



Auctor Actor Advisor



Qualifizierung des BU-Risikos mit künstlicher Intelligenz

Forum V

Versicherungsmathematisches Kolloquium
der FAU

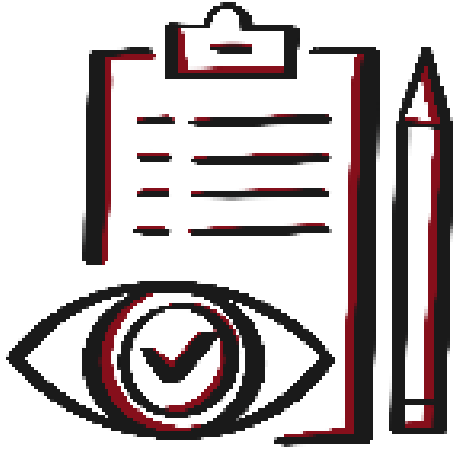
Michael Stoffels | 26.11.2024



Was ist Berufsunfähigkeit

- Berufsunfähig ist, wer seinen **zuletzt ausgeübten Beruf**, so wie er ohne gesundheitliche Beeinträchtigung ausgestaltet war, infolge Krankheit, Körperverletzung oder mehr als altersentsprechendem Kräfteverfall ganz oder teilweise **voraussichtlich auf Dauer** nicht mehr ausüben kann.
- Als weitere Voraussetzung einer Leistungspflicht des Versicherers kann vereinbart werden, dass die versicherte Person **auch keine andere Tätigkeit** ausübt oder ausüben kann, die zu übernehmen sie auf Grund ihrer Ausbildung und Fähigkeiten in der Lage ist und die ihrer bisherigen Lebensstellung entspricht.

Abstrakte Verweisbarkeit

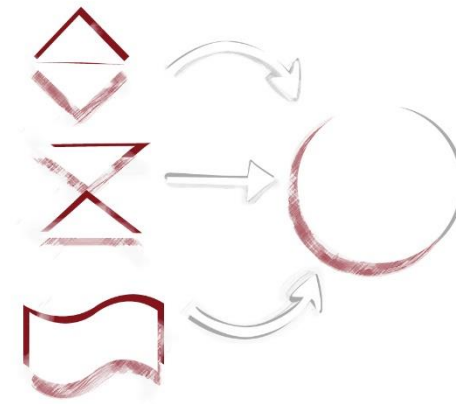


2019 Der Aktuar Data Science in der LV

- KI-Methoden sind nächster Evolutionsschritt
- Potenziell schwierige Datenlage
- Bestehende Daten nicht digital vorhanden

2019 Vorhersage der Arbeitsunfähigkeitsdauer

- Schwierigkeiten bei der Fallabgrenzung
- Denkbar, externe Daten anzuspielen
- Machine-Learning-Verfahren konnten den Mittelwert-Schätzer in den Gütemaßen nicht signifikant überbieten



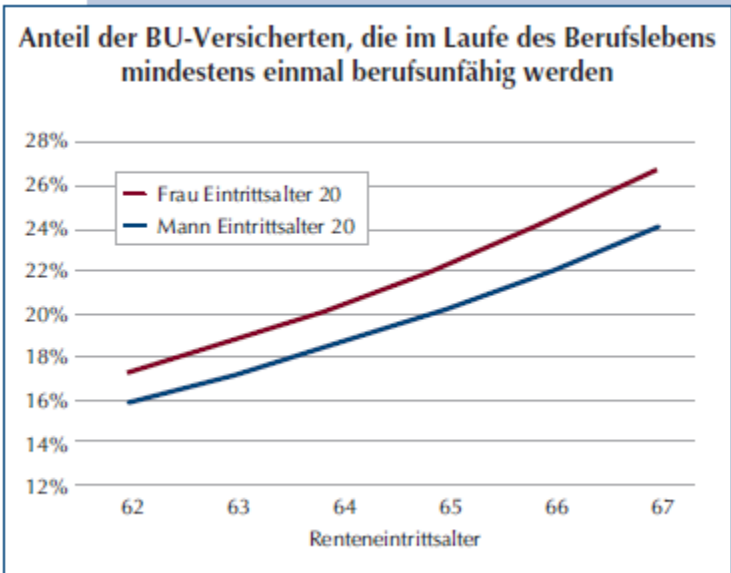
Der Aktuar 02.2021

- Machine Learning in der Berufsunfähigkeitsversicherung?
- Christian Eckert, Daniela Giesinger, Felix Müller und Antonia Schöning
- Öffentliche Daten vom National Center for Health Statistics (USA)
- Schlafverhalten, Alter und Rauchverhalten haben großen Einfluss
- Mehrwert von Resampling-Verfahren
- Logistische Regression ähnliche oder bessere Ergebnisse Random Forests oder Gradient Tree Boosting

Der Aktuar 04.2022

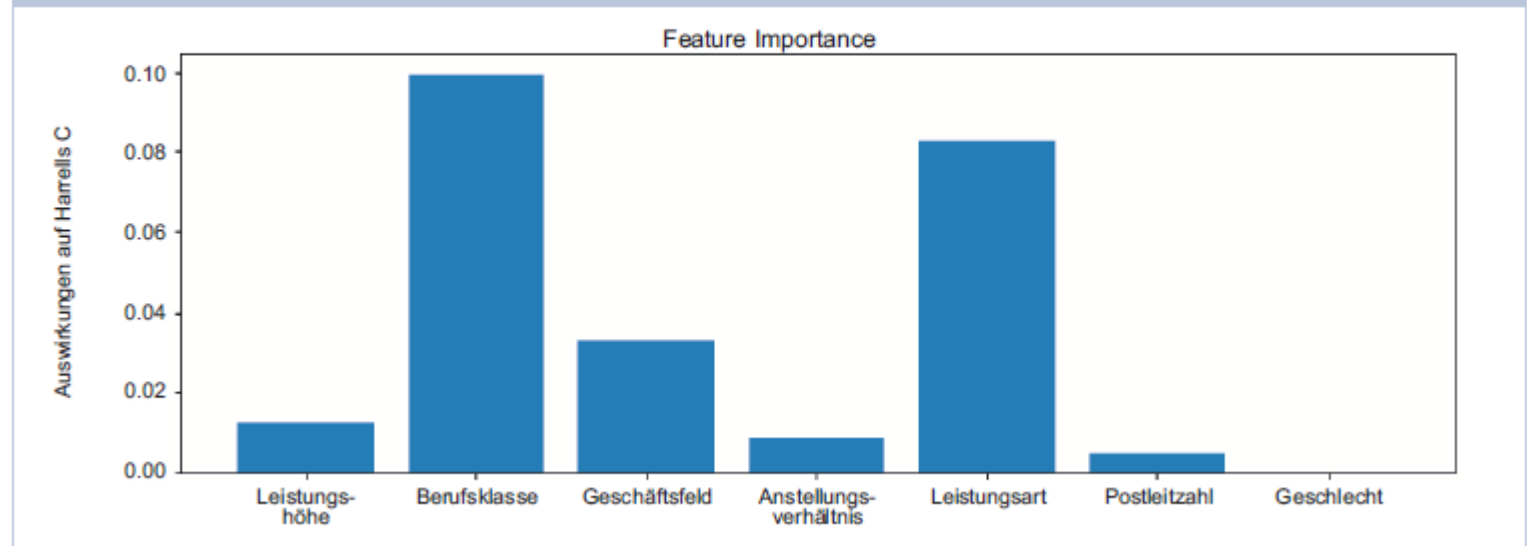
- Survival Trees für Invalidisierungswahrscheinlichkeiten in der Berufsunfähigkeitsversicherung
- Prof. Dr. Annegret Weng, Dr. Gang Feng und Dr. Felix Jörder
- Fiktiver Beispielbestand aus etwa 200.000 Datensätzen
- Der Datensatz enthält 7 Merkmale mit wenigen Ausprägungen – Nur Illustration
- „Survival Trees und Random Survival Forest eine interessante Ergänzung der bisher verwendeten Tools“

Berufsunfähigkeit:
die unterschätzte Gefahr?!



