

# Einsatz von Machine-Learning-Verfahren und deren Erklärbarkeit in der Unfallversicherung

Cem Keles, Johannes Schupp

# Inhalt

**Einführung**

**Aufbau des Use Cases**

**Eingesetzte Machine-Learning-Verfahren**

**Die Erklärbarkeit und Weiterverwendung**

**Fazit und Lessons Learned**

## Umfeld und Ausgangslage

- Unfallversicherungsmarkt in Deutschland weitgehend Verdrängungsmarkt
  - Stornosicherung sehr wichtig!
- Schon seit 15 Jahren existiert beim Unternehmen ein Modell zur Prognose von Storno auf Einzelfallbasis.
- Ziel:
  - Update, Überprüfung und Verbesserung des bisherigen Stornomodells
  - Einsatz moderner Machine-Learning-Verfahren
  - Zentrierung am Kundenbedürfnis (Kunde = Vertriebsmitarbeiter)

Jahr	Anzahl Versicherungsunternehmen	Beiträge <sup>1</sup>		Leistungen <sup>2</sup>		Schaden-Kosten-Quote <sup>3</sup>
		in Mio. EUR	Veränderung ggü. Vorjahr	in Mio. EUR	Veränderung ggü. Vorjahr	
2012	130	6.522	0,5%	3.038	-1,0%	79,3%
2013	129	6.411	-1,7%	3.092	1,8%	79,4%
2014	130	6.471	0,9%	3.167	2,4%	80,9%
2015	131	6.365	-1,6%	3.236	2,2%	79,2%
2016	127	6.450	1,3%	3.238	0,0%	77,4%
2017	123	6.479	0,5%	3.364	3,9%	81,5%
2018	123	6.545	1,0%	3.393	0,9%	77,4%

Quelle: GDV

# Wahrnehmung und Umfeld bisheriges Stornomodell

Wir haben einen Ferrari,  
können aber  
die Gangschaltung nicht  
bedienen

von Storno würde ich  
gerne rechtzeitig oder generell  
mal erfahren

Den Prognosen glaube ich nicht.  
Meine Kunden kenn' ich viel besser!

Ich sehe da Potential drin

Wir haben zu viele Programme,  
zu viele Systeme. Da wird so ein  
Auftrag auch mal übersehen



Klasse,nehm ich.  
Kommt dann zum  
01.01. oder wann?

Wieso stellen Sie nicht einfach eine  
generelle Ablaufliste der UV  
zur Verfügung [...]? Damit könnten Sie  
mir viel Arbeitszeit abnehmen

Was? Wir haben sowas?

Auf jeden Fall, würd  
ich täglich nutzen

# Inhalt

## **Einführung**

### **Aufbau des Use Cases**

Einblicke in die Zusammensetzung und Hintergründe

abgeleitetes Modellierungsziel

### **Eingesetzte Machine-Learning-Verfahren**

### **Die Erklärbarkeit und Weiterverwendung**

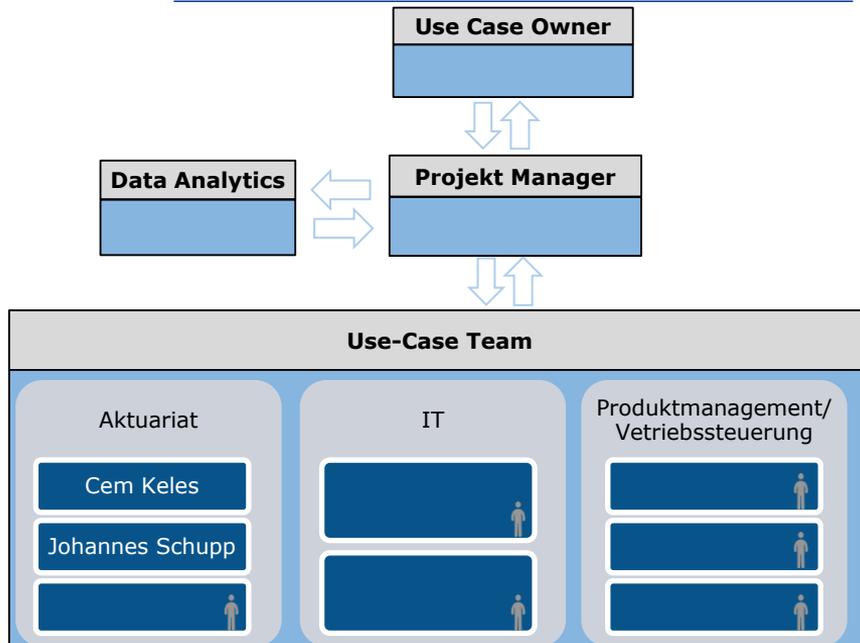
### **Fazit und Lessons Learned**

# Use Case „proVISION“ als interdisziplinäres Team

## Arbeitsweise



### Use-Case Organisation



- Agile Projektorganisation
  - Organisation in zweiwöchigen Sprints
  - Sprintplanning mit komplettem Team
  - Weekly zur Fortschrittsabstimmung
  - Falls möglich, gemeinsame Arbeit auf Projektfläche
- Jedes Use-Case Mitglied behält den Überblick und somit das Projektziel im Auge
  - gemeinsames Ziel → Teambuilding!
- Projekt Manager erkennt frühzeitig Hindernisse und beseitigt diese

## Zweigledrige Zielsetzung

Beide Treiber spielen unmittelbar auf den Erfolg des Use Cases ein

### • Modellierung & IT



- Umfangreicher Ausbau der zugrundeliegenden Datenbasis
- Modellierung mit modernen Data-Analytics-Verfahren
  - z.B. Gradient Boosting, Elastic Net, Neuronale Netze

### • Vermarktung & Außendienst



- Sicherstellung der Kundenzentrierung
- Erhöhung der Akzeptanz des Stornomodells
- Erhöhung der Akzeptanz von Machine-Learning-Verfahren
  - auch perspektivisch wichtig!

