

„Entscheidungsbäume“ – eine Möglichkeit, Präferenzen und Zufriedenheit von Patienten im Gesundheitswesen zu verstehen?

Prof. Dr. Wolfgang Himmel
Abteilung Allgemeinmedizin
Universitätsmedizin Göttingen
Humboldtallee 38
37073 Göttingen
whimmel@gwdg.de

Ulrich Reincke
Competence Center Enterprise,
Intelligence, SAS Institute
In der Neckarhelle 162
69118 Heidelberg
Ulrich.Reincke@ger.sas.com

Zusammenfassung

Hintergrund. Befragungen zur Patientenzufriedenheit sind ein etabliertes Verfahren im Gesundheitswesen – mit dem Ziel der Kundenbindung, Qualitätssicherung und wissenschaftlichen Analyse von Präferenzen. Die üblichen Auswertungsstandards lassen zumeist nicht erkennen, welche Erfahrungen in der Arztpraxis oder im Krankenhaus letztlich die Entscheidung, dort wieder hinzugehen, beeinflussen. Dies begrenzt auch die Möglichkeit einer gezielten Intervention.

Fragestellung. Aus den Antworten von ca. 87.700 Patienten zu 26 Merkmalen der Zufriedenheit mit ihrem Krankenhausaufenthalt sollten die Entscheidung zur „Weiterempfehlung des Krankenhauses“ mittels der Methode „Entscheidungsbaum“ untersucht und die Ergebnisse mit multivariaten Standardanalysen (multiple logistische Regression) verglichen werden. Besondere Beachtung sollte die große Zahl „fehlender Werte“ finden.

Methode. Primäres Kriterium für die beiden Methoden „Entscheidungsbaum“ und multiple logistische Regression war die Weiterempfehlung eines Krankenhauses durch Patienten (als Indikator für „globale Patientenzufriedenheit“). Prädiktoren waren 26 Bewertungen verschiedener Eigenschaften des Krankenhauses und der Erfahrungen während des Aufenthaltes, z. B. Behandlungs- und Betreuungsqualität oder Zimmerausstattung. Den Entscheidungsbaum haben wir im Programm „Partition“ im Modul JMP von SAS modelliert.

Ergebnisse. Als wichtiger Faktor für die Patientenzufriedenheit erwies sich die „aufmerksame Behandlung im Krankenhaus“. Dieser Faktor war das oberste Blatt im Entscheidungsbaum und hatte auch in der multiplen Regression ein adjustiertes Odds Ratio (OR) von 2,4 (95%-Vertrauensintervall: 2,3 bis 2,6). Ebenfalls sehr wichtig im Baum und in der multiplen Regression waren: gute Betreuung (OR=3,8; 3,4 – 4,3) und Behandlungserfolg (2,9; 2,6 – 3,1). Für die weitere Analyse im Entscheidungsbaum wählten wir folgende Einstellungen: pro „Blatt“ mind. 4.000 „Fälle“ (Personen), 5-fache Kreuzvalidierung sowie Einschluss fehlender Werte. Der Datensatz wurde in 70 % Trainingsdaten und 30 % Validierungsdaten gesplittet. Die Lösung, die sich aus den von uns gewählten Kriterien ergab, führte zu 11 Teilungen. Die ROC-Kurven sowohl für die Trainings- als auch Validierungs-Stichprobe hatten dann eine Fläche unter der Kurve (*area under the curve*; AUC) von knapp über 80%. Insbesondere die Kombination von aufmerksamer Behandlung, Erklärung des Befundes, Zimmerausstattung und Behandlungserfolg führte zur Zufriedenheit bei fast allen Befragten (89 %). Umgekehrt zeigte der Entscheidungsbaum, dass Defizite in der „aufmerksamen Behandlung“ durch eine „gute Gesamtbetreuung“ und „Vermeidung von

Ärger“ soweit kompensiert werden konnte, dass immerhin noch 49 % der Patienten mit dem betreffenden Krankenhaus zufrieden waren.

Diskussion. Eindrucksvoll zeigte der „Entscheidungsbaum“ die Bedeutung einer aufmerksamen Behandlung, umsichtigen Betreuung und verständlichen Informationen seitens der Ärzte, Pfleger und Schwestern – typisch ärztliche und pflegerische Leistungen, die durchaus beeinflussbar sind bzw. deren Bedeutung den im Krankenhaus Beschäftigten vermittelbar sein sollte. Für aussagefähige Ergebnisse ist die Modellierung der Voreinstellungen im Entscheidungsbaum wichtig.

Schlüsselwörter: JMP, Entscheidungsbaum, Decision Tree, Patientenzufriedenheit, Krankenhaus

1 Einleitung

Die Zufriedenheit von Patienten mit ihrer medizinischen Versorgung hat sich in den letzten Jahren zu einem Marketing-Argument entwickelt; vor allem aber gilt die Patientenzufriedenheit als sensibler und verlässlicher Indikator für die Qualität der medizinischen Behandlung [1, 2]. Zufriedenheit ist ein multidimensionales Konstrukt, in das u.a. die Bewertung der technischen und emotionalen Aspekte medizinischen Handelns einfließen, aber auch externe Bedingungen, wie die bauliche Qualität, Erreichbarkeit oder Wartezeiten. Zufriedenheit ist gewissermaßen das Endprodukt der subjektiven Bewertung dieser verschiedenen Dimensionen unter Berücksichtigung ihrer unterschiedlichen Bedeutung für den einzelnen.