Aus der Aktuar 04.2022 Survival Trees für Invalidisierungswahrscheinlichkeiten in der Berufsunfähigkeitsversicherung

Abbildung 6:
Feature Importance

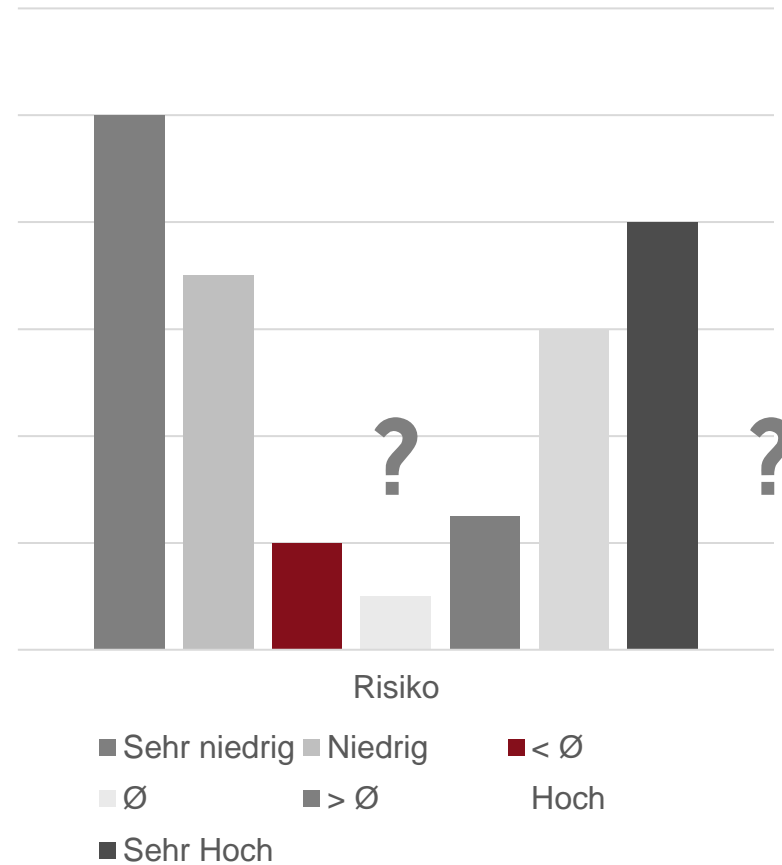


DAV 2021 I

- Versichertendaten
- Beobachtungszeitraum 2011-2015
- 59 Mio. Beobachtungsjahre
- 155.000 Neuberufsunfähige
- Marktabdeckung 85%



Typische Verteilung ?





- Nur etwa die Hälfte der Daten enthält valide Angaben zum Beruf.
- Regionale Besonderheiten werden nicht berücksichtigt.
- Keine konsistente Angabe der Rentenhöhe
- „Nur Beitragsbefreiung“ wird ignoriert



Nichts Genaues weiß man nicht !



Lücken mit KI-nahen Methoden schließen

Findet man Lücken ?



Entwicklung eines neuen KFZ-Pricing-Modells auf Basis ergänzender Kundendaten:



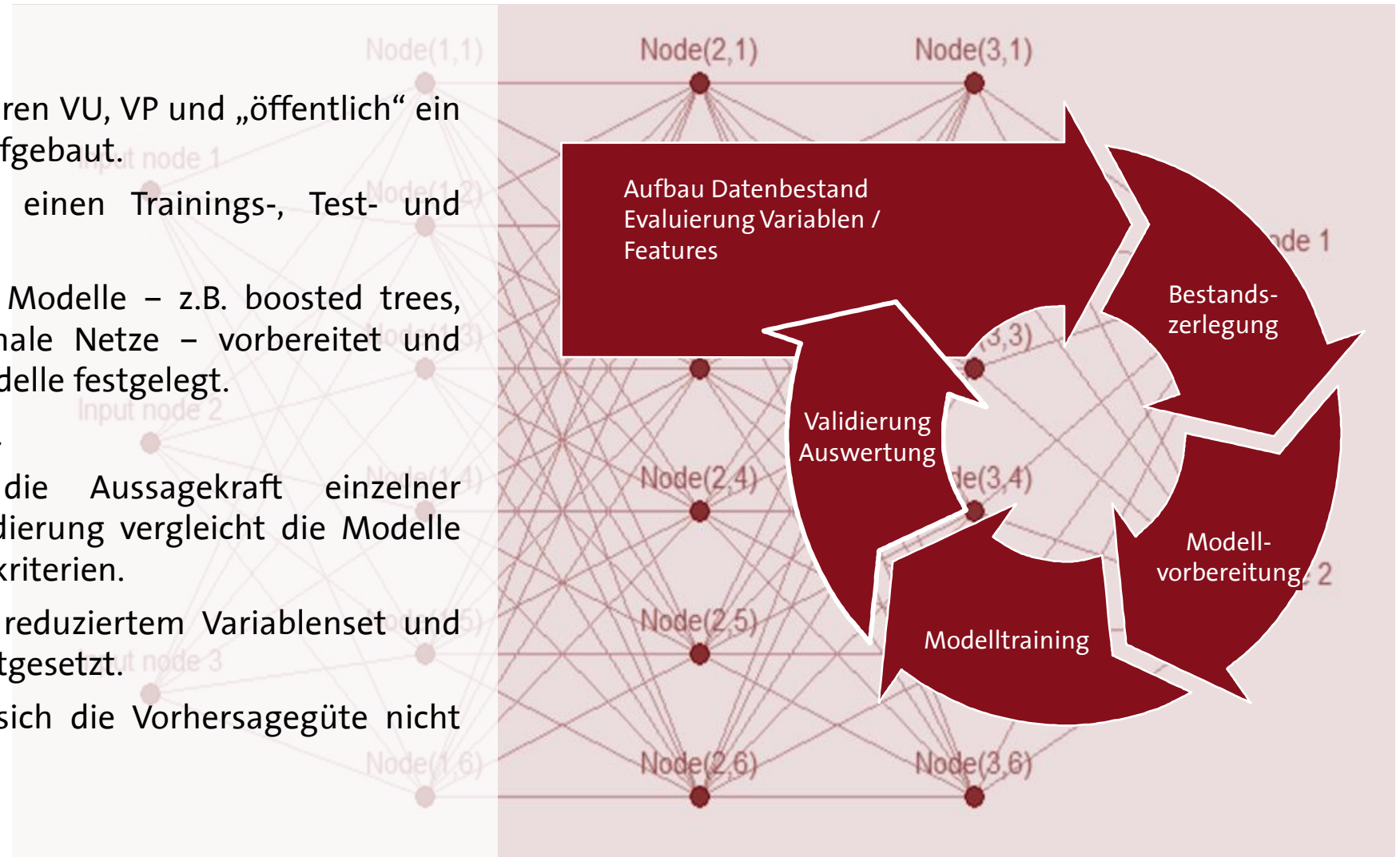
Funktioniert das auch für Berufsunfähigkeit?

Unsere Beratung:

- Konzeption eines aus integrierten Kanälen für Online- und Service-Kanäle
- Produktdesign entlang der Bedürfnisse des Vertriebswegs sowie der Kunden
- Koordination und inhaltliche Begleitung der notwendigen juristischen und datenschutzrechtlichen Klärungen zwischen den beteiligten Partnern und Dienstleistern
- Impulsgeber und Sparringspartner im Projekt, der Linie und der Organisation
 - zur strategischen Weiterentwicklung aktueller Verfahren
 - bei der (Weiter-) Entwicklung digitaler Vertriebswege und -prozesse
- Moderator und Coach
 - bei der Ausgestaltung und Verhandlung von Kooperations- und Dienstleistungsmodellen
 - bei der Aufstellung der Prozesse für ein Prototyping sowie Schulung

Mit KI-Methodik Zusammenhänge identifizieren

- Zu Beginn wird aus den Sphären VU, VP und „öffentlich“ ein erweiterter Datenbestand aufgebaut.
- Der Datenbestand wird in einen Trainings-, Test- und Validierungsbestand zerlegt.
- Es werden unterschiedliche Modelle – z.B. boosted trees, random forrest und neuronale Netze – vorbereitet und Kriterien für die Güte der Modelle festgelegt.
- Die Modelle werden trainiert.
- Die Auswertung misst die Aussagekraft einzelner Variablen/Features. Die Validierung vergleicht die Modelle anhand der definierten Gütekriterien.
- Der Kreislauf wird mit ggf. reduziertem Variablenset und neuer Bestandszerlegung fortgesetzt.
- Der Kreislauf endet, wenn sich die Vorhersagegüte nicht mehr signifikant verbessert.





