

Bewohnerklassifikation in der stationären Langzeitpflege

unter Anwendung von
Regressionsverfahren des statistischen
Lernens in R



Christian Grebe, MScN

Hintergrund und Kontext

Bedarf für ein neues System

◎ SGB XI Pflegestufen

- **normativ**
- „Fallgewichte“ sind **Verhandlungsergebnis** ohne empirische Fundierung
- **inhomogen** und **zu wenig trennscharf** (Wingenfeld & Schnabel 2002)
- wesentliche Aspekte des **Leistungsspektrums** eines Pflegeheims sind für die Einstufung nicht relevant
- träge Reaktion auf Veränderungen (**Begutachtungsverfahren**)

Fragestellung und Methodik

Fragestellungen der Studie

- 1) Wie gut erklärt ein auf den Pflegestufen des SGB XI basierendes Modell den tatsächlichen zeitlichen Aufwand für Pflege und Betreuung von Pflegeheimbewohnern?
- 2) Lässt sich auf der Basis von bewohnerbezogenen Assessment-Daten zu Fähigkeiten und Einbußen (und ggf. Leistungen) ein Modell entwickeln, welches den tatsächlichen zeitlichen Aufwand für Pflege und Betreuung von Pflegeheimbewohnern besser erklärt als das Modell auf Basis der Pflegestufen des SGB XI?

Assessmentbogen

- ◉ **dichotome** Items
- ◉ 78 Items zu Bewohnercharakteristika, vor allem Fähigkeiten und Einbußen

12	Ist der Bewohner fähig, seine Lage im Bett umfänglich zu verändern? (<i>vom Rücken auf die Seite oder umgekehrt</i>)	☐ ja	☐ nein
23	Ist der Bewohner in der Lage, seinen Rücken und seine Beine selbstständig zu waschen?	☐ ja	☐ nein
28	Ist der Bewohner in der Lage, aus einem bereit gestellten Gefäß selbstständig zu trinken?	☐ ja	☐ nein
36	Ist der Bewohner in der Lage, selbstständig (<i>ggf. mit Hilfsmittel</i>) die Toilette zu benutzen?	☐ ja	☐ nein
54	Zeigt der Bewohner mehrmals pro Woche gedrückte, depressive oder ängstliche Stimmungslagen?	☐ ja	☐ nein
60	Ist die Orientierungsfähigkeit des Bewohners zeitlich, örtlich, situativ oder zur Person eingeschränkt?	☐ ja	☐ nein

Welche Zeiten wurden erfasst?

- bewohnernah

- Jeder Kontakt zwischen Personal und Bewohner, gleich aus welchem Anlass

- bewohnerfern

- Bestimmte Zeiten, die sich konkret auf einen Bewohner beziehen:
 - Arzt- und Angehörigengespräche
 - Dokumentation
 - Pflegeplanung
 - Fallbesprechungen

- nicht berücksichtigt

- bewohnerferne Zeiten, die sich nicht konkret auf einen Bewohner beziehen
- Wegezeiten
- allgemeine Administration, Übergaben und Besprechungen (ohne Bezug zu einem bestimmten Bewohner)

Zeiterfassung

- Messung über 48h je Bewohner (die Zeiten der beiden Tage wurden zu einem 24h- Wert gemittelt)
- Messen mittels (Stopp-) Uhren und Selbstaufschreibung
- Erfasst wurden
 - Uhrzeit und Dauer des Kontakts bzw. der bewohnerfernen Leistung
 - Anzahl der involvierten Mitarbeiter und deren Qualifikation(en)
 - Anzahl der involvierten Bewohner (Gruppenzeiten wurden dem betreffenden Bewohner anteilig zugerechnet)
 - Vorherrschende Leistungsart im jeweiligen Zeitabschnitt
- nachgeschaltete Plausibilitätsprüfung