Für die Erhebung der Zufriedenheit stehen zahlreiche Instrumente, im Regelfall standardisierte Fragebogen zur Verfügung [3], welche die verschiedenen organisatorischen Gegebenheiten (z. B. in der Arztpraxis oder im Krankenhaus) berücksichtigen. Die Fragen, die dabei zum Einsatz kommen, sind im Regelfall strukturell nicht anders, als man dies von Hotel-Befragungen u. ä. kennt. Mittlerweile gelten mehrere dieser Erhebungsinstrumente als valide; die Auswertung stellt Forscher und Praktiker aber vor Probleme – gerade wenn Studien zur Patientenzufriedenheit handlungs- bzw. interventionsorientiert sein sollen.

Bei der Auswertung könnte man ganz einfach deskriptiv vorgehen und Antwort für Antwort einzeln oder als Summenscores (z. B. auf Basis einer Faktorenanalyse) auswerten. Damit können evtl. einzelne Schwachstellen einer medizinischen Einrichtung abgebildet werden. Aber aus einer solchen Auswertung lässt sich nicht oder nur unzureichend erkennen, welche Erfahrungen in der Arztpraxis oder im Krankenhaus letztlich die Entscheidung, dort wieder hinzugehen, beeinflussen. Dies begrenzt auch die Möglichkeit einer gezielten Intervention.

Scheinbar bieten Verfahren wie die multiple logistische Regression die Möglichkeit, ein Ereignis auf Basis mehrerer Variablen zu erklären. Im Fall der Patientenzufriedenheit

könnte man sich vorstellen, Patienten zu fragen, ob sie dieselbe Praxis oder dasselbe Krankenhaus Verwandten und Bekannten empfehlen würden, um dann zu prüfen, welche Fragen zur Zufriedenheit am stärksten mit dieser Empfehlung assoziiert sind. Dieses Verfahren stößt aber schnell an Grenzen, wenn die Zahl der Variablen – also die verschiedenen Fragen zur Zufriedenheit – umfangreich sind und man auch noch (ggf. mehrfache) Interaktionen zwischen den Variablen berücksichtigen möchte.

Als Alternative bietet sich der sog. Entscheidungsbaum an – ein Verfahren, mit dem der Einfluss einer beliebig großen Zahl von Variablen, einschließlich ihrer Interaktion auf eine Ziel-Variable untersucht werden kann. Um allerdings eine Entscheidungsbaum-Analyse durchzuführen und die Ergebnisse möglichst intuitiv zu verstehen, sind leistungsfähige Software-Module notwendig. Mit dem Modul JMP steht in SAS ein solches Verfahren zur Verfügung, das wir für die Auswertung eines umfangreichen Datensatzes zur Patientenzufriedenheit im Krankenhaus verwendet und im praktischen Handling überprüft haben.

2 Methode

Antworten von Patienten zur Frage der Zufriedenheit im Krankenhaus wurden explorativ mit der Methode „Entscheidungsbaum“, unter Verwendung des Moduls JMP ausgewertet.

2.1 Stichprobe und Fragebogen

Es lagen Antworten von ca. 87.700 Patienten aus verschiedenen Krankenhäusern und von verschiedenen Stationen vor. Das Alter der Patienten war erwartungsgemäß höher als im Durchschnitt der Bevölkerung (Mittelwert: 60,6; Standardabweichung: 16,3); entsprechend lag auch der Anteil der Frauen etwas höher (54 %).

Alle Befragten hatten den Zufriedenheitsbogen des Picker-Instituts ausgefüllt [4]. Dieser Fragebogen hat den Vorteil, nicht nur global nach Zufriedenheit mit den verschiedenen Personen oder Bereichen eines Krankenhauses zu fragen, sondern kritische Verhaltensweisen und Ereignisse zu erheben. Zum Beispiel wird danach gefragt, ob man verständliche Antworten erhielt, wenn man Ärzte oder Pflegepersonen fragte – mit der Antwortmöglichkeit: „Ja, immer“; „Ja, manchmal“; „Nein“ und „Ich hatte keine Fragen“. Der Fragebogen ist mit insgesamt 98 Fragen bzw. Items sehr detailliert.

2.2 Auswertung

Die Patienten wurden unter anderem danach gefragt, ob sie das betreffende Krankenhaus ihrer Familie oder ihren Freunden weiterempfehlen würden. Die Antworten auf diese Frage haben wir als globales Maß der Zufriedenheit gewertet und die Stichprobe dichotomisiert (Antwort: „Ja, ganz sicher“ = „zufrieden“ vs. Antworten: „Ja, wahrscheinlich“ oder „Nein“ = „nicht zufrieden“). Mit der Methode „Entscheidungsbaum“

sollte statistisch nachvollzogen werden, wie die Entscheidung zur Weiterempfehlung eines Krankenhauses bzw. Nicht-Empfehlung zustande kommt.

Eine gute Beschreibung der Methode findet sich z.B. bei [5]. An dieser Stelle seien nur einige Merkmale bei der Konstruktion von Entscheidungsbäumen beschrieben. Der Gesamtdatensatz (bzw. ein Trainingsdatensatz) mit seiner Aufteilung nach dem Kriterium – in unserem Fall also die Unterscheidung nach Empfehlung bzw. Nicht-Empfehlung – wird als Wurzel bezeichnet. Es werden nun schrittweise diejenigen Merkmale gesucht, mit dem sich der Datensatz gemäß dem Kriterium bestmöglich teilen lässt. Bspw. könnte dies zunächst das Merkmal „Qualität des Essens“ sein. Die Teilung des Datensatzes in Personen, die mit dem Essen zufrieden sind, und Personen, die damit unzufrieden sind, würde also – im Vergleich zu allen denkbaren anderen Teilungen – die beste Vorhersagemöglichkeit ergeben, wer ein Krankenhaus weiterempfehlen dürfte. Man hat also die vergleichsweise beste Trefferqualität bei der Annahme, dass die zufriedenen Esser das Krankenhaus wohl weiterempfehlen und umgekehrt. Die Prognose selbst muss nicht besonders gut bzw. sicher sein (vielleicht würden 40% der zufriedenen Esser das Krankenhaus gar nicht weiterempfehlen); nur ist diese erste Prognose auf Basis einer ersten Teilung immer noch besser als jede andere mögliche Aufteilung des Datensatzes.