Fachliche Expertise

- Entwickler
- Nutzer
- Management

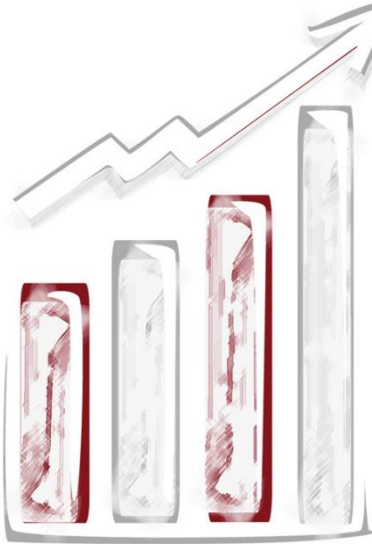
Kontrolle

- Rollen und Verantwortlichkeiten
- Kontrollen im Prozess
- Kultur

Validierung

- Aktives, regelmäßiges Überprüfen der Ergebnisse
- Unabhängigkeit von Entwicklung und Prüfung

Pricing LV
Hochrisiko



Vulnerabilitäten	Auswahl Beispiele
Merkmale	Altersgruppe, geringes Einkommen oder Armut, niedriger Bildungsstand, Migranten
Lebensereignisse	Arbeits- / Wohnungslosigkeit, Geschieden/alleinerziehend, überschuldete Personen, Unfallopfer, Opfer häuslicher Gewalt
Gesundheitszustand	Menschen mit Behinderungen, erbliche Erkrankungen (z.B. genetisch bedingt), Menschen mit psychischen Erkrankungen
Digitale Fähigkeiten	Wenig ausgeprägte digitale Fähigkeiten, Schwierigkeiten beim Zugang zu Online-Dienstleistungen/Digitaldienste

Rein fiktiv



Randverteilungen orientiert an Testbestand

- Geschlecht, Selektion, Dauern
- Beruf ohne Korrelation

Öffentliche Daten

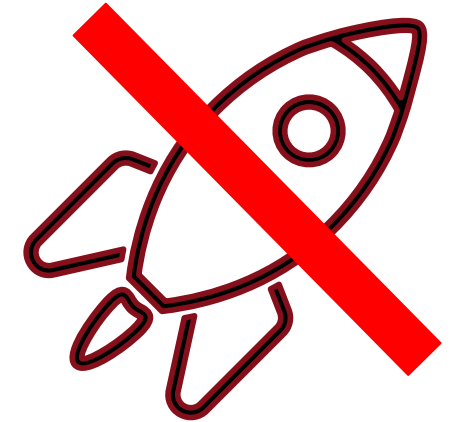
- Bundesland und PLZ (2-Stellig)
- Durchschnittseinkommen der Branche

Unberücksichtigt

- „Nur Beitragsbefreiung“, Arbeitgeber
- Kollektiv
- Neugeschäft

A³ Testbestand - Datenfelder

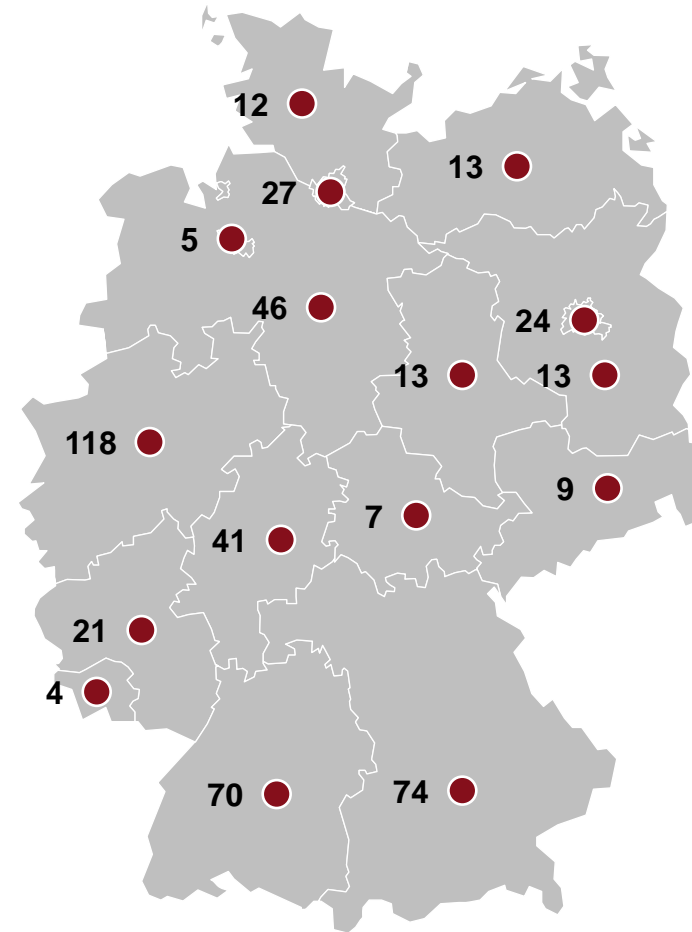
Feld	Ausprägung	Verteilung
Tarif	BUBR Risiko/BUBR sonst, BUB, SBU	Wie Poolbestabd
Geschlecht	m/W	Wie Poolbestand
Meldejahr	2011 -2013	Wie Poolbestand
Aktuelles Alter	Ganzjährig 15-66	Wie Poolbestand
Arbeitgeber	6 Klassen	Beschäftigte nach Unternehmensgröße
Bundesland	1 - 16	Häufigkeit sozialversicherungspflichtige AN
PLZ	2-Stellig	Wohnsitz
Beruf	20 häufigste Berufe + Unbekannt	Wie Poolbestand
Dauer	Max 35 Jahre	Wie Poolbestand
Endalter	Bis 68	Wie Poolbestand
Selektion	1 – 10+	Wie Poolbestand
Verdienst	5 Klassen	Stark überdurchschnittlich bis stark unterdurchschnittlich relativ zur Branche
Zustand	Aktiv, BU, Tod, Storno, Ablauf	Alle Personen starten als Aktiv, Reaktivierung tritt nicht auf



MS Excel
R-Studio

Fiktiver Bestand orientiert am Poolbestand

- Ca 140.000 Datensätze
- 510 BU-Fälle – Stochastisch simuliert
- Inzidenzen auf dem Niveau der 2. Ordnung DAV
- 508,6 BU-Fälle erwartet – Das Modell ist scheinbar i.O.





Modellannahmen zur BU

Modellannahme	Verdienst			Bundesland		
	Ausprägung	Anteil am Bestand	Inzidenz	Name	Anteil am Bestand	Inzidenz
Wer mehr verdient, wird seltener berufsunfähig	Stark überdurchschnittlich	5%	-5,0%	Baden-Württemberg	13,85%	-10%
	überdurchschnittlich	10%	-2,5%	Bayern	16,89%	-9%
	Durchschnittlich	70%	±0,0%	⋮	⋮	
Wer länger lebt, wird seltener berufsunfähig	Unterdurchschnittlich	10%	+2,5%	Sachsen-Anhalt	2,55%	+9%
	Stark unterdurchschnittlich	5%	+5,0%	Thüringen	2,50%	+10%

$$E(\text{Modell}) = E(\text{Pool})$$

$$V(\text{Modell}) \approx V(\text{Pool})$$

Erkennt ein Random Forrest die Effekte, oder sieht man den Wald vor lauter Bäumen nicht?



Der Zufall ist stärker als erwartet!

