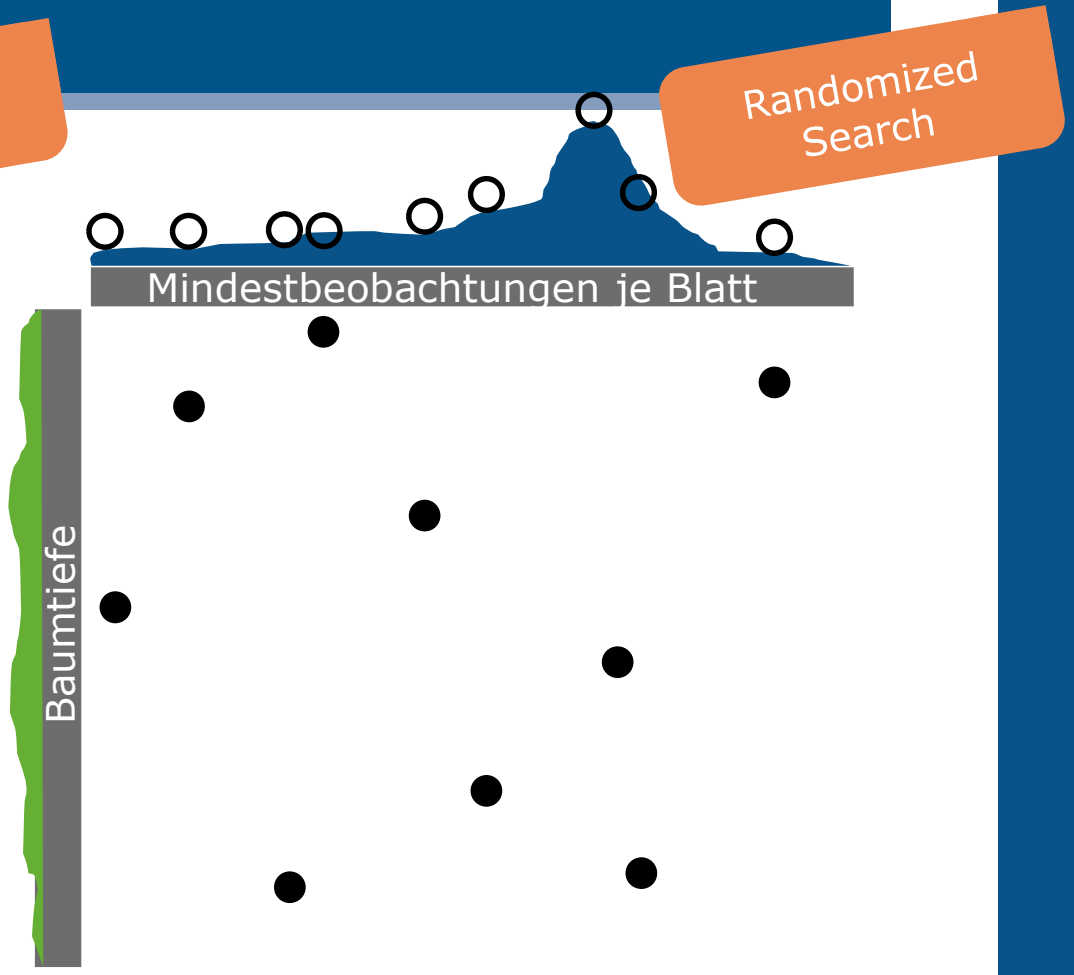


ifa

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Hyperparameter Tuning

Grid Search vs. Randomized Search

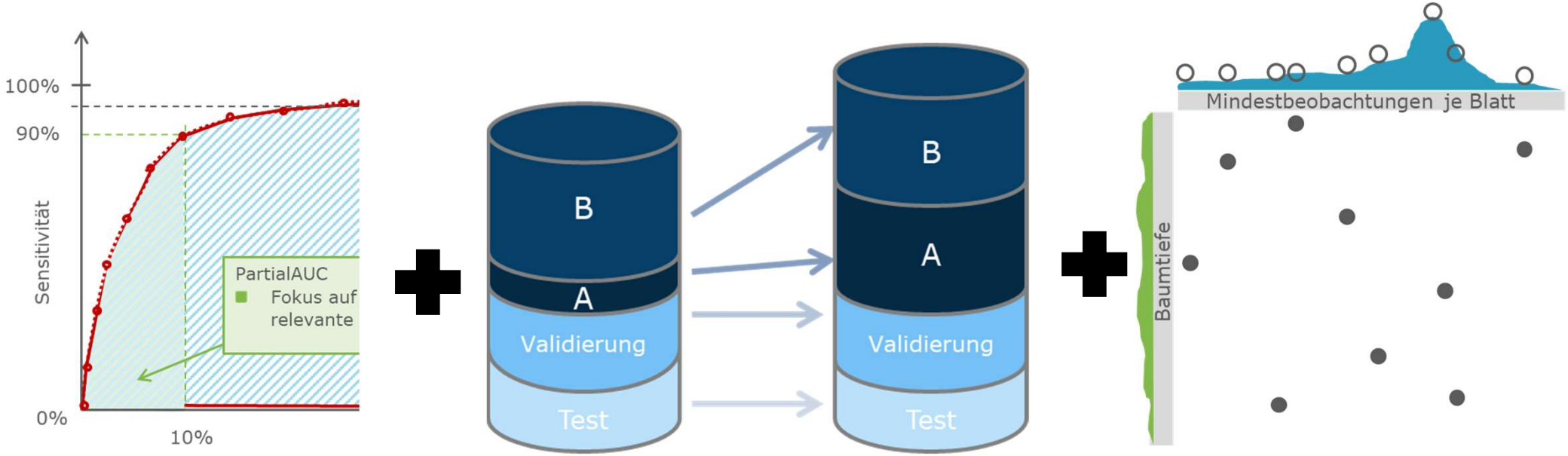


Bei einer Randomized Search werden (bei gleichem Zeitaufwand!) Hyperparameter ausführlicher analysiert!

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Finales Modell

Erst das Tuning mit drei wesentlichen Bausteinen liefert ein gutes Ergebnis!




```
In [ ]: # Balanced Random Forest

brfc = imb.ensemble.BalancedRandomForestClassifier(random_state=INTEGER,
                                                   n_estimators=ESTIMATOR,
                                                   max_depth=DEPTH,
                                                   min_samples_split=SPLIT,
                                                   max_features=FEATURES,
                                                   min_impurity_decrease=IMPURITY,
                                                   bootstrap=BOOLEAN_1,
                                                   sampling_strategy=STRATEGY,