In einem weiteren Teilungsschritt würde man nun versuchen, entweder die Gruppe der zufriedenen oder unzufriedenen Esser nach einem weiteren Merkmal aufzuteilen, um für diese Untergruppen die prognostische Qualität hinsichtlich des Kriteriums zu verbessern. Um noch einmal beim Beispiel zu bleiben: Man könnte die Gruppe der unzufriedenen Esser danach aufteilen, ob sie mit der ärztlichen Information zufrieden oder unzufrieden waren. Die beiden Merkmale Unzufriedenheit mit dem Essen und der ärztlichen Information werden vermutlich die Wahrscheinlichkeit erhöhen, das Krankenhaus nicht weiterzuempfehlen – und damit auch unsere prognostische Fähigkeit, dieses Kriterium vorauszusagen.

Schon an dieser Stelle deuten sich drei wesentliche Eigenschaften der Methode „Entscheidungsbaum“ an: (1) Es bedarf eines Kriteriums für die Teilung, d.h. an welcher Stelle des Entscheidungsbaums weitergeteilt und welches Kriterium dafür gewählt wird; (2) es bedarf genauso eines Abbruchkriteriums, wann nicht weitergeteilt werden soll, weil natürlich jede weitere Teilung immer noch – und sei es minimal – die Prognose verbessert; (3) das Verfahren wird relativ schnell unübersichtlich. Die Vorteile des Entscheidungsbaums, aber auch die zuletzt genannten Probleme möchten wir an unserem Beispiel-Datensatz erläutern.

3 Ergebnisse

Unter der Annahme, dass eine vorbehaltlose Weiterempfehlung des Krankenhauses eine globale Zufriedenheit ausdrückt, waren 67,3 % der Befragten mit ihrem Krankenhausaufenthalt zufrieden. Um einen ersten Eindruck zu erhalten, welche Erfahrungen im

Krankenhaus mit der globalen Zufriedenheit assoziiert sind, zeigt Tabelle 1 die 15 wichtigsten Variablen, einschließlich ihrer Odds Ratios als Effektmaß. Ebenfalls zur Information finden sich dort auch die Ergebnisse einer multiplen logistischen Regression mit Einschluss dieser zehn wichtigsten Variablen. Den stärksten Effekt haben die eingeschätzte Betreuung im Krankenhaus, das Vertrauen in die Ärzte, Behandlungserfolg sowie die Erfahrung, von Ärzten aufmerksam behandelt worden zu sein. Bei dieser Auswertungsvariante gab es bereits fast 23.000 „fehlende Werte“. Hätte man übrigens alle 31 Bewertungen in die logistische Regression eingeschlossen, wäre die Zahl fehlender Werte auf fast 42.000 gestiegen.

Im Vergleich dazu nun die Vorgehensweise und Ergebnisse mit der Methode „Entscheidungsbaum“. Das Modul JMP ist in SAS eingebunden, so dass der für die multiple Regression verwendete SAS-Datensatz unmittelbar auch für JMP verwendbar ist. Etwas verwirrend ist am Anfang, dass man im Menü-gesteuerten Programm eine Prozedur „Entscheidungsbaum“ bzw. „decision tree“ nicht findet. Das vergleichbare statistische Verfahren firmiert bei JMP als „partition“ unter dem Haupt-Menü „Modellierung“.

Tabelle 1: Faktoren für eine Weiterempfehlung eines Krankenhauses

Variable *	Odds Ratio	Adjustiert	
		Odds Ratio	(95%-Intervall)
Gute Betreuung	32,0	3,8	3,4 - 4,3
Behandlungserfolg	9,0	2,8	2,6 - 3,1
Aufmerksame Behandlung	8,0	2,4	2,3 - 2,6
Vertrauen in Ärzte	9,8	2,4	2,3 - 2,6
Von anderen empfohlen	2,6	2,0	1,9 - 2,3
Empfohlen durch Familie	2,3	2,0	1,9 - 2,2
Gute Zimmerausstattung	3,9	1,8	1,7 - 1,9
Gutes Essen	3,3	1,7	1,6 - 1,8
Kenne Krankenhaus	1,5	1,7	1,6 - 1,8
Krankenhaus nicht in der Nähe	1,5	1,7	1,6 - 1,8
Keinen Ärger	4,8	1,7	1,6 - 1,8
Ergebnisse erklärt	4,6	1,6	1,5 - 1,7
Gute Nasszellen	2,9	1,6	1,5 - 1,7
Symptome erklärt	3,8	1,5	1,4 - 1,6
Antworten der Ärzte verständlich	6,4	1,5	1,4 - 1,6

* sortiert nach adjustierten Odds Ratios; missing values: n = 22.691

Nach Aufrufen des entsprechenden Menü-Punktes müssen das Kriterium (bei uns also Krankenhaus-Empfehlung) und die Liste möglicher entscheidungsrelevanter Variablen eingegeben werden. Wie für diese Methode nicht unüblich, haben wir alle zur Verfügung stehenden Variablen in das Model mitaufgenommen, vor allem natürlich die einzelnen Zufriedenheitsaspekte, aber auch Geschlecht und Alter, Größe des Krankenhauses etc. Alle Variablen lagen kategorial vor bzw. wurden (z.B. im Falle des Alters) ka-

tegorial gruppiert, viele Variablen mit Mehrfach-Ausprägung dichotomisiert. Das wäre übrigens nicht erforderlich gewesen, da alle Varianten von Variablen im Entscheidungsbaum möglich sind, eine Dichotomisierung bzw. eine Begrenzung der Variablen auf wenige Ausprägungen Variablen aber am einfachsten „lesbar“ ist.