Modellannahme

Wer mehr verdient, wird seltener berufsunfähig

Wer länger lebt, wird seltener berufsunfähig

Land	Erwartete BU	Streuung	Testfall 1	Testfall 2	Testfall 3	Testfall 3	Testfall 4	Testfall 5
Schleswig-Holstein	18,2	66% - 126%	23	16	21	18	19	12
Hamburg	11,9	42% - 227%	14	8	9	14	5	27
Niedersachsen	48,2	71% - 118%	44	43	36	57	34	46
Bremen	3,7	109% - 163%	6	6	4	5	5	5
Nordrhein-Westfalen	108,4	94% - 111%	105	120	102	113	109	118
Hessen	38,9	95% - 121%	46	47	46	37	39	41
Rheinland-Pfalz	24,4	74% - 115%	20	25	27	28	18	21
Baden-Württemberg	68,8	92% - 121%	83	67	79	67	63	70
Bayern	85,1	78% - 99%	84	84	83	83	66	74
Saarland	5,8	52% - 121%	6	3	6	3	7	4
Berlin	21,9	105% - 137%	25	23	30	23	25	24
Brandenburg	14,2	56% - 127%	9	18	8	18	14	13
Mecklenburg-Vorpommern	10,2	79% - 128%	8	11	11	10	12	13
Sachsen	23,6	63% - 127%	23	28	15	30	26	21
Sachsen-Anhalt	13,1	61% - 107%	9	10	8	14	12	12
Thüringen	12,3	73% - 179%	18	22	15	13	11	9
Deutschland	508,6	91% - 105%	523	531	500	533	465	510

Erkennt ein Random Forrest die Effekte, oder sieht man den Wald vor lauter Bäumen nicht?

Modell

- BU Eintritt korreliert mit Lebenserwartung im Bundesland $\pm 10\%$
- BU Eintritt hängt ab vom relativen Verdienst $\pm 5\%$



Max und Moritz, gar nicht träge,
Sägen heimlich mit der Säge,
Ritzeratze! voller Tücke,
In die Brücke eine Lücke.

Simulation

- Aus der Hubschrauberperspektive trifft die Simulation das Modell
 - Verteilung Verträge auf Bundesland
 - Verteilung BU-Fälle auf Meldejahre
- Random Forrest, GBM und Neuronales Netz reagieren je nach Simulation zwar unterschiedlich, aber ihnen ist gemeinsam
 - Mal wird PLZ als wichtig erkannt, mal das Bundesland

**Mehr Daten
Noch mal sägen**

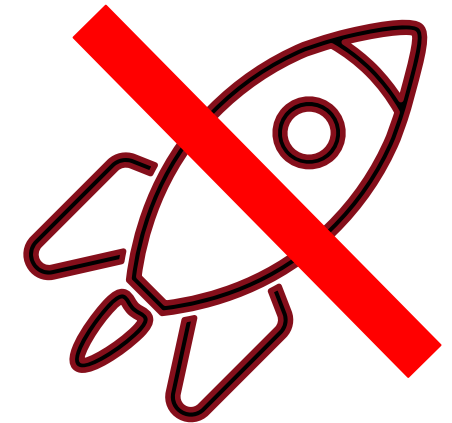
Fiktiver Bestand vergrößert und reduziert

- Ca 400.000 Datensätze
- Ca. 3.000 BU-Fälle – Stochastisch simuliert
- BU Eintritt hängt ab vom relativen Verdienst $\pm 10\%$
- Inzidenzen auf dem Niveau der 2. Ordnung DAV
Weniger Selektion, gealterte Verträge
- Übergang von Berufen auf Berufsgruppen
- Bilden von Altersgruppen
- (Weglassen von) PLZ und Bundesland



A³ Testbestand – Anpassung Datenfelder

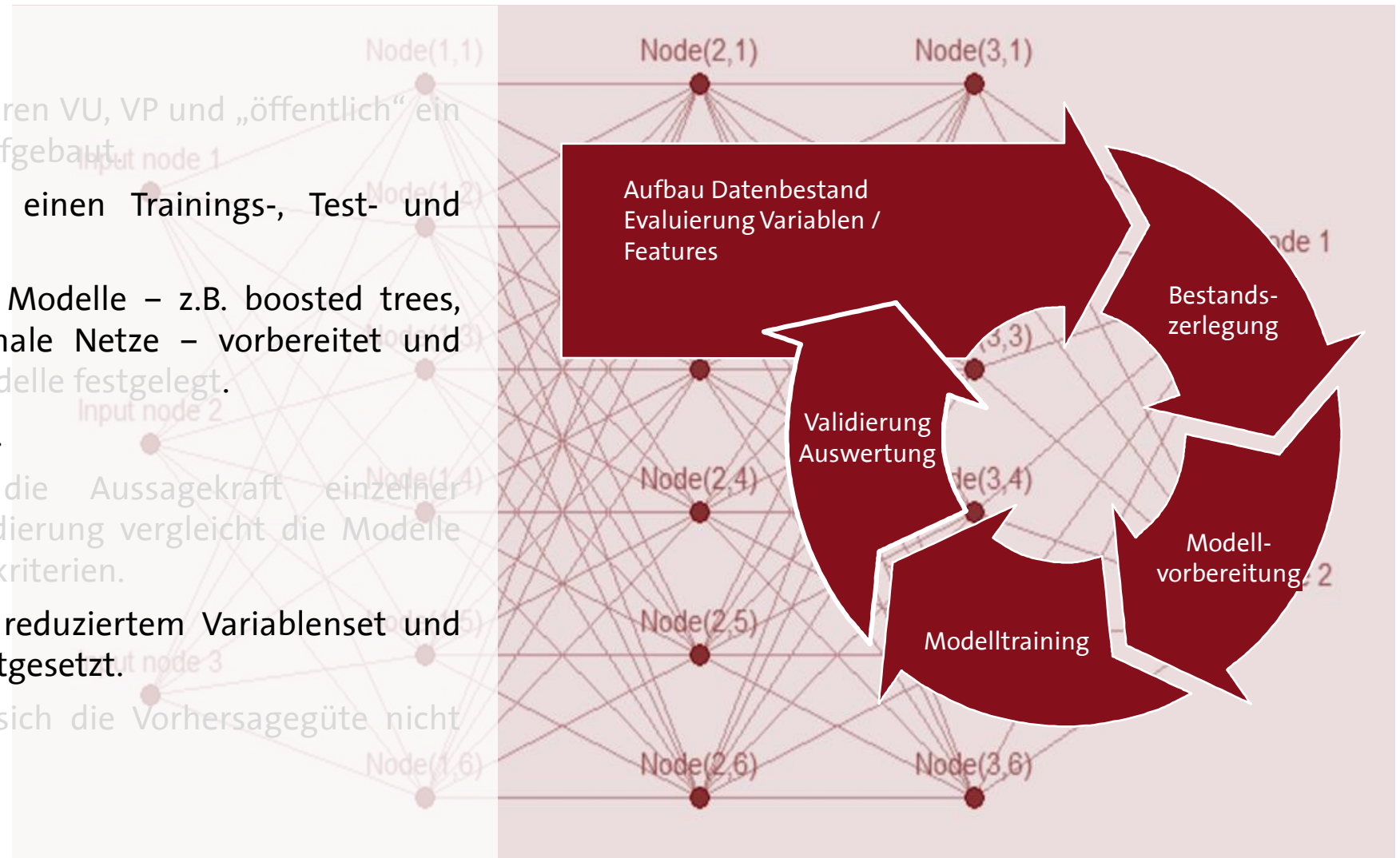
Feld	Ausprägung	Verteilung
Tarif	BUBR Risiko/BUBR sonst, BUB, SBU	Wie Poolbestabd
Geschlecht	m/w	Wie Poolbestand
Meldejahr	2011 -2013	Wie Poolbestand
Aktuelles Alter	Gealterter Bestand	Wie Poolbestand
Arbeitgeber	6 Klassen	Beschäftigte nach Unternehmensgröße
Bundesland	1 - 16	Entsprechend sozialversicherungspflichtige AN
PLZ	2-Stellig	Wohnsitz
Berufsklasse	5 Berufsklassen	Wie Poolbestand
Dauer	Max 35 Jahre	Wie Poolbestand
Endalter	Bis 68	Wie Poolbestand
Selektion	8 – 10+	Wie Poolbestand
Verdienst	5 Klassen	Stark überdurchschnittlich bis stark unterdurchschnittlich relativ zur Branche
Zustand	Aktiv, BU	Alle Personen starten als Aktiv, Reaktivierung tritt nicht auf