- Das interdisziplinäre Team bearbeitet beide Zielsetzungen gleichzeitig und gemeinsam!
- Anforderung an Modellierung:
  - nicht nur bestmögliches Modell, sondern **Erklärung der Modellprognosen** auf Einzelfallbasis
  - besser **verständlich und genutzt** als **komplex und nicht verwendet**

# Inhalt

## **Einführung**

## **Aufbau des Use Cases**

## **Eingesetzte Machine-Learning-Verfahren**

Regularisierte GLM

XGBoost

## **Die Erklärbarkeit und Weiterverwendung**

## **Fazit und Lessons Learned**

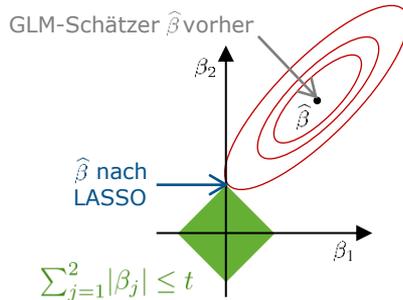
# Erklärbare Modelle

## Regularisierte GLM

$$\ell_{\text{pen}}(\beta; y_i, x_i, \lambda) = \ell(\beta; y_i, x_i) + \lambda \Omega(\beta_1, \dots, \beta_m) \rightarrow \min$$

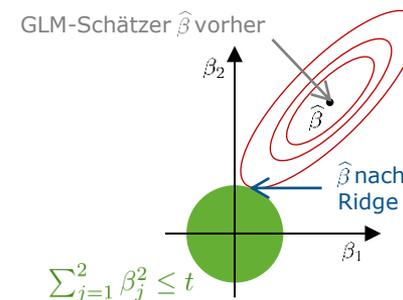
### LASSO<sup>1</sup>

$$\Omega(\beta_1, \dots, \beta_m) = \|(\beta_1, \dots, \beta_m)\|_1 = \sum_{j=1}^m |\beta_j|$$



### Ridge<sup>2</sup>

$$\Omega(\beta_1, \dots, \beta_m) = \|(\beta_1, \dots, \beta_m)\|_2^2 = \sum_{j=1}^m \beta_j^2$$



### Elastisches Netz: Mischung beider Normen

$$\Omega(\beta_1, \dots, \beta_m; \alpha) = \alpha \|(\beta_1, \dots, \beta_m)\|_1 + \frac{1}{2} (1 - \alpha) \|(\beta_1, \dots, \beta_m)\|_2^2$$

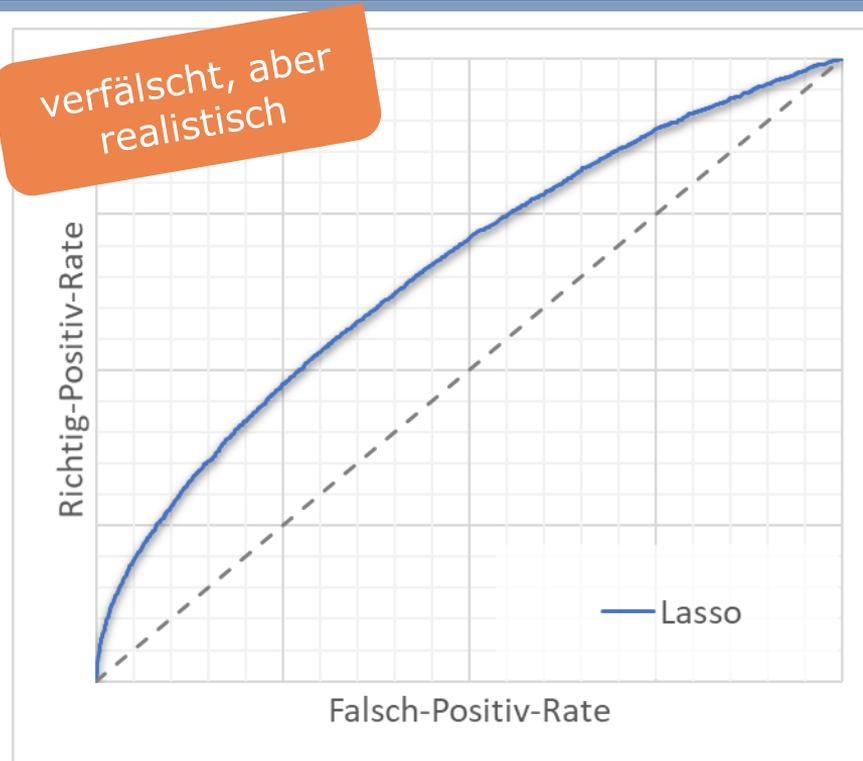
<sup>1</sup> Least Absolute Selection and Shrinkage Operator; nach Tibshirani (1996)

<sup>2</sup> engl. „Bergrücken“; benannt in Anlehnung zu Ridge-Analysen (Hoerl 1970), auch Thikonov-Regularisierung

# Erklärbare Modelle

## GLM mit Regularisierung

- Ausgabe und Interpretation analog zum bekannten GLM-Output
- Unterschied liegt in der regularisierten Parameterschätzung
- Erklärbarkeit der Prognose direkt über  $\beta_i$  gegeben
  
- Ergebnisformate
  - AUC
  - Konfusionsmatrix
  - Dezilplots
  - ...
  - Sehr ähnliche Ergebnisse für LASSO, RIDGE, ELASTIC NET