Die erste Teilung des Datensatzes erfolgt – etwas abweichend von den Ergebnissen der logistischen Regression – nicht anhand der Variablen „Gute Betreuung“, sondern anhand „Aufmerksame Behandlung durch Ärzte“ (Abbildung 1). Die Information, die man durch diese Teilung des Datensatzes bzw. an dieser Stelle des Entscheidungsbaums erhält, könnte man folgendermaßen lesen: Wenn man lediglich weiß, dass ein Patient mit der ärztlichen Aufmerksamkeit im Krankenhaus zufrieden war, wird man mit 78 % Wahrscheinlichkeit voraussagen können, dass diese Person das Krankenhaus weiterempfiehlt. In der nächsten Stufe des Entscheidungsbaums bzw. durch die nächste Teilung des Datensatzes wird man versuchen, dasjenige Merkmal zu finden, das zum größten Informationsgewinn in der Prognose beiträgt. Im konkreten Fall würde eine weitere Teilung optimal am Knoten „Zufriedenheit mit aufmerksamen Behandlung“ ansetzen – und nicht (was auch denkbar gewesen wäre) bei der Unzufriedenheit mit der Behandlung. Und das optimalste Trennungmerkmal ist „Vertrauen in die Ärzte“ (Abbildung 2). Mit anderen Worten: Wenn man weiß, dass eine Person sich aufmerksam behandelt fühlte und Vertrauen hatte, kann man nun schon mit einer Sicherheit von 80 % davon ausgehen, dass diese Person das Krankenhaus weiterempfiehlt – oder anders formuliert: Von allen Personen, für die diese Kombination zutrifft, werden 80 % eine Weiterempfehlung aussprechen. Damit deutet sich dann auch die unterschiedliche Betrachtungsweise im Vergleich zur logistischen Regression an.

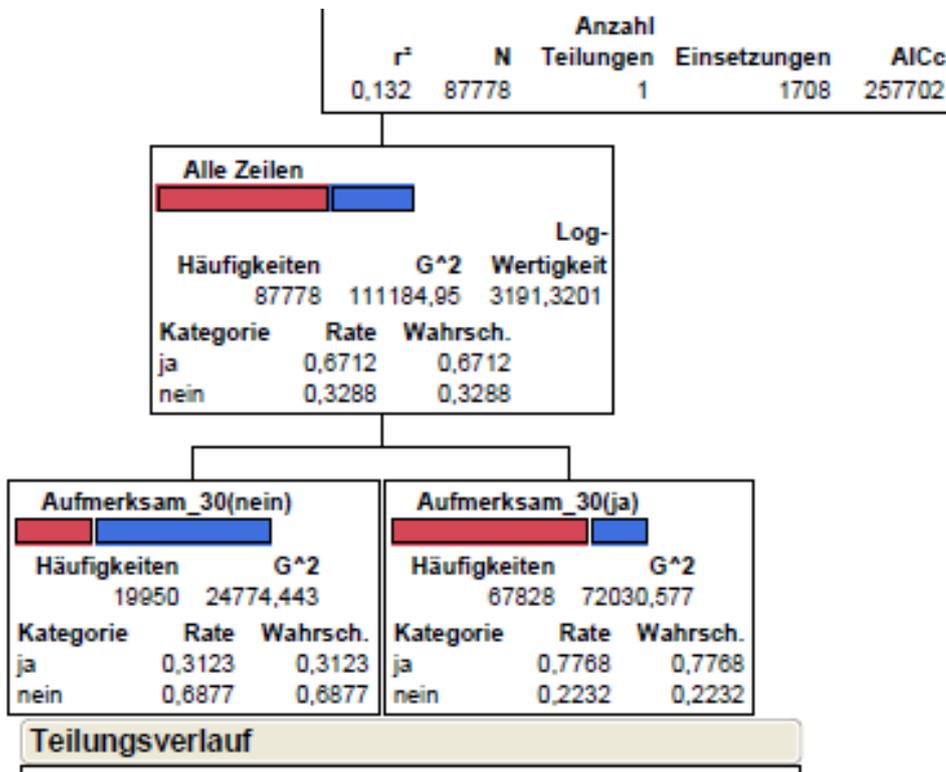


Abbildung 1: Entscheidungsbaum – 1. Teilung

Bisher haben wir den Entscheidungsbaum rein intuitiv und ohne definierte Kriterien wachsen lassen. Im Folgenden soll es darum gehen, wie weit man die Kriterien für die Teilung steuern kann – und soll das gleich im Programm selbst demonstriert werden. Das inhärente Problem der Methode „Entscheidungsbaum“ ist, dass fast zwangsläufig jede weitere Teilung zu einer Verbesserung der Prognose führt – zumindest für die Trainingsdaten, diese Verbesserung sich aber nicht unbedingt für die Validierungsdaten verallgemeinern lässt. Ab einer bestimmten Progression der Teilungen tritt das sogenannte „overfitting“ auf den Trainingsdaten ein [6]. Abbildung 3 veranschaulicht dieses Phänomen. Schon deshalb sollte man die Stichprobe in Trainings- und Validierungsdaten teilen, um sicher zu stellen, dass weitere Teilungen im Entscheidungsbaum einen messbaren Informationsgewinn auch für die Validierungs-Stichprobe bedeuten.

Ein weiteres mögliches Kriterium ist eine Mindestzahl an Beobachtungen für jedes Blatt des Entscheidungsbaums. Aufgrund der großen Stichprobe konnten wir dieses Kriterium großzügig mit $n=4000$ definieren. Damit ist garantiert, dass hinter jeder Entscheidungsregel mit 5% der Gesamtheit auch eine praktisch relevante Zahl von Probanden bzw. Patienten steht.