MS Excel
R-Studio

Mit KI-Methodik Zusammenhänge identifizieren

- Zu Beginn wird aus den Sphären VU, VP und „öffentlich“ ein erweiterter Datenbestand aufgebaut.
- Der Datenbestand wird in einen Trainings-, Test- und Validierungsbestand zerlegt.
- Es werden unterschiedliche Modelle – z.B. boosted trees, random forrest und neuronale Netze – vorbereitet und Kriterien für die Güte der Modelle festgelegt.
- Die Modelle werden trainiert.
- Die Auswertung misst die Aussagekraft einzelner Variablen/Features. Die Validierung vergleicht die Modelle anhand der definierten Gütekriterien.
- Der Kreislauf wird mit ggf. reduziertem Variablenset und neuer Bestandszerlegung fortgesetzt.
- Der Kreislauf endet, wenn sich die Vorhersagegüte nicht mehr signifikant verbessert.





Gradient Boosting Machine: Verdienst $\pm 10\%$, Bundesland $\pm 10\%$

■ Alle Variablen – Relative Information

■ AktAlter	58.9
■ Beruf	27.7
■ PLZ	4.8
■ Bundesland	2.5
■ Verdienst	2.2
■ AG	1.6
■ Tarif	1.3
■ GS	1.0

■ Ohne Bundesland – Relative Information

■ AktAlter	58.7
■ Beruf	30.9
■ PLZ	6.1
■ Verdienst	1.4
■ AG	1.3
■ GS	0.8
■ Tarif	0.8

■ Ohne PLZ – Relative Information

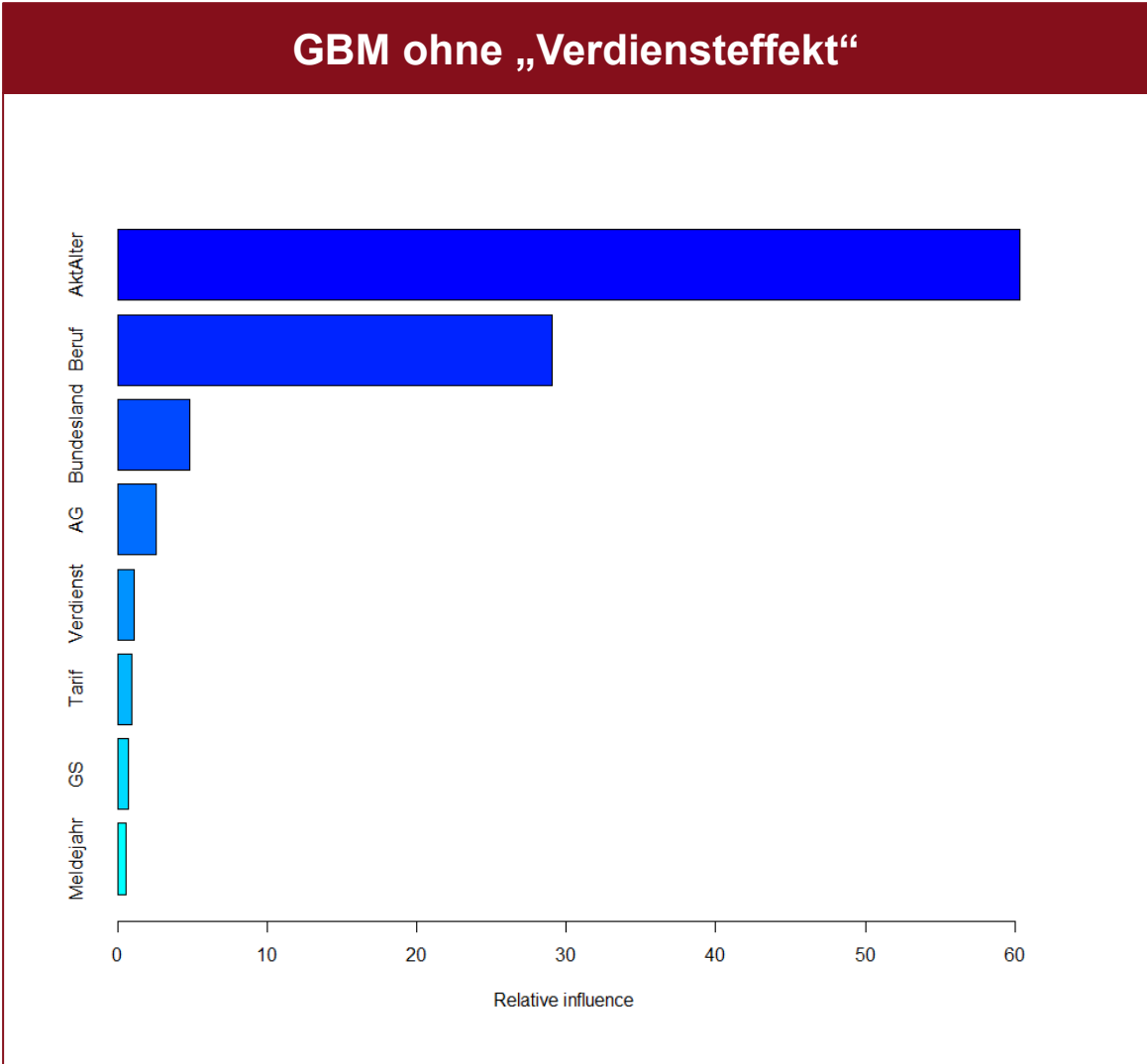
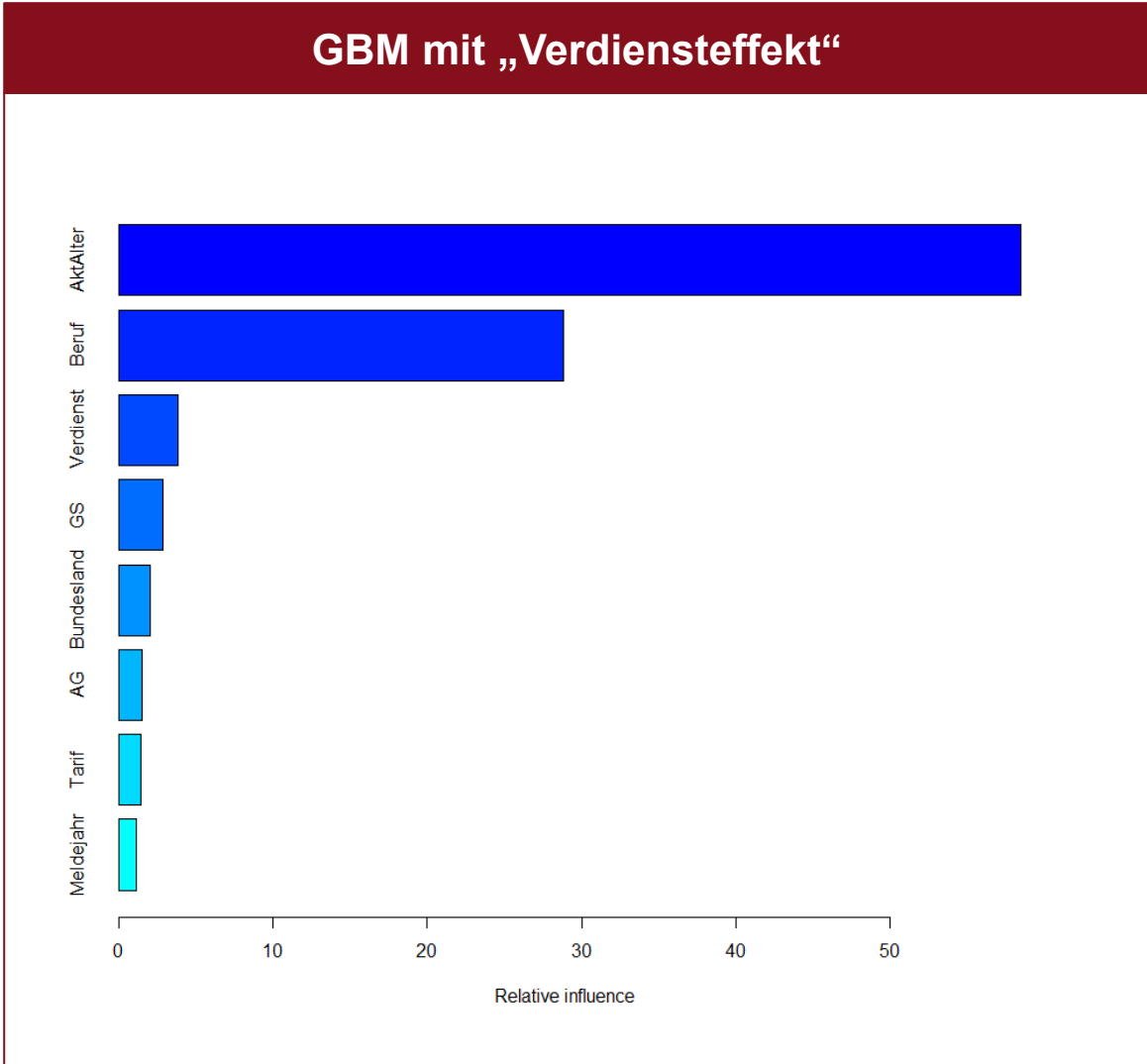
■ AktAlter	60.2
■ Beruf	28.8
■ Bundesland	4.0
■ Verdienst	2.4
■ AG	1.8
■ GS	0.9
■ Tarif	1.7