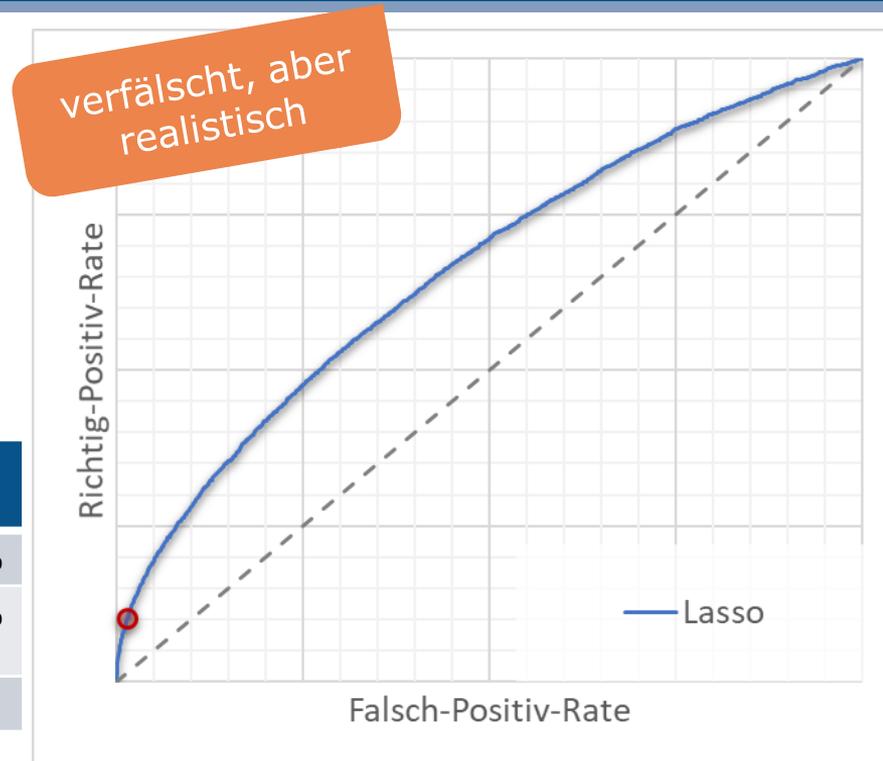
# Erklärbare Modelle

## GLM mit Regularisierung

- AUC / ROC-Kurven in Ergebnisdarstellung allein zu kompliziert
- zusätzlich: lege Schwellenwert so fest, dass x% Stornierer erkannt werden (z.B. x=10%)
  - Vergleich einer Zahl: fälschlicherweise als Storno vorhergesagte Nichtstornierer
  - ökonomisch: Anzahl Fehlansprache zu minimieren!

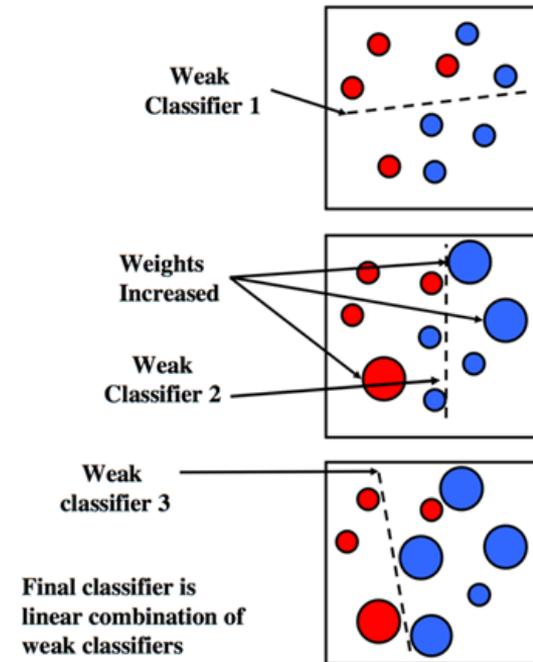
LASSO	Storno- prognose	Keine Stornoprognose	korrekt
Storno	2.172	19.548	10%
Kein Storno	5.615	260.811	97,9%
Total	7.787	280.359	

- ABER: geringes Prognosepotenzial (Linearität)



# Gradient Boosting und Regularisierung

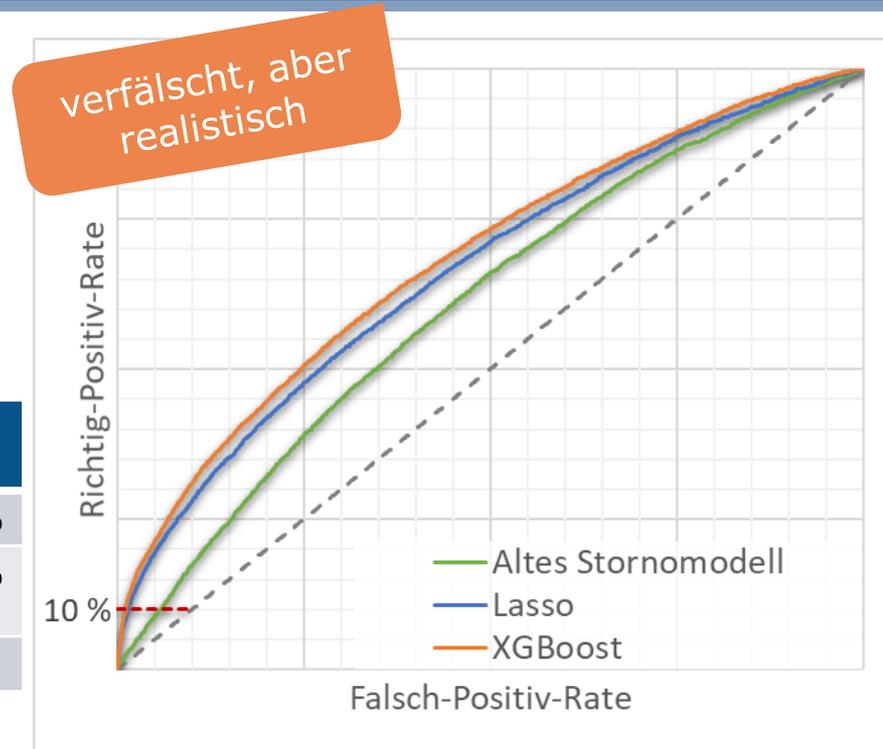
- GBM: Bestimmung der Residuen bei jedem Klassifikationsbaum,
  - schlechte Vorhersagen in Folgebäumen sequentiell und sukzessive verbessern
- XGBoost (eXtreme Gradient Boosting) gibt optimierten Algorithmus des GBM's mit Regularisierung analog zu Lasso und Ridge
  - niedrigere Laufzeiten, sparsamerer Arbeitsspeicherungang, bessere Ergebnisse im Vergleich zu GBM
- (automatisiertes) Erkennen von Nichtlinearen Strukturen in den Daten