                                                   replacement=BOOLEAN_2)
```


ifa

Agenda

Einführung

Daten in der PKV

Data-Analytics-Verfahren zur Stornoprognose

Modelltuning und Optimierung

Ergebnis und Anwendung

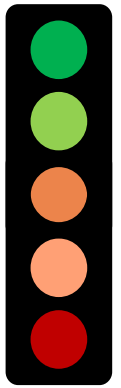
Fazit & Lessons Learned

DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Projektergebnisse

Entwicklung eines Scorewertes

- Ist es möglich, ein Prognosemodell für die individuelle Stornierungswahrscheinlichkeit eines vollversicherten Kunden zu erstellen?
  - **Ja, es ist möglich!**



- Ziel der Modellkalibrierung: die „rote Ampel" soll möglichst gut sein!

Finale Vorhersage des Modells für Testdatensatz:

Score-wert	*Stornogefahr*	*Anteil*
0	keine	*70,7%*
1	kaum	*13,1%*
2	gering	*5,5%*
3	erhöht	*5,0%*
4	groß	*4,9%*
5	sehr groß	*0,8%*
gesamt		100%

Tatsächliche Beobachtung:

Scorewert-Stornoquote
X_0%
X_1%
X_2%
X_3%
X_4%
X_5%
X_{quer}%

Finale Vorhersage des Modells für 2021:

Score-wert	*Stornogefahr*	*Anteil*
0	keine	*71,1%*
1	kaum	*12,3%*
2	gering	*5,8%*
3	erhöht	*4,3%*
4	groß	*4,9%*
5	sehr groß	*1,5%*
gesamt		100%