■ Ohne PLZ/BL – Relative Information

■ AktAlter	60.1
■ Beruf	31.8
■ Verdienst	2.4
■ AG	1.6
■ GS	1.2
■ Tarif	2.9



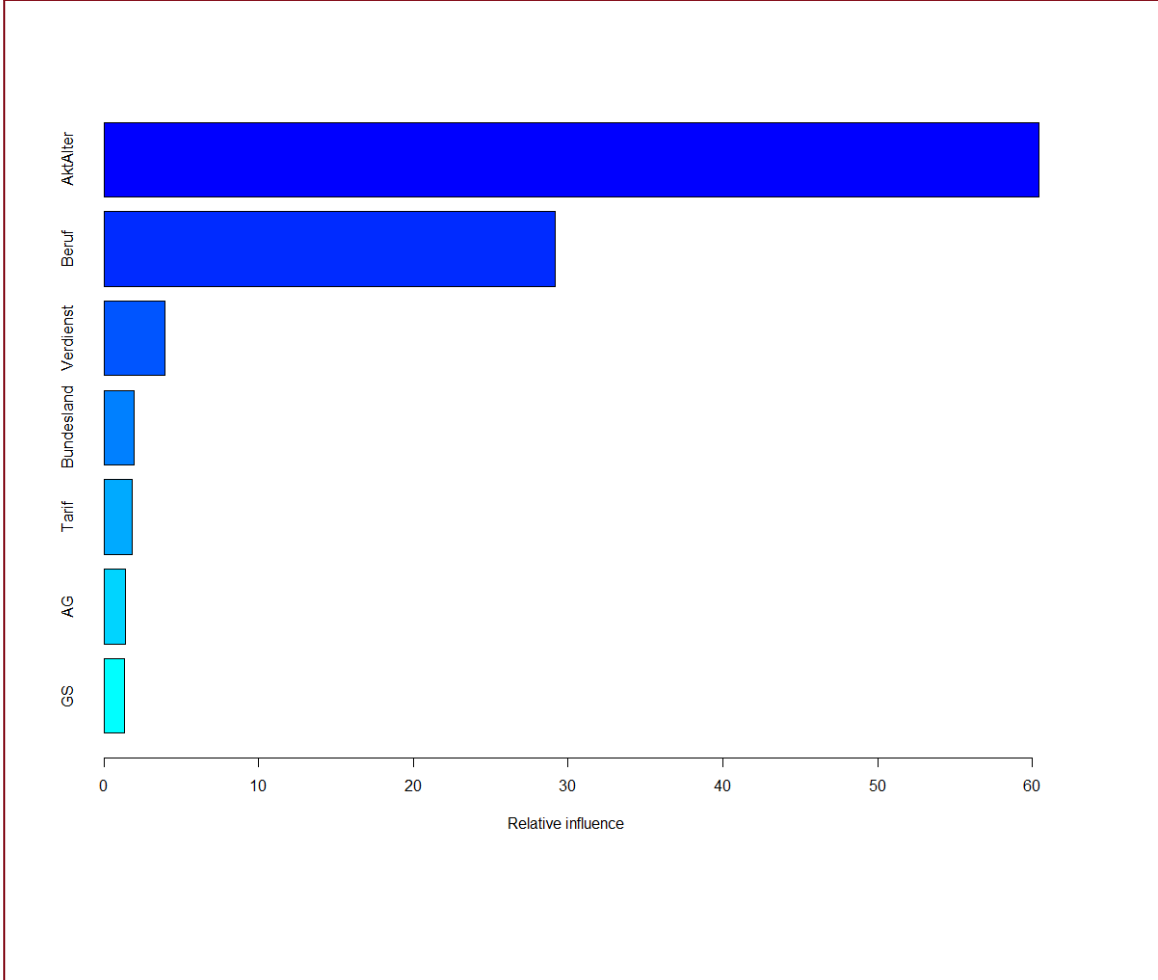
Gradient Boosting Machine – Ohne „Bundeslandeffekt“