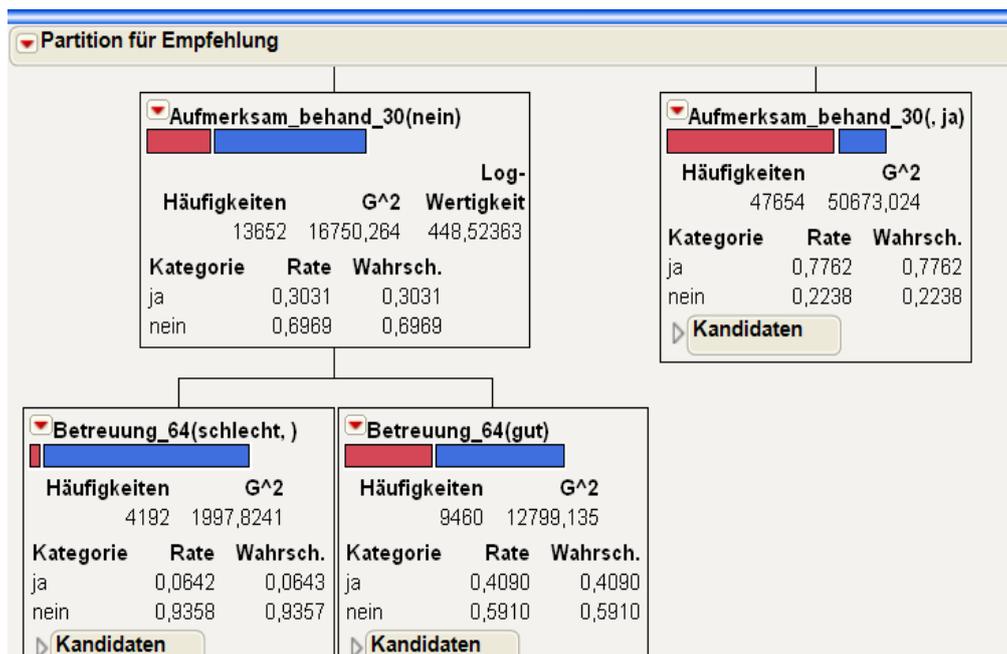


Abbildung 2: Entscheidungsbaum – 2. Teilung

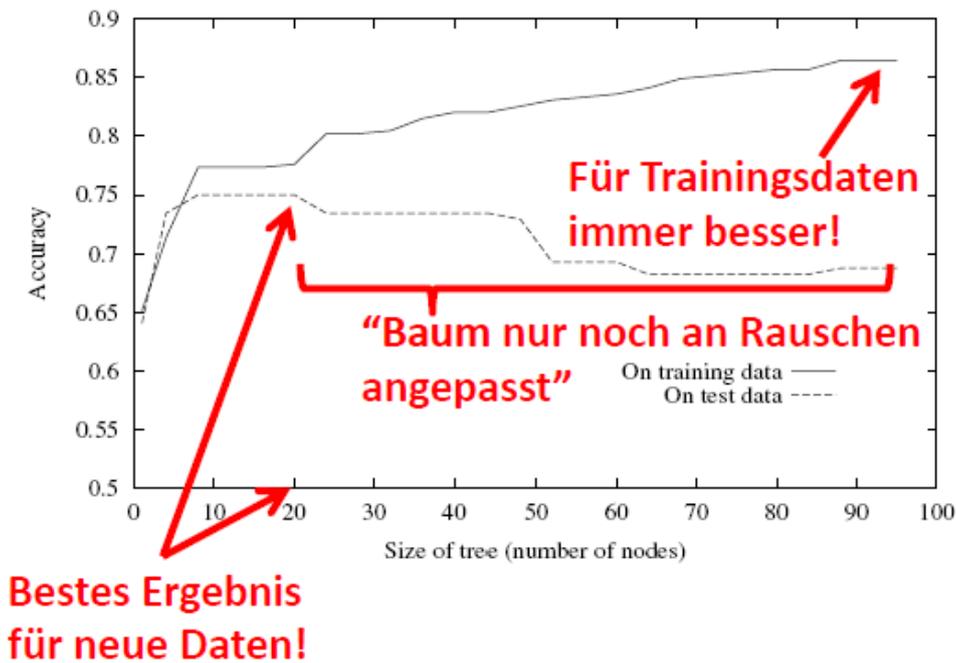


Abbildung 3: Entscheidungsbaum – Overfitting der Trainingsdaten nach [6]

Ein weiteres, gerade für unsere Stichprobe sinnvolles Kriterium im „Entscheidungsbaum“ ist die Art der Berücksichtigung fehlender Werte. Sie können *per Zufall* einer der beiden Teilungsvarianten zugeteilt werden (als „Einsetzungen“ in Abbildung 1 gekennzeichnet) – sicher unproblematisch bei nur wenigen fehlenden Werten. Bei einer Vielzahl fehlender Werte ist dagegen das Kriterium sinnvoll, die fehlenden Werte *sichtbar* einer der beiden Teilungsvarianten zuzuweisen. Ein „Blank“ bei der betreffenden Ausprägung einer Kategorie symbolisiert „fehlende Werte“;. In Abbildung 4 wären also die „fehlenden Werte“ bei der Kategorie „Betreuung_64 (schlecht;)“ und „Antwort_Verstand_9“ (nein, , keine Fragen). Gerade diese Maßnahme trägt im Gegensatz zu logistischen Regressionen zur Transparenz der Methode bei.