Stichprobe

N= 594 Bewohner/-innen
aus 13 Pflegeheimen

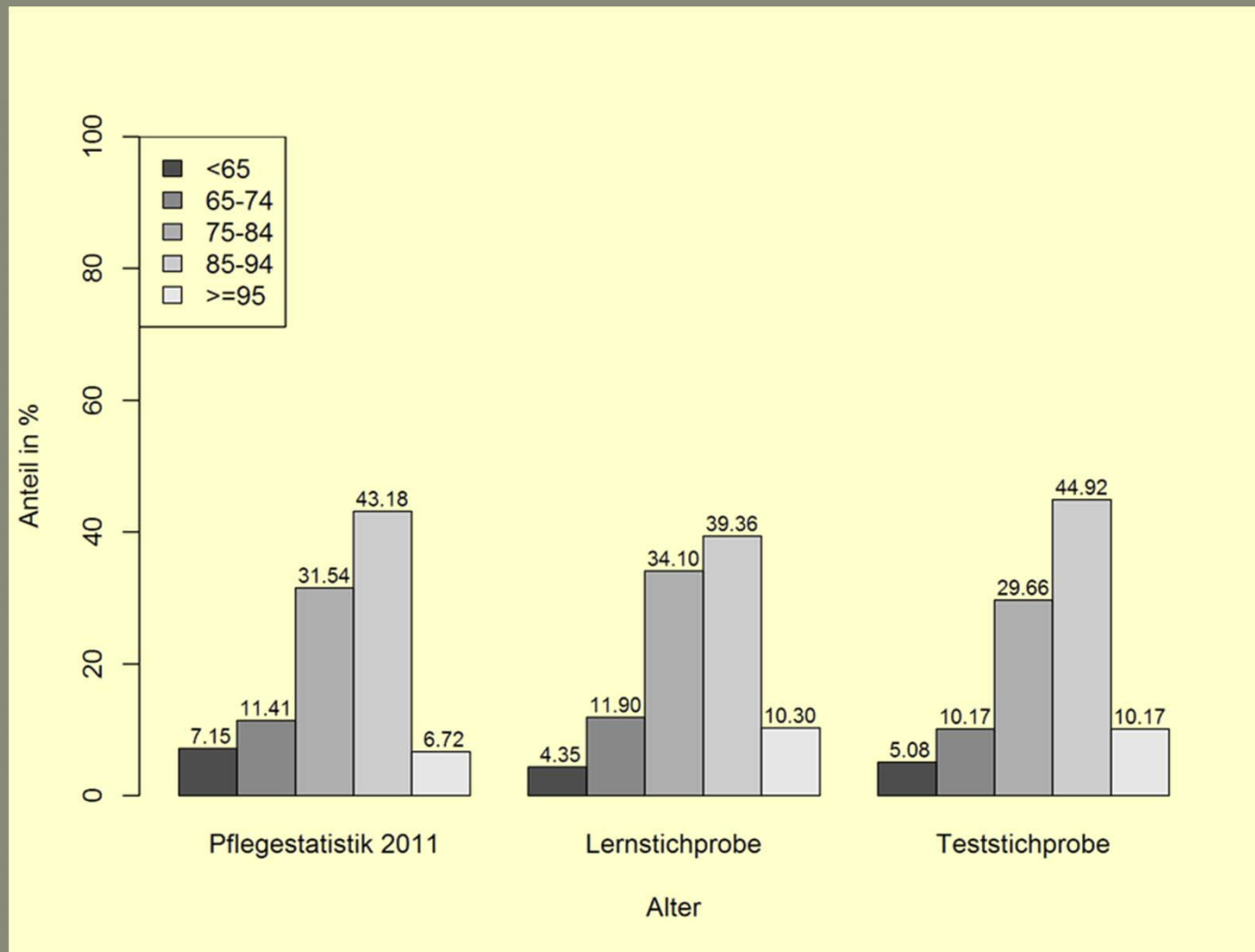
Fehlende Werte

- ◉ Die folgenden eingesetzten Algorithmen können (zumindest in ihrer R-Implementierung) nicht mit Missing Values umgehen):
 - MARS (earth), Random Forest (randomForest)
- ◉ CART (rpart) und cTree (party) verwenden **Surrogate Splits**
 - Für jeden Splitpunkt wird berechnet, welche anderen Splits zu einer möglichst gleichen Partitionierung führen (prozentuale Übereinstimmung)
 - Bei fehlendem Wert werden diese (ggf. der Reihe nach) als Split verwendet
- ◉ Gradient Boosting (gbm) baut bei Missing Values den jeweiligen Baum nicht weiter aus
- ◉ Vergleich der Modelle setzt in dieser Studie identische Lern- und Teststichproben voraus
 - Daher: Eliminierung aller Fälle mit Missing Values unter den Assessmentvariablen und/oder SGB XI- Variablen (38 Fälle)
 - Die **Datenanalyse basiert somit auf N=556 Fällen**