# XGBoost Modell

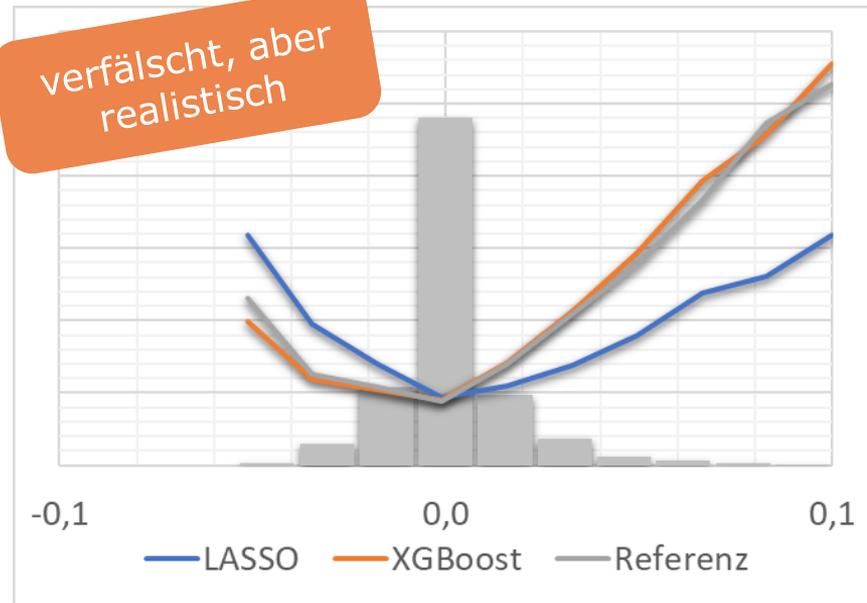
- Bessere Prognosegüte hinsichtlich aller Gütemaße, z.B.
  - um 10% Stornierer zu kontaktieren, werden 1.180 eigentliche Nichtstornierer weniger kontaktiert (15% weniger)
  - Ohne Tuning: Direkte Verbesserung AUC
  - Feintuning: weitere Verbesserung AUC

LASSO	Storno- prognose	Keine Stornoprognose	korrekt
Storno	2.172	19.548	10%
Kein Storno	<b>4.435</b>	261.991	98,3%
Total	6.607	281.539	



# XGBoost Modell

- Weitere Modellvergleiche:
  - weitere Maße:
    - Area Under Precision-Recall Curve
    - Devianz/Likelihood
    - ...
  - Visualisierung
    - **Dezilplots**
    - Liftplots
    - ...



Wir haben nun also ein gutes Stornomodell – wie erklären wir das?

## **Einführung**

### **Aufbau des Use Cases**

### **Eingesetzte Machine-Learning-Verfahren**

### **Die Erklärbarkeit und Weiterverwendung**

Warum erklärbar?