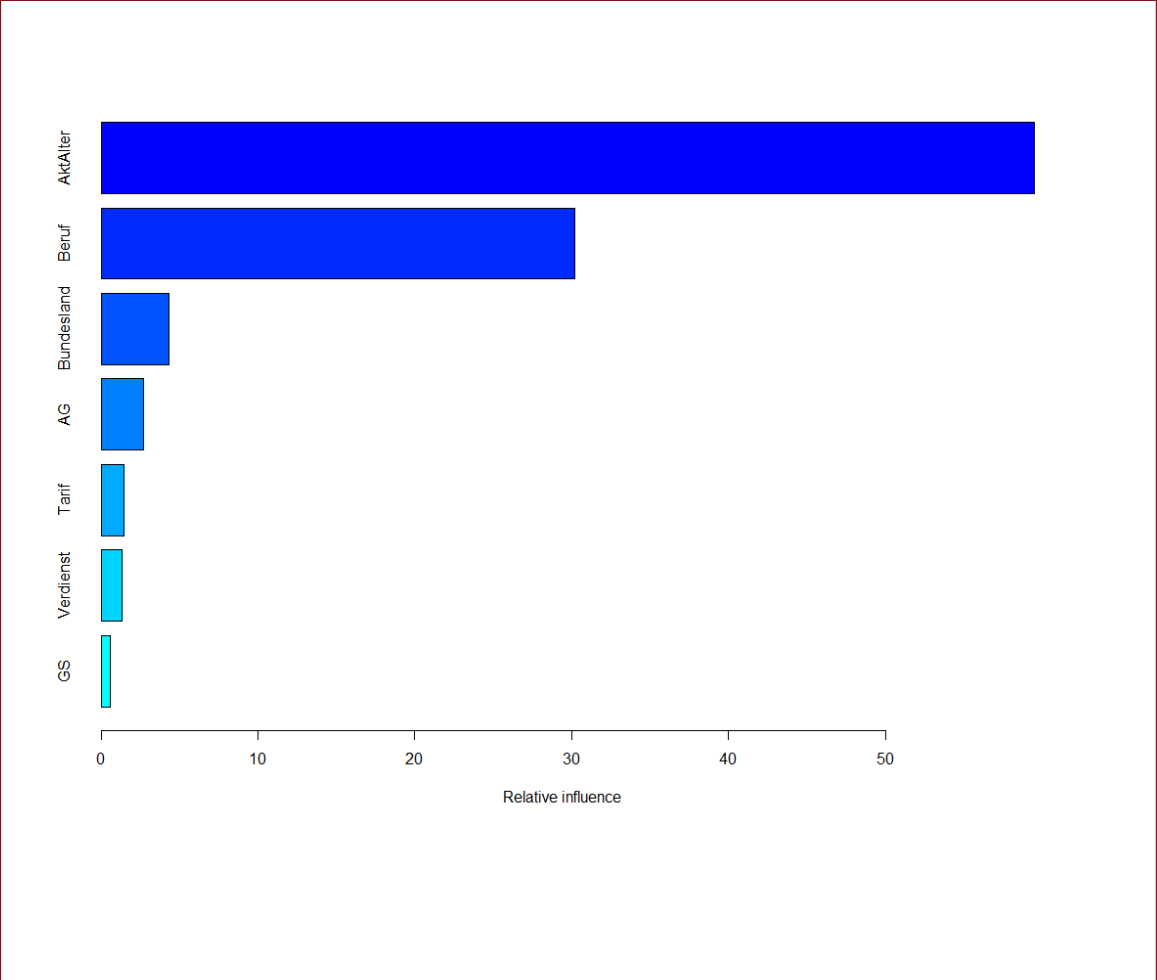


Gradient Boosting Machine – Ohne Meldejahr

GBM nur „Verdiensteffekt“



GBM ohne „Verdiensteffekt“





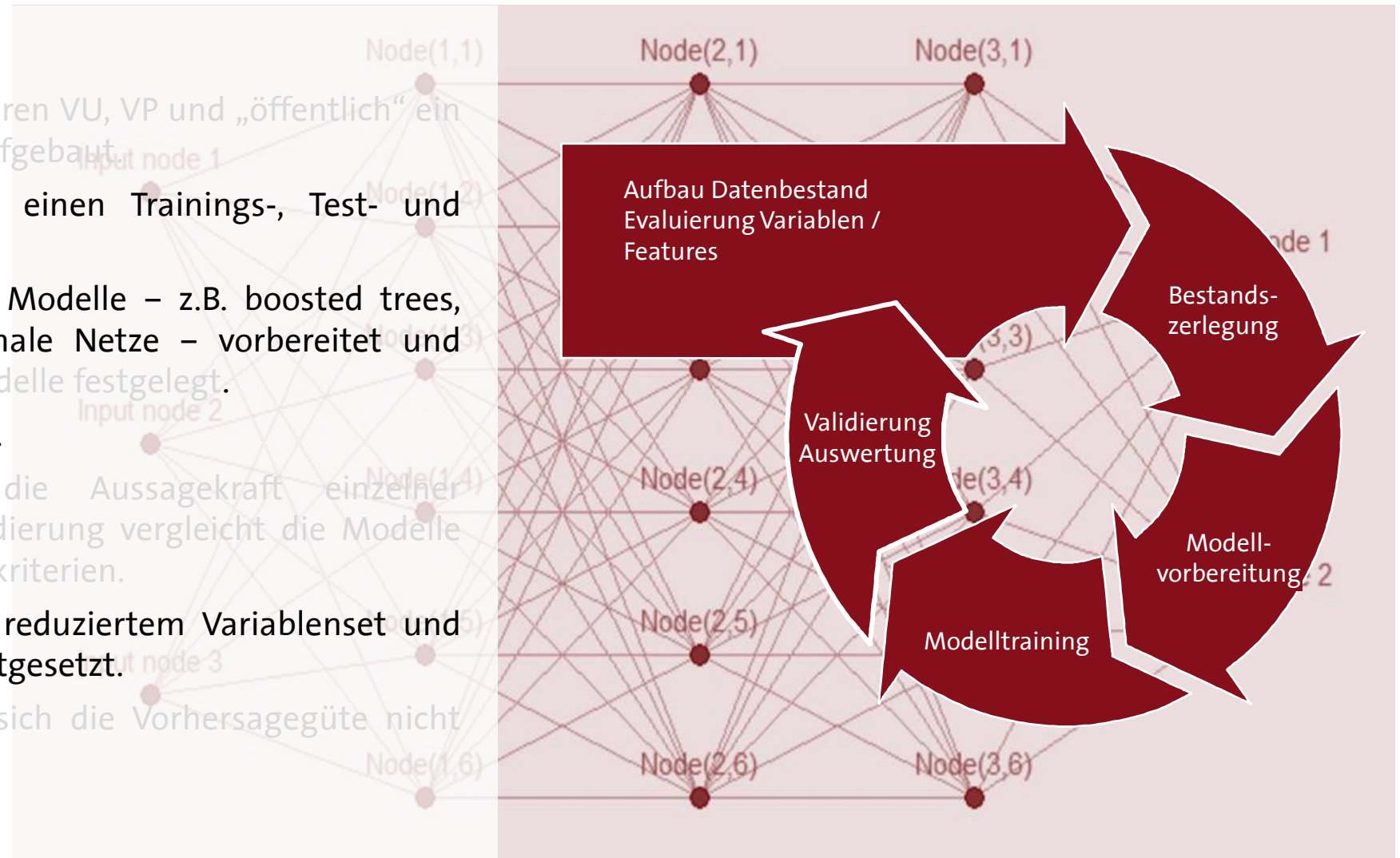
Der Zufall ist stärker als erwartet! Ist er stärker als das Gesetz der großen Zahlen?

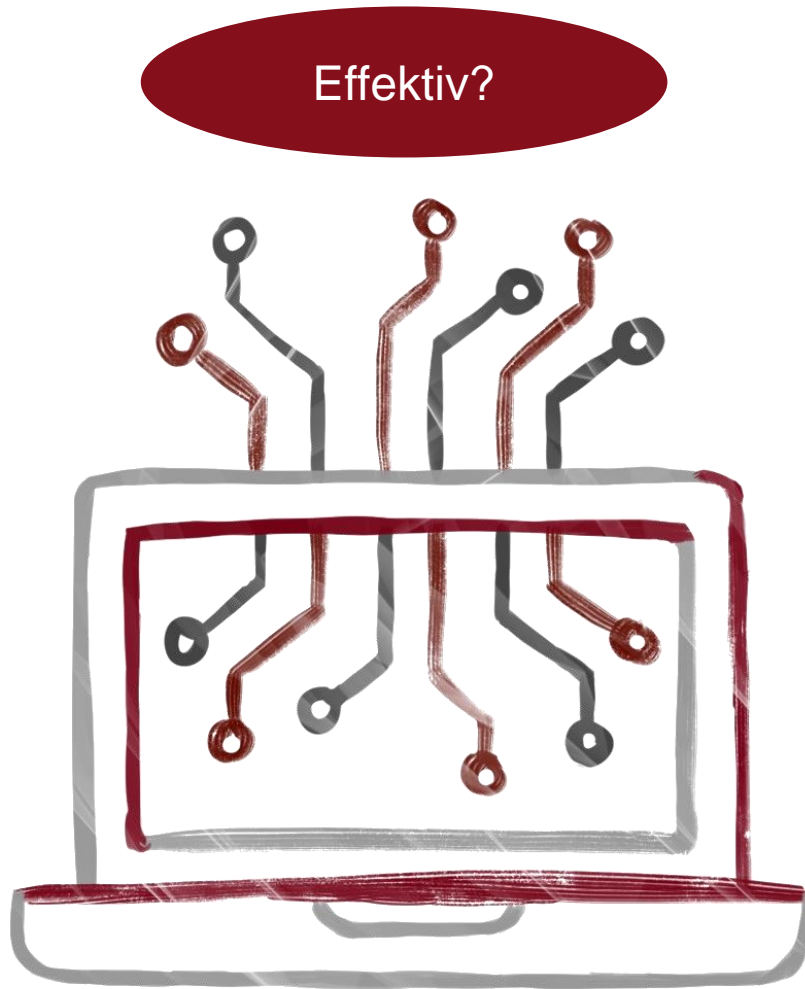
Struktur			400k		500K	
Verdienst	Vorgabe Anteil am Bestand	Vorgabe BU-Risiko	Modelliert	Eingetreten / Erwartet roh	Eingetreten / Erwartet roh	Eingetreten / Erwartet Modell
1	5%	-20%	5,0%	78,8%	76,3%	95,1%
2	10%	-10%	10,1%	95,9%	93,1%	103,6%
3	70%	±0%	70,1%	99,8%	100,2%	100,2%
4	10%	+10%	9,8%	113,7%	112,8%	102,5%
5	5%	+20%	5,0%	129,8%	125,0%	104,0%
Gesamt	100%		100%	101,2%	100,8%	100,8%

Erkennt ein Random Forrest die Effekte,
oder sieht man den Wald vor lauter Bäumen nicht?