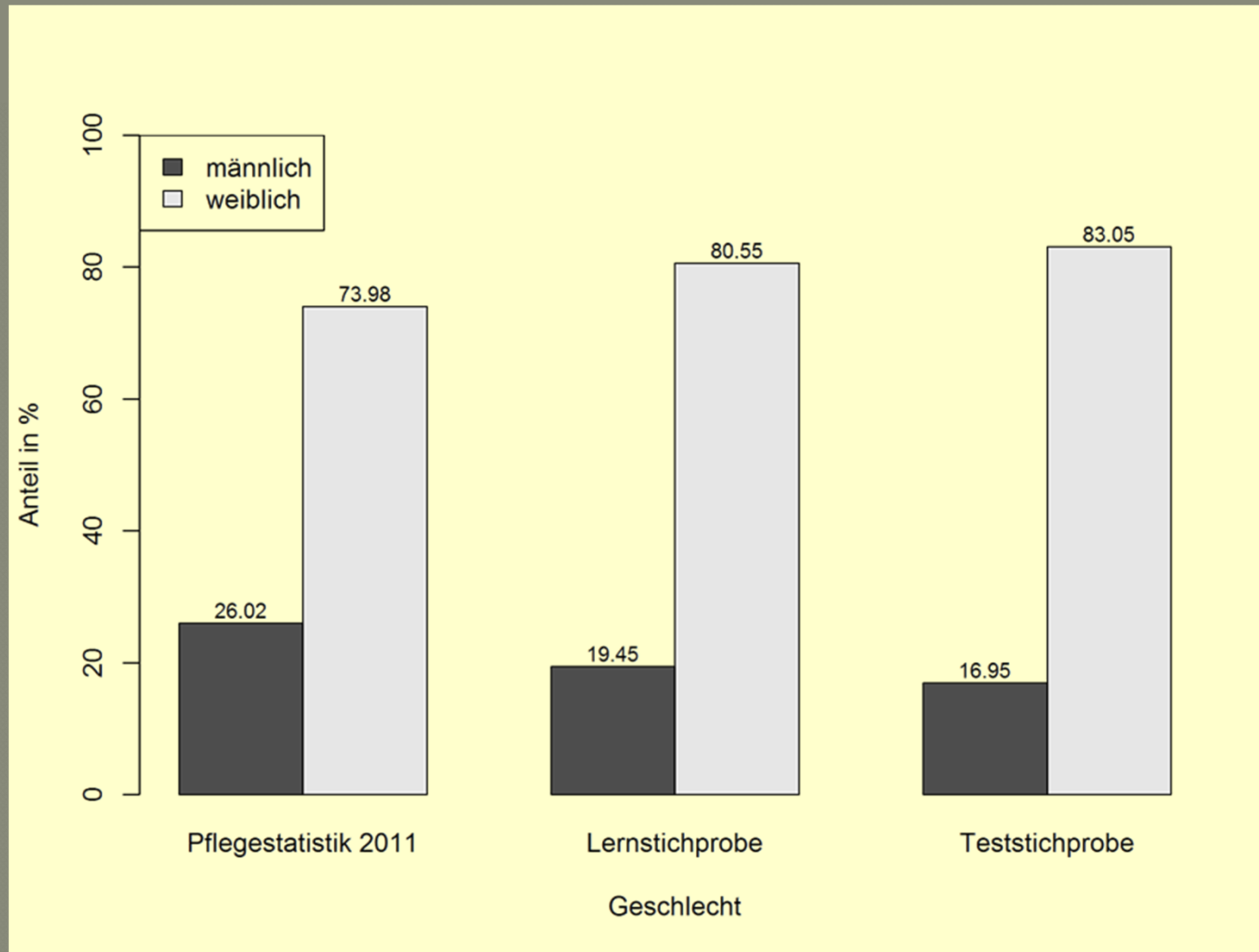
Lern- und Teststichprobe

- Vergleichsgröße der verschiedenen Modelle ist die Varianzreduktion (R^2) bezogen auf eine einheitliche Teststichprobe
- Die N=556 Fälle wurden daher randomisiert aufgeteilt in
 - Lernstichprobe (80%): N= 438
 - Teststichprobe (20%): N= 118
- Bei 14 der 76 Bewohnercharakteristika unterscheidet sich die Verteilung in der Teststichprobe signifikant von jener in der Lernstichprobe

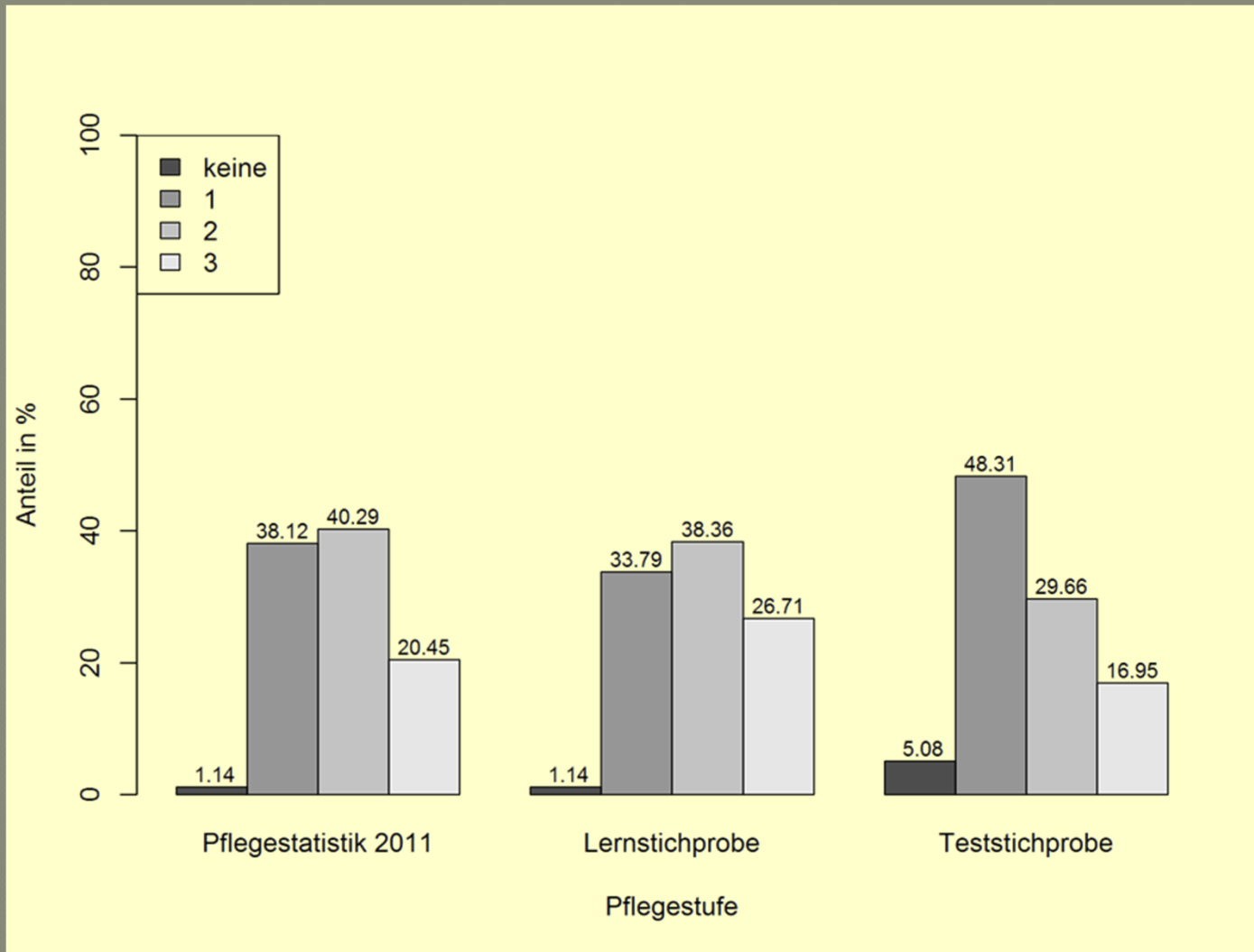
Stichprobe: Alter



Stichprobe: Geschlechter



Stichprobe: Pflegestufen



Modellierungs- Algorithmen

eingesetzte Verfahren des statistischen Lernens

Methode A: Entscheidungsbäume

- Es kamen 2 verschiedene Algorithmen zum Einsatz:
 - Classification and Regression Trees (**CART**) (Breiman et al. 1984)
 - Conditional Inference Trees (**cTree**) (Hothorn et al. 2010)