Modelle zur Erklärbarkeit der Modellprognose

Modelle zur Analyse des Einflusses eines Merkmals

Verwendung dieser Modelle im Vertriebscockpit

### **Fazit und Lessons Learned**

# Wie erklärt man das XGBoost Modell?

## Möglichkeiten

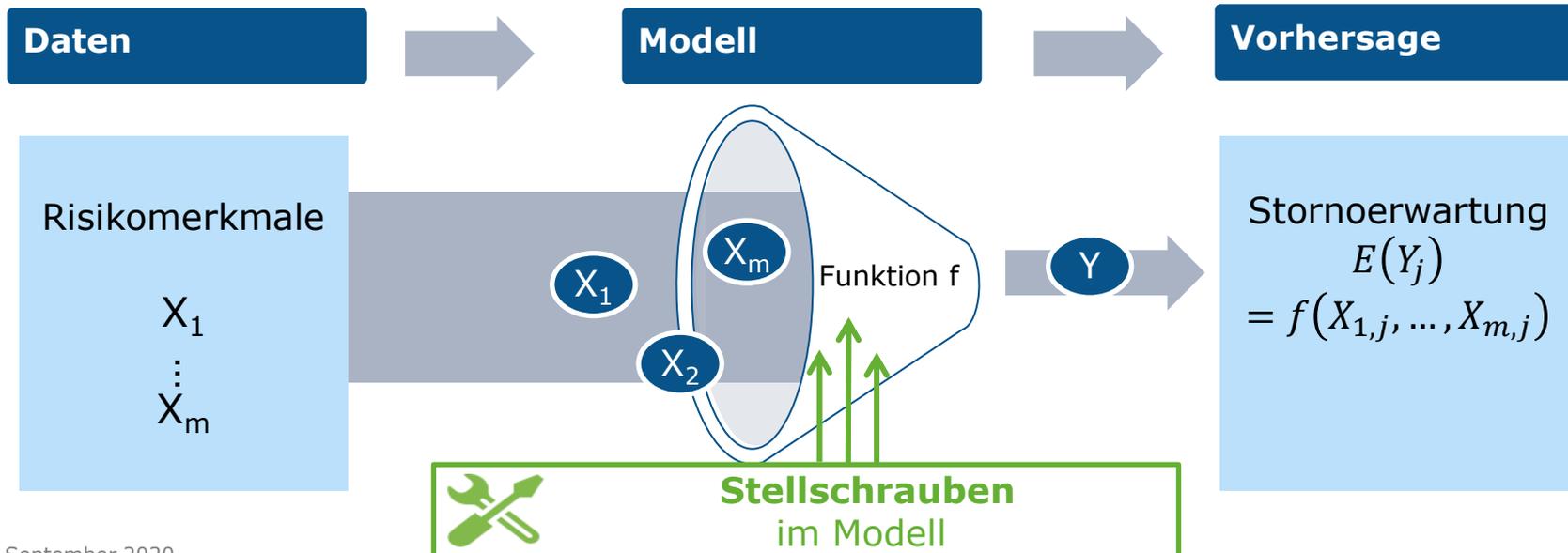
- Zentral in Kommunikation mit Vertrieb:
  - **Warum hat ein Vertrag eine hohe Stornowahrscheinlichkeit?**
- Antwort GLM/Lasso:
  - Die 3-5 Merkmale mit den absolut größten Betas nennen (bei Standardisierung der Merkmale)
  - Nicht: Liste mit allen  $\beta_i$ ! (Stichwort: Akzeptanz und Verständnis)
- Antwort XGBoost (und auch RandomForest, Neuronales Netz, SVM, ...)
  - Einsatz **Explainerverfahren** zur Visualisierung
    - zur Erklärung der Einzelvorhersage
    - zur Erklärung des Einflusses eines Merkmals

# Explainermodelle

(high-level) Theorie

- Beobachtung  $j \in \{1, \dots, n\}$  von Risikomerkmale  $X_{1,j}, \dots, X_{m,j}$  einer Zielgröße  $Y_j$
- Machine-Learning-Verfahren versuchen auf unterschiedliche Weise eine parametrische Funktion  $f$  zu kalibrieren, sodass die Zielgröße möglichst gut approximiert wird:

$$f(X_{1,j}, \dots, X_{m,j}) = \hat{Y}_j = E(Y_j)$$



# Wie erklärt man das XGBoost Modell?

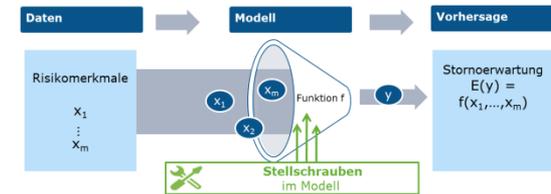
## Möglichkeiten

- Einsatz **Explainerverfahren**

- zur Erklärung der Einzelvorhersage
  - Kalibrierte Funktion  $f$  ist zu kompliziert.
  - globale oder lokale Approximation der Funktion  $f$
  - Ziel: Erklärbarkeit einer Prognose
  - z.B. LIME Modelle, Surrogate Trees, modellspezifische Ansätze
- zur Erklärung des Einflusses eines Merkmals
  - Approximation der partiellen Ableitungen von  $f$ , d.h. zum Beispiel bzgl. des ersten Merkmals

$$\frac{df(X_1, \dots, X_m)}{dX_1}$$

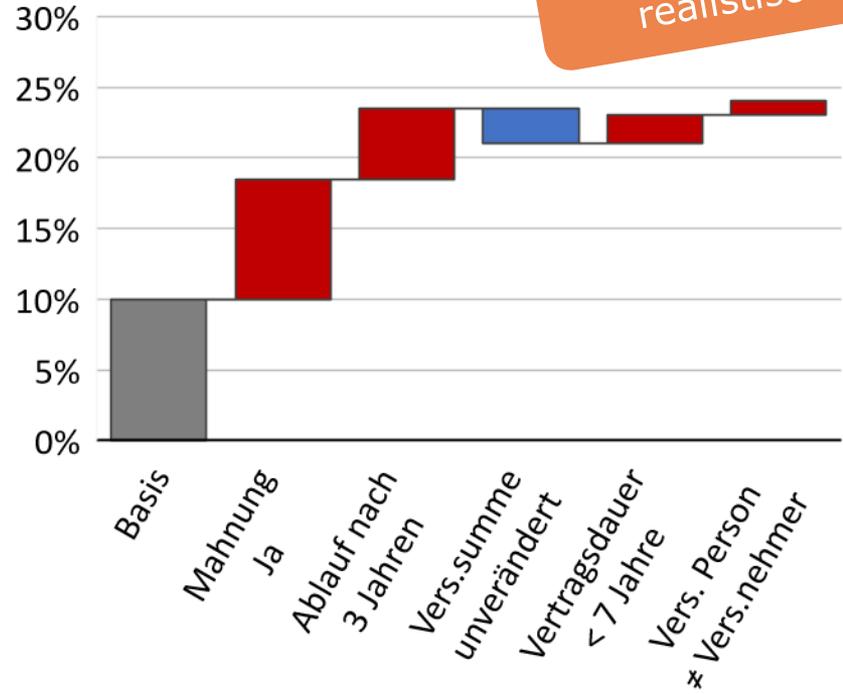
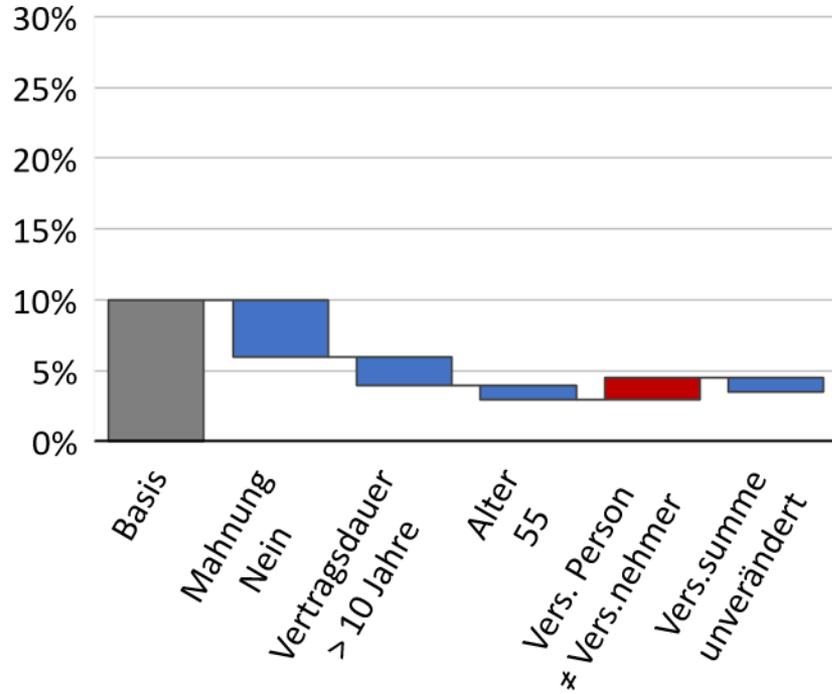
- Ziel: Plausibilisierung des Prognosemodells
- z.B. PDP, ICE, ALE