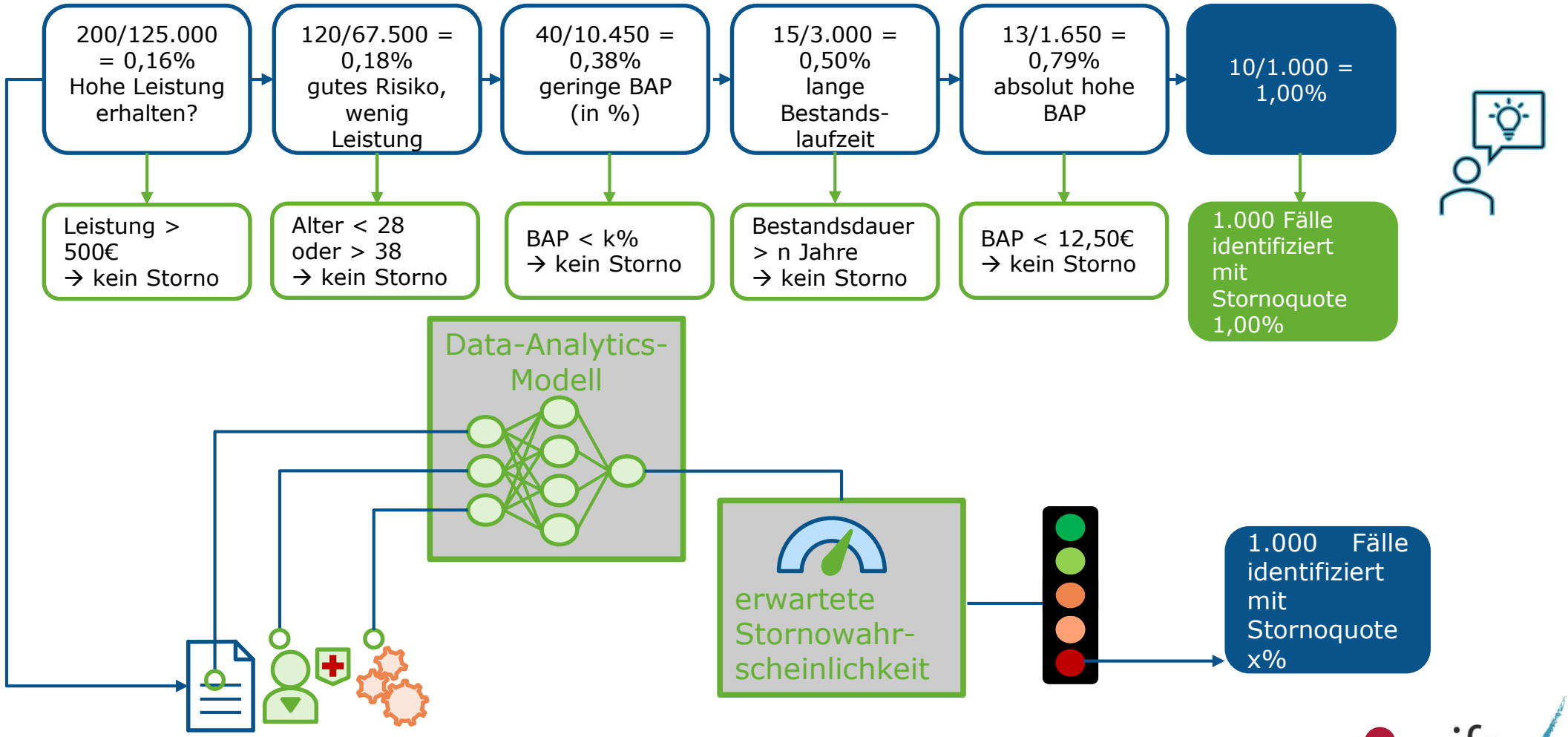
Tatsächliche Beobachtung:

Scorewert-Stornoquote
Y_0%
Y_1%
Y_2%
Y_3%
Y_4%
Y_5%
Y_{quer}%

ifa

Verbesserung durch Data-Analytics-Verfahren

Es ist eine Kundenbindungsaktion mit einem Budget für 1.000 Kunden geplant. Auftrag für Fachexperte: „Bitte selektieren Sie 1.000 Fälle aus 125.000, die stark stornogefährdet sein könnten!“



Anwendungen von Stornomodellen in der PKV

- Langfristige Bindung an das Versicherungsunternehmen ist nicht automatisch gegeben. Die Bereitschaft zum Unternehmenswechsel ist jedoch vorhanden.
- Es ist wünschenswert, dass die abwanderungsgefährdeten Bestandskunden erkannt und durch geeignete Maßnahmen an das Unternehmen gebunden werden.
- Im Vergleich zu klassisch abgeleiteten Stornotafeln ermöglichen die Stornomodelle eine bessere Steuerung der Vorhersagen auf Einzelkundensicht.
- Ziel: Kunden gemäß ihrer Stornogefahr zu sortieren, um anschließend eine gewisse Kundenanzahl (z. B. abhängig vom Budget) zur Stornoprävention auszuwählen.