CART

◉ Grundlegender Algorithmus:

- geprüft werden alle möglichen Varianten, die Gesamtmenge der Fälle in 2 Gruppen zu partitionieren (**binäre Splits, exhaustiver Algorithmus**)
- gewählt wird jener Split, durch den die Residuenquadratsumme minimal wird
- **Abbruchkriterium** :
 - minimale Anzahl der Fälle im Knoten
 - minimale Anzahl der Fälle in den Tochterknoten
 - Mindestmaß an zusätzlicher erreichter Reduktion der Residuenquadrate
- **Pruning** auf den Teilbaum mit dem geringsten Kreuzvalidierungsfehler ($k=10$)

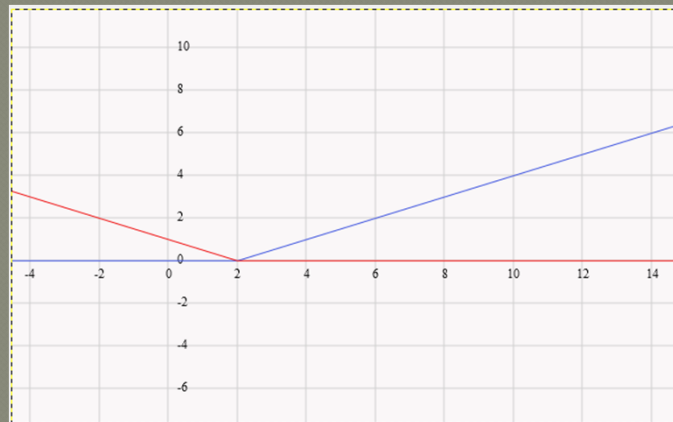
◉ R-Package: rpart

Unterschied cTree - CART

- ◉ cTree arbeitet bei der Auswahl der Splits mit Signifikanztests
- ◉ jeweils Test auf H_0 : Unabhängigkeit zwischen erklärender und abhängiger Variable (**Permutationstest**)
- ◉ Wenn die Nullhypothese nicht verworfen werden kann, stoppt der Algorithmus
- ◉ **kein Pruning** erforderlich
- ◉ weitere Abbruchkriterien wie bei CART (Anzahl der Fälle in Knoten) sind möglich
- ◉ R-Package für cTree: party

Methode B: MARS- Modelle (Friedman 1991)

- ähnlich dem linearen Regressionsmodell, an die Stelle der Kovariate treten aber Basisfunktionen (Hinge- Funktionen)
 - Paarweise hinzugefügte Hinge – Funktionen: $\max(0, x - t)$ bzw. $\max(0, t - x)$



- Auswahl analog zu CART (exhaustive Suche über alle möglichen Knotenpunkte, selektiert wird jene, die die Residuenquadratsumme minimiert)
- Abbruchkriterium: zusätzliche Varianzreduktion bleibt unter einem wählbaren Wert
- Anschließend Rückwärtsdurchlauf (Modellwahl: **generalisierter Kreuzvalidierung**)

Methoden C: Ensemblemethoden

◉ Random Forest (Breiman 2001)

- Modellierung einer großen Zahl von **vollausgebauten** CART- Bäumen (hier: 500)
- für jede Iteration: Randomisierung der Lernstichprobe und der Prädiktorvariablen

◉ Conditional Inference Forest (Hothorn et al. 2010)

- wie Random Forest, aber basierend auf cTree- Bäumen (hier: 500)
- Randomisierung

◉ Gradient Boosting (Friedman 2001)

- mit verschiedenen Verfahren kombinierbar, hier: mit CART- Bäumen (hier: 2500, jeweils mit nur einem Split = Stumps)
- **Gradientenverfahren**
- **Regularisierung**: Gewichtung mit Shrinkage- Parameter (hier: 0.01) drosselt die Lerngeschwindigkeit

Ergebnisse

Modelle auf der Basis von Methoden des maschinellen Lernens

Referenz- und Konkurrenzmodelle

◉ Referenzmodelle

- nutzen als erklärende Variable die SGB XI- Informationen, die im derzeitigen Status quo als Basis für die Personalschlüssel verwendet werden:
 - Pflegestufe
 - eingeschränkte Alltagskompetenz nach §87b (PEA)

◉ Konkurrenzmodelle

- wird als Bezeichnung für die neu entwickelten Modelle verwendet

Modelltypen

- ◉ Alle Modelle wurden für 2 Arten der abhängigen Variable (zeitlicher Pflegeaufwand) berechnet:
 - a) Gesamtzeiten
 - b) Nursing only (nur Grund- und Behandlungspflege, siehe nächste Folie)

- ◉ Zudem werden für die Konkurrenzmodelle 2 weitere Arten unterschieden:
 - a) ausschließlich zustandsbezogene Assessment-Items
 - b) Zusätzlich 3 Items mit engem Bezug zu Zeitaufwand:
 - aufwändiges Anreichen des Essens
 - Teilnahme an Tagesgruppe
 - mehrmals wöchentlich soziale Betreuung)

Modelle für „Nursing only“

- Für diese Modelle wurden ausschließlich bewohnernahe Grund- und Behandlungspflegeleistungen berücksichtigt
- nicht berücksichtigt wurden Zeitabschnitte, in denen folgende Leistungsarten die jeweils vorherrschende war:
 - Gespräch/ soziale Einzelbetreuung
 - Therapie- oder Rehaübungen
 - Visite/ Angehörigengespräch
 - Dokumentation
 - Pflegeplanung
 - Fallbesprechung
 - alle Gruppenangebote