Mit KI-Methodik Zusammenhänge identifizieren

- Zu Beginn wird aus den Sphären VU, VP und „öffentlich“ ein erweiterter Datenbestand aufgebaut.
- Der Datenbestand wird in einen Trainings-, Test- und Validierungsbestand zerlegt.
- Es werden unterschiedliche Modelle – z.B. boosted trees, random forrest und neuronale Netze – vorbereitet und Kriterien für die Güte der Modelle festgelegt.
- Die Modelle werden trainiert.
- Die Auswertung misst die Aussagekraft einzelner Variablen/Features. Die Validierung vergleicht die Modelle anhand der definierten Gütekriterien.
- Der Kreislauf wird mit ggf. reduziertem Variablenset und neuer Bestandszerlegung fortgesetzt.
- Der Kreislauf endet, wenn sich die Vorhersagegüte nicht mehr signifikant verbessert.





Neuronales Netz vergleichsweise rechenintensiv

- Vielzahl an möglichen Parametern wie Layer, Algorithmus, Aktivierungsfunktion, Fehlerfunktion, Backpropagation macht das neuronale zu einem sehr variablen Tool
- Nachteil: „Optimales“ Netz für diese Fragestellung noch nicht gefunden

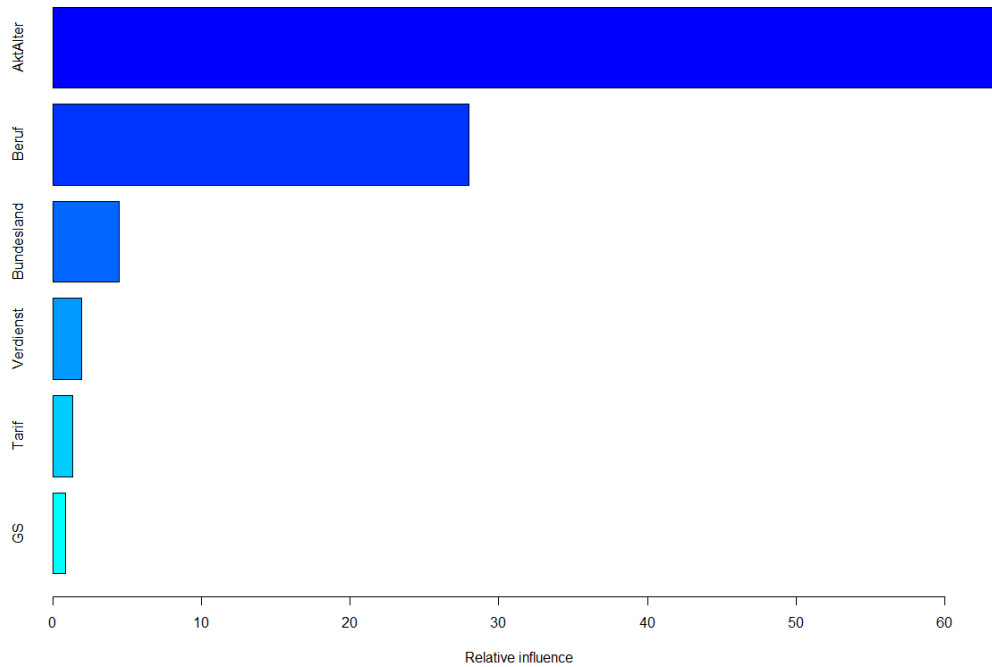
Die Abhängigkeit von „Verdienst“ wird nicht problemlos erkannt

- Die Performance wird schnell besser, wenn Variablen ohne Einfluss im Modell nicht berücksichtigt werden
 - Tarif, Arbeitgeber, PLZ



Gradient Boosting Machine – „survivor bias“

GBM



Bewertung

- Der Effekt durch das Bundesland wird hinreichend gut erkannt
- Der Verdienst beeinflusst zu viele Datensätze überhaupt nicht, sodaß der Effekt nicht sicher identifiziert werden kann
- Das Geschlecht bleibt mir ein Rätsel

Neuronales Netz

- Downsampling und Gewichte liefern im ersten Lauf in etwa gleich (schlechte) Ergebnisse
- Für 400k Datensätze wurde ein gutes Modell für diese Fragestellung noch nicht gefunden.
- Für 500k Datensätze wurde ein gutes Modell für diese Fragestellung noch nicht gefunden.

Random Forrest

- Downsampling und Gewichte liefern im ersten Lauf in etwa gleich (schlechte) Ergebnisse
- Für 400k Datensätze verbessert sich die Klassifikation etwas
Downsampling für diese Fragestellung von Vorteil
- Für 500k Datensätze verbessert sich die Klassifikation nicht wirklich
Downsampling für diese Fragestellung von Vorteil

GBM

- Downsampling und Gewichte liefern im ersten Lauf in etwa gleich (schlechte) Ergebnisse
- Für 400k Datensätze verbessert sich die Klassifikation etwas
Downsampling für diese Fragestellung von Vorteil
- Für 500k Datensätze verbessert sich die Klassifikation nicht wirklich
Downsampling für diese Fragestellung von Vorteil

(1) Das Ergebnis wird als gut angesehen, wenn es Bundesland und Verdienst als Risikotreiber identifiziert.



Bonus – Modell Nur „Verdiensteffekt ± 20 “

Modellannahmen

- Neben Alter, Geschlecht und Beruf hat nur der Verdienst Auswirkungen auf die BU
 - Stark überdurchschnittlich -20%
5% des Bestandes
 - Überdurchschnittlich -10%
10% des Bestandes
 - Durchschnittlich $\pm 0\%$
70% des Bestandes
 - Unterdurchschnittlich +10%
10% des Bestandes
 - Stark unterdurchschnittlich +20%
5% des Bestandes

Erkenntnisse

- Neuronales Netz
 - Der Effekt wird schon von einem einfachen neuronalem Netz erkannt
 - Das neuronale Netz scheint mehr Probleme mit Variablen zu haben, die keinen Beitrag leisten, als die anderen Modelle (Tarif, Arbeitgeber)
- Random Forrest und GBM
 - Beide Werkzeuge liefern gute Ergebnisse
 - Verdienst wird schnell als Einflussgröße identifiziert



Wann wird ein Effekt erkannt?

- Der Effekt muss die abhängige Variable in der Höhe ausreichend beeinflussen
- Der Effekt muss ausreichend häufig im Bestand auftreten
- Der Effekt muss „meßbar“ sein

Wann wird ein Effekt nicht oder falsch erkannt?

- Das Modell passt nicht
- Zu viele Merkmale
- Zu viele Ausprägungen eines Merkmals
- Merkmale ohne Einfluss.

```
warnmeldung:  
Algorithm did not converge  
> |
```



Thomas Binder

Geschäftsführender Gesellschafter

AAA Auctor Actor Advisor GmbH

Nymphenburger Str. 4 | 80335 München | www.3-a-3.de

Telefon +49 (0) 89 208 027 285 | Telefax +49 (0) 89 288 90 45

Mobil +49 (0) 163 333 88 90 | Email Thomas.Binder@3-a-3.de



Michael Stoffels

Principal Consultant

AAA Auctor Actor Advisor GmbH

Im Mediapark 8 | 50670 Köln | www.3-a-3.de

Telefon +49 (0) 221 55 40 54 11 | Telefax +49 (0) 221 55 40 54 5

Mobil +49 (0) 178 333 88 13 | Email Michael.Stoffels@3-a-3.de