# LIME-plots

## Kommunikation zum Vertrieb

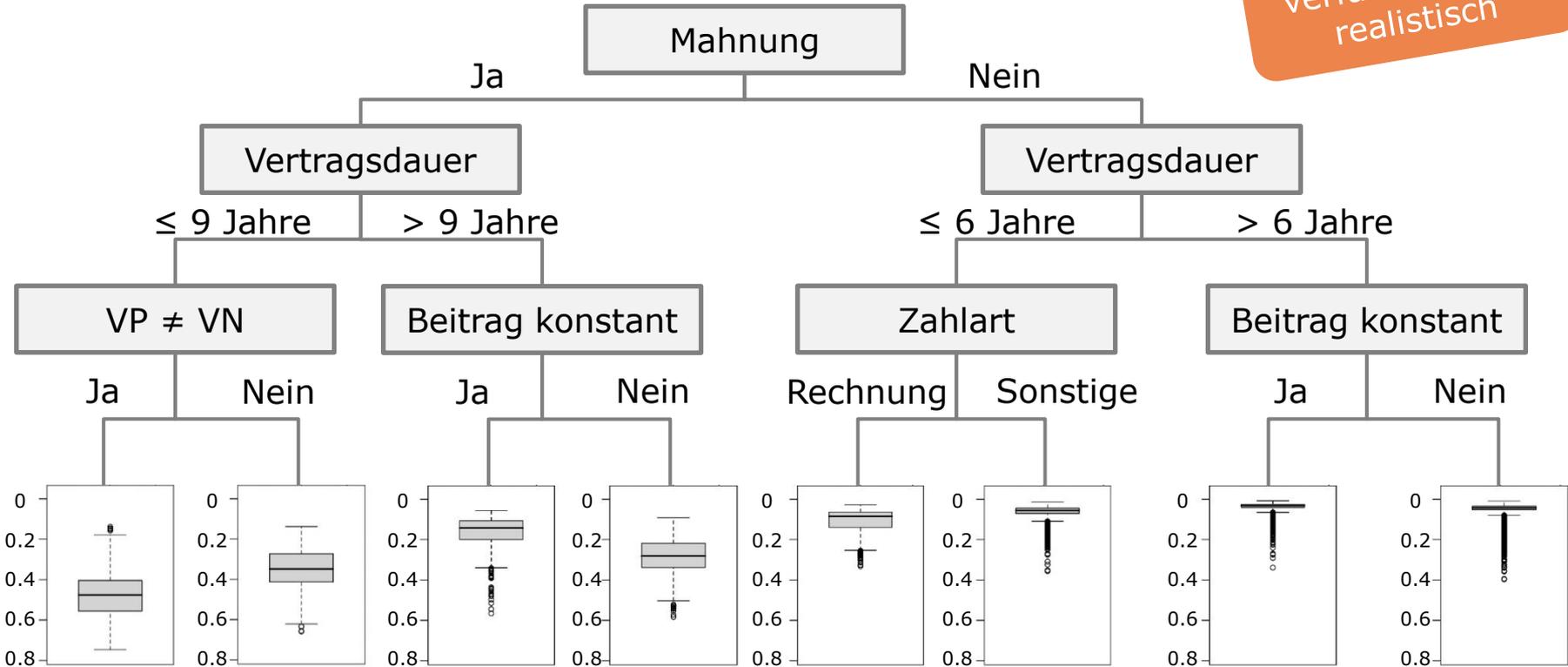
verfälscht, aber  
realistisch



# Surrogate Tree

Kommunikation im Aktuarat

verfälscht, aber  
realistisch

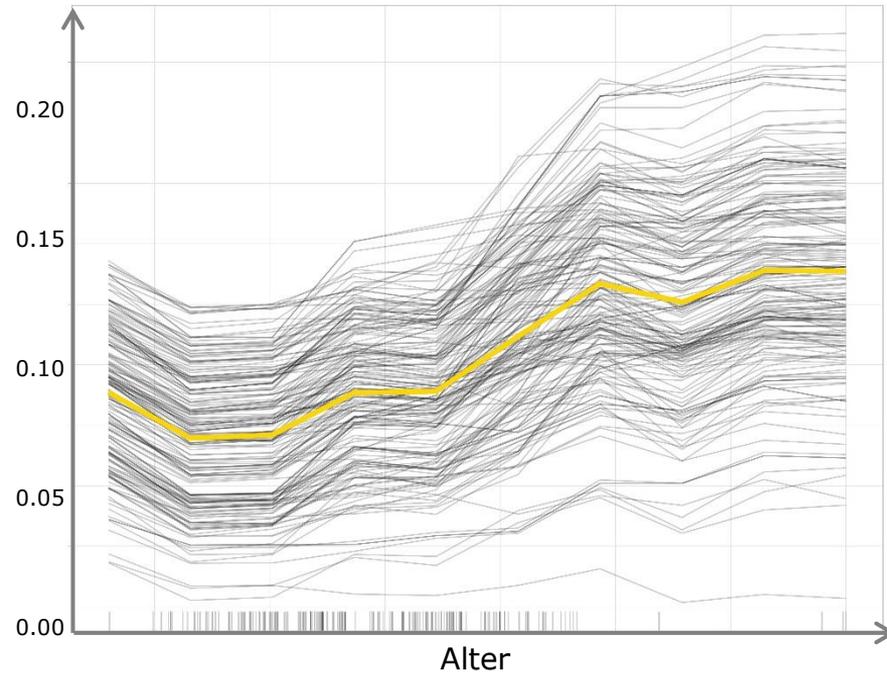


# Partial Dependence Plots

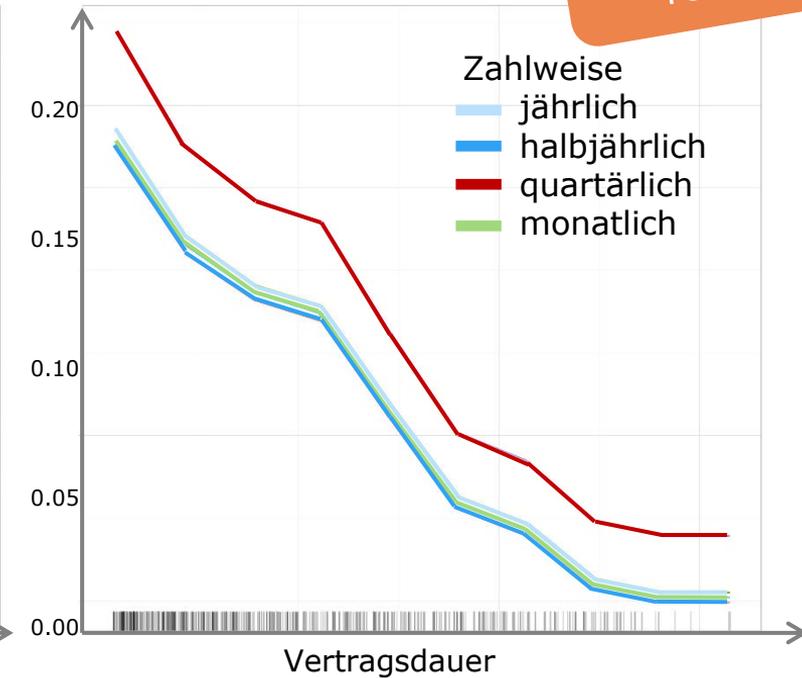
## Kommunikation im Aktuarat

verfälscht, aber  
realistisch

Stornowahrscheinlichkeit



Stornowahrscheinlichkeit



# Einbinden in Vertriebscockpit

**Umsatzentwicklung**

UV ↗ KV ↘

LV ↘ PKV ↗

KFZ → WGV ↑

---

Gesamt 43.050€

Vormonat -3%

Vorjahr +7%

**Wohngebäude**

Ziel

92%

(erwarteter) Bestandsverlauf

Provision: 4362€

Storno: 6,2%

Ablauf: 4,3%

**Vorschlag 1:** Ablauf von 4 Verträgen. Produktempfehlung: TG2019 (Provision: 254€)

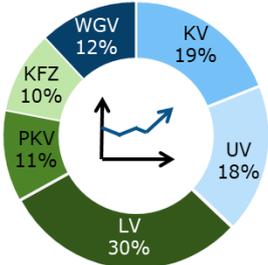
**Vorschlag 2:** Verträge mit hoher Stornowahrscheinlichkeit identifiziert. (Provision: -325€)

**Vorschlag 3:** Absicherungslücke identifiziert bei Hans Müller (Provision: 34€)

**Vorschlag 4:** Umstellung auf Seniorenarief von Horst Maier (Provision: 16€)

Filtern    Sortieren    Suchen

Name	Vorname	Geburtsdatum	Vertrag	Prämien-summe	Provision	Storno-gefahr	Kunden-wert	letztes Gespräch
Schöpf	Johanna	18.05.1990	UV	3.450	138	Hoch	1	2019
Hafer	Bernd	16.02.1974	KV	2.507	100	Gering	3	2018
Becks	Sandy	05.05.1977	LV	105.355	4214	Gering	2	2010
Becks	Sandy	05.05.1977	KV	6.504	260	Mittel	1	2010
Ziesel	Hans-Peter	03.03.1960	UV	15.035	601	Gering	2	2016
Blombe	Sandro	04.11.1975	PKV	69.073	2763	Hoch	3	2015