Referenzmodell mit MARS

	Coefficients
(Intercept)	63.25013
ps_ohne_haertefall2	46.88307
ps_ohne_haertefall3	82.38187
A5	23.59461

Selected 4 of 5 terms, and 3 of 4 predictors
Number of terms at each degree of interaction: 1 3 (additive model)
GCV 2847.47 RSS 1207646 GRSq 0.3368391
RSq 0.3549245 cv.rsq 0.3277589

Referenzmodell
SGBXI_earth3

Varianzaufklärung

● Gesamtzeiten, nur Zustandsvariablen

- bestes Referenzmodell: 0.3051
- bestes Konkurrenzmodell: 0.4325 (CART)
- bestes Ensemble: 0.4707 (Random Forest)

● Gesamtzeiten, mit Leistungsvariablen

- bestes Konkurrenzmodell: 0.5748
- bestes Ensemble: 0.5635 (Random Forest)

● Nursing only, nur Zustandsvariablen

- bestes Referenzmodell: 0.3870 (MARS)
- bestes Konkurrenzmodell: 0.5803 (CART)
- bestes Ensemble: 0.6215 (Random Forest)

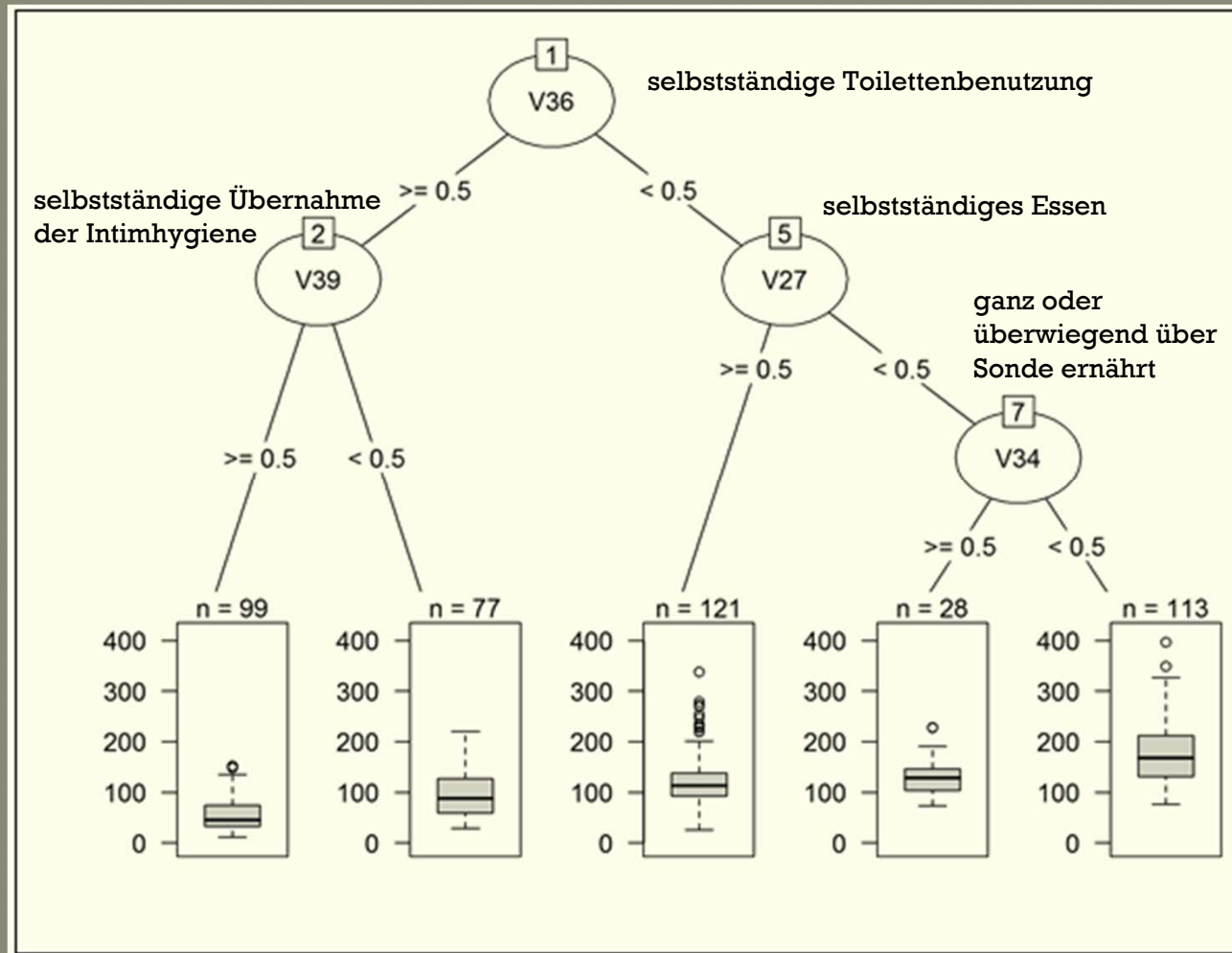
● Nursing only, mit Leistungsvariablen

- bestes Konkurrenzmodell: 0.6155 (cTree)
- bestes Ensemble: 0.6569 (Random Forest)