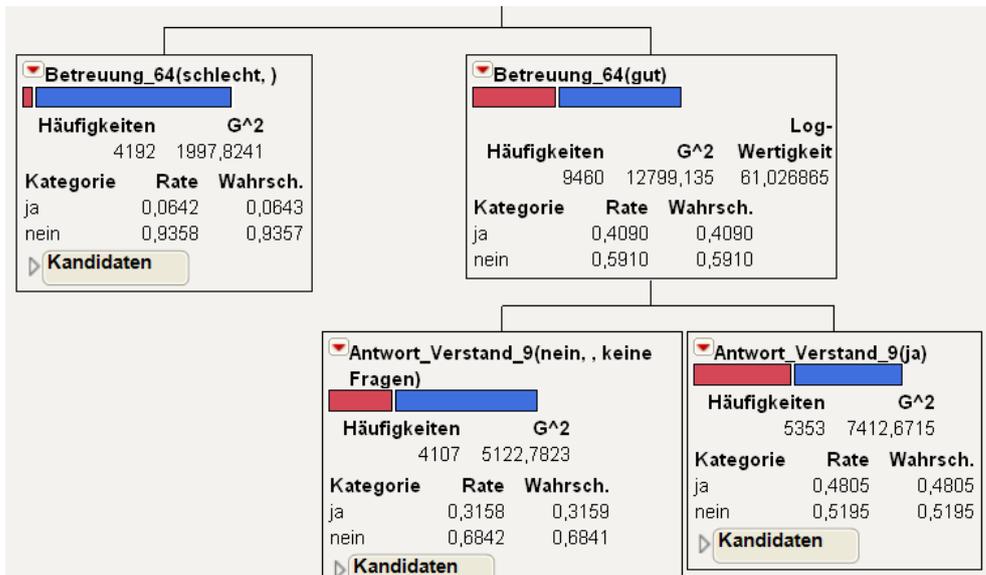
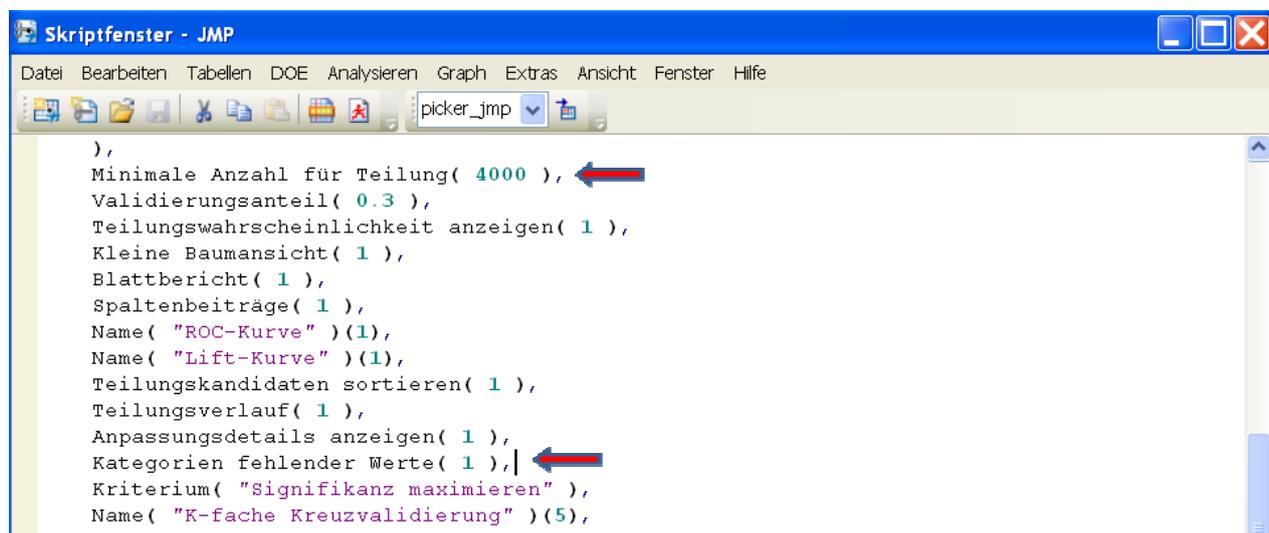


Abbildung 4: Fehlende Werte

Auch im Sinne gezielter Qualitätssicherung ist eine weitere Option in JMP sinnvoll: nämlich die „Sperrung“ von Variablen für die weitere Analyse. Bewusst haben wir in unserem ersten Anlauf der Datenanalyse mehr oder minder alle Variablen, die uns zur Verfügung standen, in das Modell eingeschlossen (einschl. Alter und Geschlecht der Patienten oder auch die Größe des jeweiligen Krankenhauses). Das kann im Einzelfall natürlich interessant sein (übrigens spielte bei uns nur das Alter der Patienten eine Rolle). Unter dem Aspekt der Qualitätssicherung aber würde man möglicherweise der Größe des Krankenhauses oder auch dem Alter der Befragten nicht unbedingt Bedeutung beimessen, weil man diese Faktoren nicht ändern kann. Es könnte sich dann als sinnvoll herausstellen, diese Faktoren aus dem Modell auszuschließen und erneut das Modell zu rechnen (beides ist in JMP leicht umsetzbar).