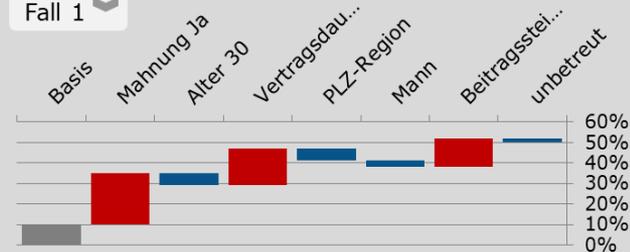
# Einbinden in Vertriebscockpit

**Vorschlag 2:** Verträge mit hoher Stornowahrscheinlichkeit identifiziert. (Provision: **-325€**)

## Erklärung

- Teilbestand mit um **Faktor 9** erhöhter Stornowahrscheinlichkeit
- Prognostizierte Stornowahrscheinlichkeit: **45%** (Bestandsdurchschnitt: 5%)
- Hauptgründe für erhöhte Stornoprognose: Inkassoschreiben, Dauer < 5 Jahre, geringe Versicherungssumme

Fall 1



Informationen zu den gefährdeten Verträgen [hier](#)



**Vorschlag 1:** Ablauf von 4 Verträgen. Produktempfehlung: TG2019 (Provision: **254€**)

**Vorschlag 2:** Verträge mit hoher Stornowahrscheinlichkeit identifiziert. (Provision: **-325€**)

**Vorschlag 3:** Absicherungslücke identifiziert bei Hans Müller (Provision: **34€**)

**Vorschlag 4:** Umstellung auf Seniorentarif von Horst Maier (Provision: **16€**)

Kombination mit weiteren Machine-Learning-Entwicklungen wichtig und sinnvoll, z.B.

- Steuerung der Wiederanlage
- Cross-/Upselling
- Erweiterung um erwartete Provisionen, Kundenwertmodell usw.

# Inhalt

**Einführung**

**Aufbau des Use Cases**

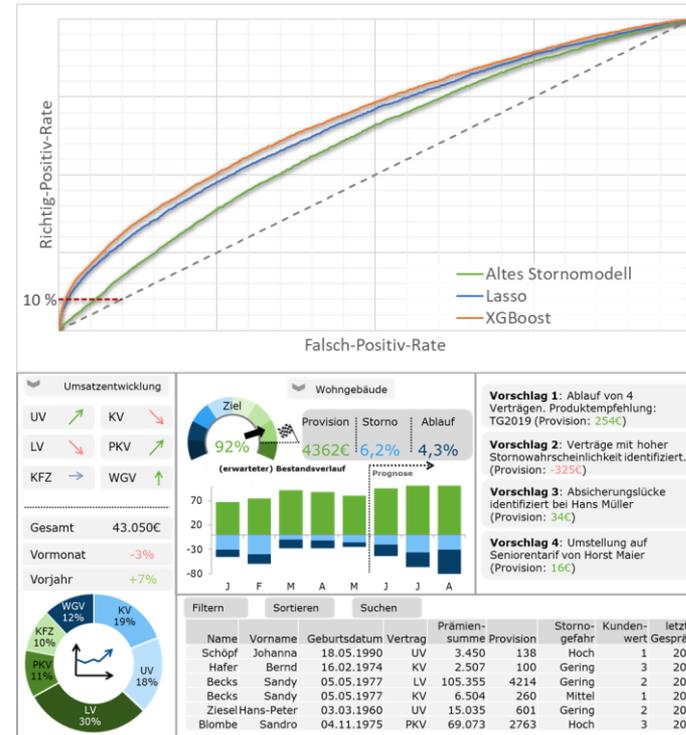
**Eingesetzte Machine-Learning-Verfahren**

**Die Erklärbarkeit und Weiterverwendung**

**Fazit und Lessons Learned**

# Fazit

- Ziele erreicht?
  - Update, Überprüfung und Verbesserung des bisherigen Stornomodells
  - → Prognosegüte bisheriges Stornomodell deutlich übertroffen
  - Einsatz moderner Machine-Learning-Verfahren
  - → XGBoost gilt als eines der besten Machine-Learning-Verfahren
  - Zentrierung am Kundenbedürfnis (Kunde = Vertriebsmitarbeiter)
  - → Gezielter Einsatz von Explainerverfahren zur Erklärung und Visualisierung der Kundenbedürfnisse



## Lessons Learned

- Aufteilung in Test-/Trainingsdaten:
  - Mehrere VP unterhalb eines VN müssen bei Aufteilung der Daten (auch innerhalb Kreuzvalidierung) gemeinsam behandelt werden, da sehr stark abhängig
  - andernfalls: Modell lernt alle VP auswendig
- Feature Engineering
  - Das XGBoost verbessert sich deutlich, wenn man die sonst übliche Gruppierung numerischer Merkmalen (z. B. Altersgruppen) dem Modell überlässt
- Datenfilterung
  - Eliminierung von Personen, die gar nicht stornieren können, z.B. weil sie einen mehrjährigen Vertrag abgeschlossen haben
    - verschlechtert zwar z.B. den AUC, weil „kein Storno“ schlechter prognostiziert wird
    - dafür kann sich das Modell auf die Fälle konzentrieren, die überhaupt für Storno in Frage kommen
- Agile Projektorganisation
  - hat wesentlich zum Projekterfolg beigetragen
  - erfordert Erfahrung und Sachverstand zur Auswahl geeigneter Agiler Methoden
  - macht Spaß!

# Kontaktdaten

**SIGNAL IDUNA** 

Institut für Finanz- und  
Aktuarwissenschaften

**ifa**



**Cem Keles**  
+49 (40) 4124-4350  
cem.keles@signal-iduna.de



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