Konkurrenzmodelle ohne Leistungsvariablen

Modell	Pflegeaufwand	Gruppen/ Terme	R^2_{Lern}	R^2_{Kreuz}	R^2_{Test}
rpart_1	Gesamtzeiten	5	0.4388	0.3736	0.4327
ctree_1	Gesamtzeiten	7	0.4614	----	0.4248
earth_1	Gesamtzeiten	44	0.7624	0.0263	0.1028
earth_2	Gesamtzeiten	11	0.5245	0.3933	0.4043
earth_3	Gesamtzeiten	14	0.5436	0.4173	0.3487
earth_4	Gesamtzeiten	10	0.5133	0.4027	0.4029
rpart_earth_hybrid1	Gesamtzeiten	58	0.7757	0.0624	-0.3989
rpart_earth_hybrid2	Gesamtzeiten	11	0.5604	0.3412	0.4170
rpart_earth_hybrid3	Gesamtzeiten	14	0.5436	0.4113	0.3487
rpart_earth_hybrid4	Gesamtzeiten	9	0.5052	0.3804	0.3850
RandomForest_1	Gesamtzeiten	----	0.4541	----	0.4707
cforest_1	Gesamtzeiten	----	0.4773	----	0.4653
gbm_1	Gesamtzeiten	----	0.5513	----	0.4571
rpart_Nurs1	„Nursing only“- Zeiten	5	0.5608	0.5364	0.5803
ctree_Nurs1	„Nursing only“- Zeiten	6	0.4614	----	0.4248
earth_Nurs1	„Nursing only“- Zeiten	46	0.8071	0.0213	0.4197
earth_Nurs2	„Nursing only“- Zeiten	7	0.6006	0.5510	0.5277
earth_Nurs3	„Nursing only“- Zeiten	14	0.6403	0.5461	0.5380
earth_Nurs4	„Nursing only“- Zeiten	7	0.6006	0.5493	0.5277
rpart_earth_hybrid_Nurs 1	„Nursing only“- Zeiten	47	0.8114	-0.0132	0.3672
rpart_earth_hybrid_Nurs 2	„Nursing only“- Zeiten	7	0.6006	0.5481	0.5277
rpart_earth_hybrid_Nurs 3	„Nursing only“- Zeiten	14	0.6403	0.5473	0.5380
rpart_earth_hybrid_Nurs 4	„Nursing only“- Zeiten	7	0.6006	0.5468	0.5277
RandomForest_Nurs1	„Nursing only“- Zeiten	----	0.5447	----	0.6215
cforest_Nurs1	„Nursing only“- Zeiten	----	0.5750	----	0.5981
gbm_Nurs1	„Nursing only“- Zeiten	----	0.6346	----	0.5951