Maßnahme bestimmt das Modell

- Entscheidend für die Feinjustierung des Modells ist die geplante Maßnahme
- Auswahl möglicher Anwendungsszenarien:

Ampel

- Systematische Anzeige eines Scorewertes als Hilfestellung bei Kundengesprächen
- zusätzlich wichtig: entscheidende Merkmale für hohen Scorewert

Fokus im Use Case

Kundenbindung (Aktionen)

- Ermittlung einer begrenzter Anzahl von Kunden
- z.B. 1.000 Kunden sollen kontaktiert werden

Bestandsführung

- gezielte Maßnahmen für bestimmte Kundengruppen
- Stornogründe benötigt
- z.B. durch computergestützte Rechnungsprüfsysteme

Maßnahme bestimmt das Modell

- Bei Anwendung des Modells und einer Maßnahme zur Stornoprävention sollte die **Maßnahme kontrolliert** werden.
- Mögliches Vorgehen: gemäß Modell die abwanderungsgefährdeten Bestandskunden in zwei Gruppen zufällig und gleichmäßig einteilen

Experimentgruppe	Kontrollgruppe
• zu kontaktierende Kunden • Beobachtung, ob die eingeleitete Maßnahme Storno verhindern konnte • Hat die Maßnahme etwas gebracht?	• keine Maßnahmen durchführen • Beobachtung von Storno ohne potentielle Beeinflussung • War die Prognose korrekt?

- Wichtig für die Weiterentwicklung des Modells
- Eine regelmäßige Neujustierung anhand aktueller Daten notwendig

Agenda

Einführung

Daten in der PKV

Data-Analytics-Verfahren zur Stornoprognose

Modelltuning und Optimierung

Ergebnis und Anwendung

Fazit & Lessons Learned

Fazit & Lessons Learned

- Data-Analytics-Verfahren können umfangreiche Datenmengen auf Muster untersuchen.
  - Für Storno besonders relevant: Alterungsrückstellung, BAP-Historie, Bestandsdauer
- Abhängig von der **zuvor festgelegten** Maßnahme kann das Modell optimiert werden.
- entscheidend sind dabei
  - **versicherungstechnisches Wissen** → welche relevanten Daten kann man bekommen?
  - **versicherungswirtschaftliches Wissen** → welche Maßnahme macht Sinn?
  - **statistisches Wissen** → wie werden Modelle eingestellt und optimiert?
  - **Programmierkenntnisse** → Umsetzung in Python
- Im Use Case ist das Data-Analytics-Verfahren um ein Vielfaches besser als der Fachexperte.

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Kontakt

Vielen Dank für Ihre Aufmerksamkeit!

Ella Russer
ella.maria.russer@hallesche.de

Kinga Böhm
kinga.boehm@hallesche.de

Dr. Johannes Schupp
+49 (731) 20 644-241
j.schupp@ifa-ulm.de

Eugenia Jung
eugenia.jung@hallesche.de

ifa

Quellenverzeichnis

- 15: seaborn.pairplot, https://seaborn.pydata.org/generated/seaborn.pairplot.html, 03.05.2021.
- 23, 24: Hastie et al. (2009). The Elements of Statistical Learning – Data Mining, Inference, and Prediction.
- 25: Biodiversity and Climate Change Virtual Laboratory (2016), Random Forest, https://support.bccvl.org.au/support/solutions/articles/6000083217-random-forest#header-page3, 31.03.2021.

Rechtliche Hinweise

Gerne überlassen wir Ihnen diese Präsentation zu Informationszwecken. Bitte beachten Sie aber, dass die darin enthaltenen Informationen allgemeiner Natur sind und eine Beratung im konkreten Einzelfall nicht ersetzen können.

Diese Unterlage haben wir nach bestem Wissen erstellt und die Inhalte sorgfältig erarbeitet. Gleichwohl kann man Fehler nie ganz ausschließen. Bitte haben Sie deshalb Verständnis dafür, dass wir keine Garantie und Haftung für die Aktualität, Richtigkeit und Vollständigkeit übernehmen. Infolgedessen haften wir nicht für direkte, indirekte, zufällige oder besondere Schäden, die Ihnen oder Dritten entstehen. Der Haftungsausschluss gilt nicht für vorsätzliches oder grob fahrlässiges Handeln oder bei Nichtvorhandensein zugesicherter Eigenschaften.

In die Zukunft gerichtete Aussagen sind naturgemäß mit Ungewissheiten verbunden. Deshalb können die tatsächlichen Ergebnisse von diesen abweichen. Eine Verpflichtung zur Aktualisierung von Zukunftsaussagen wird nicht übernommen.

Bei Kapitalanlage-Produkten gilt zusätzlich: Die Präsentation stellt keine Anlageberatung dar und sollte auch nicht als Grundlage für eine Anlageentscheidung dienen. Aus den gegebenenfalls dargestellten Wertentwicklungen der Vergangenheit können keine Rückschlüsse auf zukünftige Wertsteigerungen gezogen werden.

Unsere Marken und Logos sind international markenrechtlich geschützt. Es ist nicht gestattet, diese Marken und Logos ohne unsere vorherige schriftliche Zustimmung zu nutzen.