So intuitiv und leicht zugänglich die Menüsteuerung bei JMP ist, so energierend kann das immer gleiche Anklicken bestimmter Optionen werden, wenn man kleine Varianten durchrechnen möchte. Hier bietet ein Skriptfenster (Abbildung 5) den Vorteil den Programm-Code darzustellen, in dem dann selbst an den entscheidenden Stellen eine Variante eingefügt werden kann. In der erwähnten Abbildung zeigen die Pfeile beispielhaft wo man die minimale Anzahl für die Teilung bzw. die Berücksichtigung fehlender Werte ändern kann.



```

Skriptfenster - JMP
Datei Bearbeiten Tabellen DOE Analysieren Graph Extras Ansicht Fenster Hilfe
picker_jmp
),
Minimale Anzahl für Teilung( 4000 ),
Validierungsanteil( 0.3 ),
Teilungswahrscheinlichkeit anzeigen( 1 ),
Kleine Baumansicht( 1 ),
Blattbericht( 1 ),
Spaltenbeiträge( 1 ),
Name( "ROC-Kurve" )(1),
Name( "Lift-Kurve" )(1),
Teilungskandidaten sortieren( 1 ),
Teilungsverlauf( 1 ),
Anpassungsdetails anzeigen( 1 ),
Kategorien fehlender Werte( 1 ),
Kriterium( "Signifikanz maximieren" ),
Name( "K-fache Kreuzvalidierung" )(5),

```

Abbildung 5: Skriptfenster



Abbildung 6: Entscheidungsregeln

Für die Überprüfung der Güte des Entscheidungsbaums haben wir die Daten in 70 % Trainingsdaten und 30 % Validierungsdaten aufgeteilt. Die Lösung, die sich aus den von uns gewählten Kriterien ergab, führte zu 11 Teilungen. Abbildung 6 zeigt die Entscheidungsregeln, die sich als am prognostisch effizientesten bei dieser Anzahl von Teilungen herausgestellt haben. Abbildung 7 zeigt dann die ROC-Kurven sowohl für die Trainings- als auch Validierungs-Stichprobe mit einer recht akzeptablen Fläche unter der ROC-Kurve (*area under the curve*; AUC) von knapp über 80 %. Das ist auch nach mehrfach gezielten Durchrechnen mit verschiedensten Teilungsvarianten der derzeit beste Kompromiss.

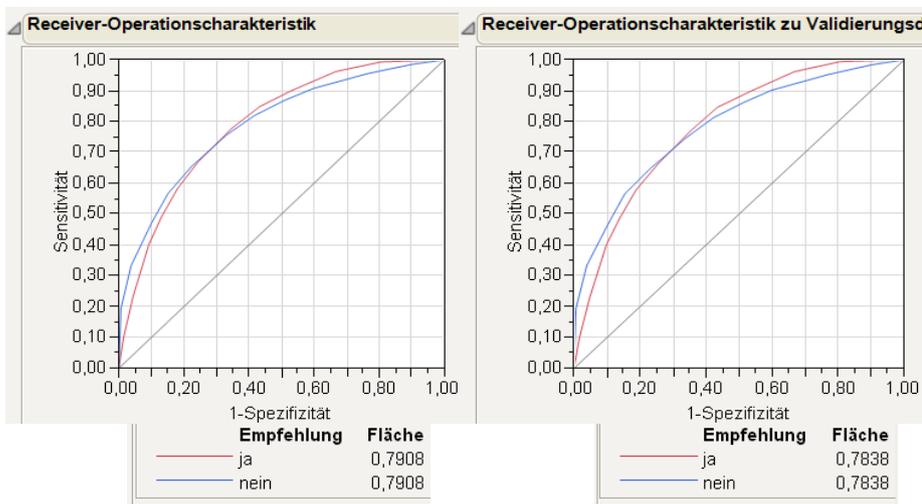


Abbildung 7: ROC-Kurven für Trainings- und Validierungsdaten

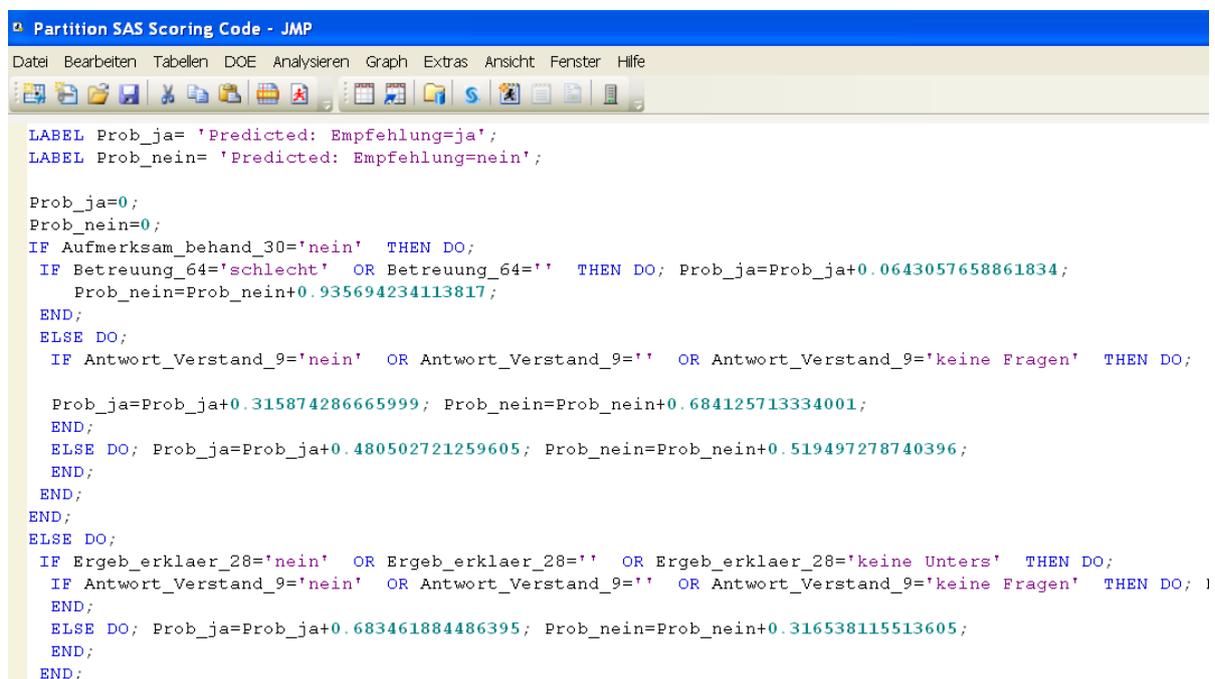
Die gefundene Entscheidungsbaum-Variante kann man übrigens als SAS-Code mit den entsprechend verschachtelnden *do-if-end-loops* als SAS-Code schreiben lassen (ausschnittsweise in Abbildung 8). Auch das kann man als klaren Transparenz-Vorteil sehen, insoweit tatsächlich die Entscheidungsregeln für jede Person des Datensatzes eine errechnete Wahrscheinlichkeit für die Weiterempfehlung des Krankenhauses ergeben, die dann mit der tatsächlichen Entscheidung verglichen werden kann (Abbildung 9). So

zeigt sich bei Patient 11986 eine falsche Zuordnung, die man bei Bedarf dann genauer überprüfen könnte.