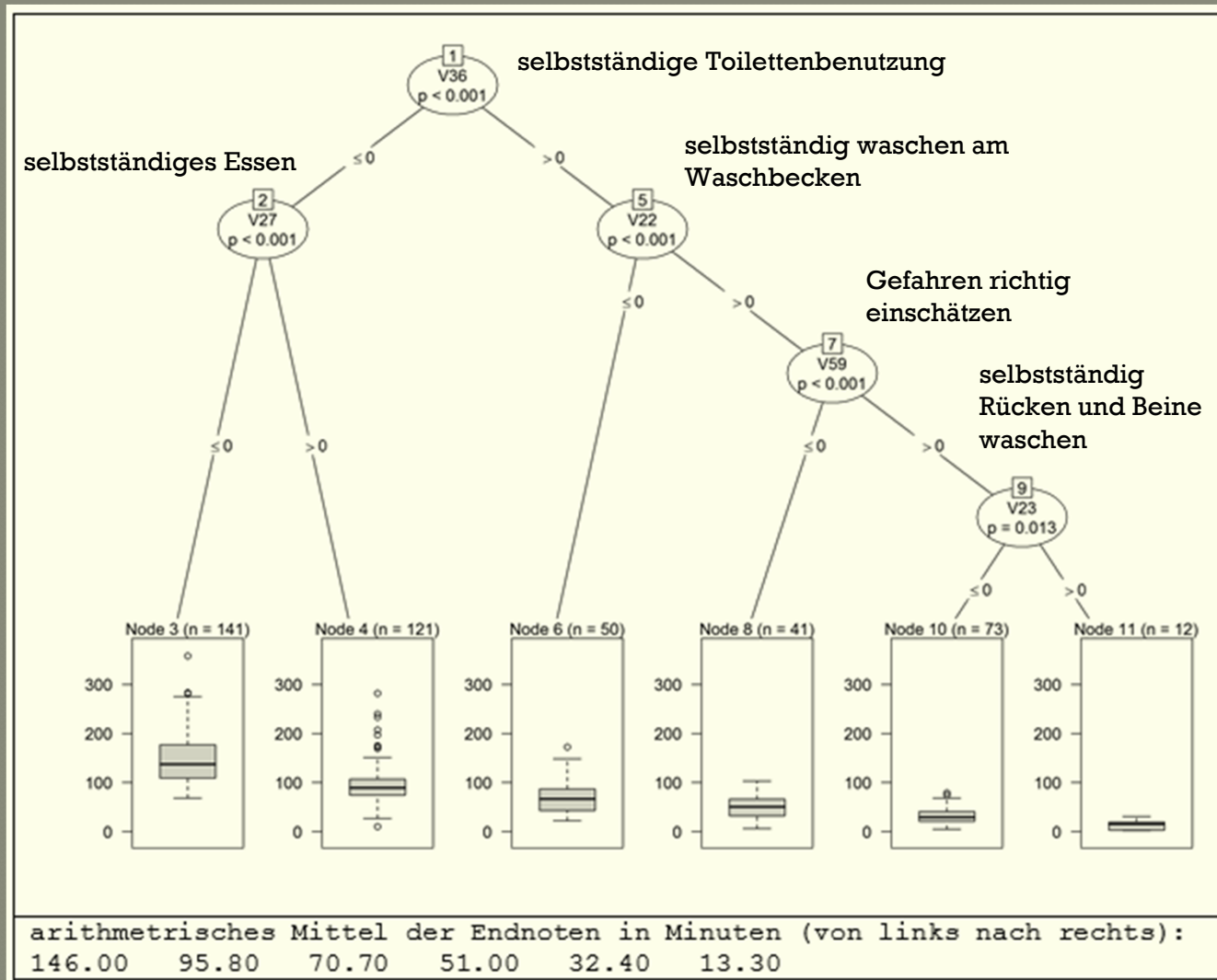
Beispiel: Modell rpart₁



$$R^2_{\text{Test}} = 0.4325$$

Bestes Referenzmodell:
0.3051

Modell ctree_Nurs₁ – Nursing only



$$R^2_{\text{Test}} = 0.6142$$

Bestes Referenzmodell:
0.3870

Konkurrenzmodelle mit Leistungsvariablen

Modell	Pflegeaufwand	Gruppen/ Terme	R^2_{Lern}	R^2_{Kreuz}	R^2_{Test}
rpart_LV1 (xerror)	Gesamtzeiten	6	0.4970	0.4033	0.4665
rpart_LV1 (xstd)	Gesamtzeiten	4	0.4447	0.3777	0.4299
cree_LV1	Gesamtzeiten	8	0.5196	----	0.4587
earth_LV1	Gesamtzeiten	46	0.8239	-0.0357	0.0636
earth_LV2	Gesamtzeiten	8	0.5788	0.5252	0.5748
earth_LV3	Gesamtzeiten	13	0.6001	0.5205	0.5045
earth_LV4	Gesamtzeiten	8	0.5663	0.5037	0.5512
rpart_earth_hybrid_LV1	Gesamtzeiten	44	0.8269	0.1698	0.0636
rpart_earth_hybrid_LV2	Gesamtzeiten	7	0.5810	0.5068	0.5748
rpart_earth_hybrid_LV3	Gesamtzeiten	14	0.6021	0.5070	0.5045
rpart_earth_hybrid_LV4	Gesamtzeiten	8	0.5635	0.5342	0.5512
RandomForest_LV1	Gesamtzeiten	----	0.5254	----	0.5635
cforest_LV1	Gesamtzeiten	----	0.5072	----	0.4908
gbm_LV1	Gesamtzeiten	----	0.6065	----	0.5532
rpart_NursLV1	„Nursing only“- Zeiten	5	0.5774	0.5292	0.6087
cree_NursLV1	„Nursing only“- Zeiten	7	0.5881	----	0.6155
earth_NursLV1	„Nursing only“- Zeiten	48	0.8689	-0.0092	0.1847
earth_NursLV2	„Nursing only“- Zeiten	7	0.6133	0.5517	0.5943
earth_NursLV3	„Nursing only“- Zeiten	14	0.6568	0.5781	0.5912
earth_Nurs4LV	„Nursing only“- Zeiten	6	0.6033	0.5699	0.5813
rpart_earth_hybrid_Nurs LV1	„Nursing only“- Zeiten	55	0.8878	-0.0469	0.1879
rpart_earth_hybrid_Nurs LV2	„Nursing only“- Zeiten	7	0.6133	0.5496	0.5943
rpart_earth_hybrid_Nurs LV3	„Nursing only“- Zeiten	14	0.6568	0.5777	0.5912
rpart_earth_hybrid_Nurs LV4	„Nursing only“- Zeiten	6	0.6033	0.5845	0.5813
RandomForest_NursLV1	„Nursing only“- Zeiten	----	0.5677	----	0.6569
cforest_NursLV1	„Nursing only“- Zeiten	----	0.5917	----	0.6043
gbm_NursLV1	„Nursing only“- Zeiten	----	0.6555	----	0.6188

Modell earth_LV₁

	coefficients
(Intercept)	133.17800
A14	32.72312
A15	37.97353
V27	-29.57421
V36	-34.14351
V52	20.62525
A14 * V46	40.23133
V22 * V59	-37.93445

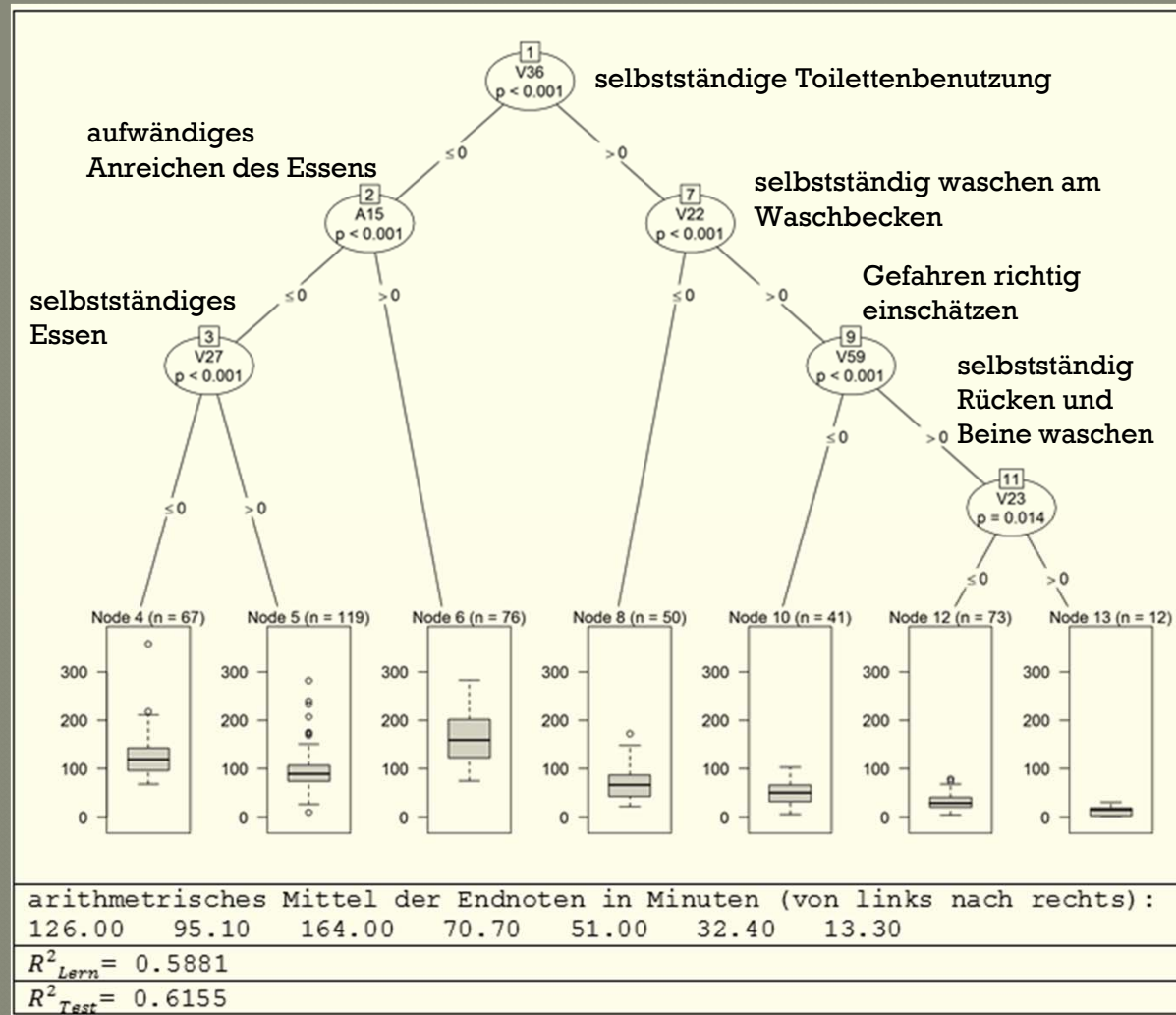
Number of cases: 438
Selected 8 of 9 terms, and 8 of 76 predictors
Number of terms at each degree of interaction: 1 5 2
GCV 1962.663 RSS 788561.5 GRSq 0.542906
RSq 0.5787824 cv.rsq 0.5252018
$R^2_{Test} = 0.5748$

$$R^2_{Test} = 0.5748$$

Bestes Referenzmodell:
0.3051

A14	nimmt mehr als 3h am Tag an einer Tagesgruppe teil
A15	aufwändiges Anreichen des Essens
V27	selbstständiges Essen
V36	selbstständige Toilettenbenutzung
V52	mehrmals pro Woche Teilnahme an Angeboten der sozialen Betreuung bzw. kulturellen Angeboten
V46	sozial inadäquates Verhalten
V22	selbstständig waschen am Waschbecken
V59	Gefahren richtig einschätzen

Modell ctree_NursLV₁ (Nursing only)



$$R^2_{Test} = 0.6155$$

Bestes Referenzmodell:
0.3870

Ensemblemethoden: Variable Importance

relative Bedeutsamkeit der erklärenden Variablen, standardisiert
(Gesamtzeiten, keine Leistungsvariablen)

RandomForest_1 (IncNodePurity)		cForest_1 (varImp)		gbm_1 (relative influence)	
selbstständige Toilettenbenutzung	100,00	selbstständig essen	100,00	selbstständig essen	100,00
selbstständig essen	0,91	selbstständige Toilettenbenutzung	0,72	selbstständige Toilettenbenutzung	0,89
selbstständig waschen am Waschbecken	0,80	selbstständig waschen am Waschbecken	0,59	selbstständig waschen am Waschbecken	0,63
Intimhygiene selbstständig	0,32	Intimhygiene selbstständig	0,49	Stuhlinkontinenz	0,27
Stuhlinkontinenz	0,29	Stuhlinkontinenz	0,46	Intimhygiene selbstständig	0,26
Sehbeeinträchtigung	0,16	umfängliche Lageveränderungen im Bett vornehmen	0,30	sozial inadäquates Verhalten	0,21
Gefahren richtig einschätzen	0,14	selbstständig Oberbekleidung an- und ausziehen	0,29	umfängliche Lageveränderungen im Bett vornehmen	0,16
selbstständige Transfers aus dem Bett	0,14	Stehen ohne Hilfe	0,25	kann Gefahren richtig einschätzen	0,12
ist gesellig	0,14	Gefahren richtig einschätzen	0,23	überwiegend Sondenernährung	0,12
umfängliche Lageveränderungen im Bett vornehmen	0,14	Wissen erwerben	0,23	Diäten verordnet	0,11

Diskussion

Diskussion

- Keine klare Tendenz zwischen CART, cTree und MARS
 - je nach Modellart liefert einer der 3 Algorithmen das Modell mit dem höchsten R^2_{Test}
- MARS-Modelle mit bis zu 3 Interaktionen und thresh=.001 zeigen massives Overfitting
 - GCV verhindert dies in diesem Fall nicht
- Hybrid-Modelle aus CART und MARS nicht besser als MARS
- Zeiten für Grund- und Behandlungspflege (Nursing only) lassen sich deutlich besser erklären als die Gesamtzeiten
 - Geeignete Prädiktoren für soziale Betreuung?