4 Diskussion

Die Erfahrungen im Krankenhaus und die Zufriedenheit mit dem dortigen Aufenthalt standen zwar nicht im Vordergrund dieses Beitrages, aber - noch weit deutlicher als in der logistischen Regression - zeigte sich im Entscheidungsbaum, dass für die Patienten aufmerksame Behandlung, umsichtige Betreuung und verständliche Informationen seitens der Ärzte, Pfleger und Schwestern am wichtigsten sind. Natürlich wissen wir, dass auch scheinbar äußerliche Faktoren wie die Ausstattung des Zimmers und die Qualität des Essens wichtig sind, aber sie stehen hier erst an dritter und vierter Stelle.

Das bedeutet aber auch, dass der Entscheidungsbaum uns deutlich macht, dass die Kriterien für eine letztendliche Bewertung des Krankenhauses zu komplex sind, als dass eine einfache Manipulation wie z. B. Zimmerkosmetik oder etwas besseres Essen schon nachhaltig die Beurteilung des Krankenhauses ändern würde. Umgekehrt wird deutlich, dass es typisch ärztliche und pflegerische Leistungen sind, die über die Bewertung von Patienten entscheiden und in der Tat auch durch die Krankenhausleitung zu ändern sind bzw. deren Bedeutung den im Krankenhaus Beschäftigten vermittelbar sein sollte.



```

LABEL Prob_ja= 'Predicted: Empfehlung=ja';
LABEL Prob_nein= 'Predicted: Empfehlung=nein';

Prob_ja=0;
Prob_nein=0;
IF Aufmerksam_behand_30='nein' THEN DO;
  IF Betreuung_64='schlecht' OR Betreuung_64='' THEN DO; Prob_ja=Prob_ja+0.0643057658861834;
    Prob_nein=Prob_nein+0.935694234113817;
  END;
ELSE DO;
  IF Antwort_Verstand_9='nein' OR Antwort_Verstand_9='' OR Antwort_Verstand_9='keine Fragen' THEN DO;

  Prob_ja=Prob_ja+0.315874286665999; Prob_nein=Prob_nein+0.684125713334001;
  END;
ELSE DO; Prob_ja=Prob_ja+0.480502721259605; Prob_nein=Prob_nein+0.519497278740396;
  END;
END;
ELSE DO;
  IF Ergeb_erklaer_28='nein' OR Ergeb_erklaer_28='' OR Ergeb_erklaer_28='keine Unters' THEN DO;
  IF Antwort_Verstand_9='nein' OR Antwort_Verstand_9='' OR Antwort_Verstand_9='keine Fragen' THEN DO;
  END;
ELSE DO; Prob_ja=Prob_ja+0.683461884486395; Prob_nein=Prob_nein+0.316538115513605;
  END;
END;

```

Abbildung 8: SAS-Scoring-Code

	Geschlecht	Versicher	Betten	jungalt	Empfehlung	Wahrsch. (Empfehlung==ja)	Wahrsch. (Empfehlung==nei)	Blattnummer (
11984	Mann	Kasse	<999	< 65	ja	0,8785624996	0,1234375004	10
11985	Frau	Kasse	<999	> 64	ja	0,7928264323	0,2071735677	8
11986	Mann	Kasse	<999	< 65	nein	0,734779346	0,265220654	7
11987	Mann	Kasse	<999	< 65	nein	0,0643057659	0,9356942341	1
11988	Frau	Kasse	<999	< 65	ja	0,626339329	0,373660671	6
11989	Frau	Kasse	<999	> 64	ja	0,3158742867	0,6841257133	2
11990	Frau	Kasse	<999	> 64	ja	0,734779346	0,265220654	7
11991	Frau	Kasse	<999	< 65	ja	0,9328694052	0,0671305948	12
11992	Frau	Kasse	<999	< 65	ja	0,7928264323	0,2071735677	8
11993	Mann	Privat	<999	< 65	ja	0,626339329	0,373660671	6
11994	Mann	Kasse	<999	< 65	ja	0,6834618845	0,3165381155	5
11995	Frau	Kasse	<999	> 64	ja	0,9328694052	0,0671305948	12
11996	Frau	Kasse	<999	> 64	nein	0,0643057659	0,9356942341	1
11997	Frau	Kasse	<999	< 65	ja	0,8785624996	0,1234375004	10
11998	Mann	Kasse	<999	< 65	ja	0,7928264323	0,2071735677	8

Abbildung 9: SAS-Scoring; verglichen mit tatsächlicher Entscheidung („Empfehlung“)

Eine deutliche Einschränkung zur Validität der Ergebnisse bzw. zur Validität der Methode „Entscheidungsbaum“ muss gemacht werden. Diese Methode ist eine rein fiktive Nachbildung eines Entscheidungsprozesses und besagt nichts über den tatsächlichen Ablauf einer Entscheidung – und schon gar nichts über den Einzelfall. Wir sehen im Entscheidungsbaum, nicht anders als in der multiplen Regression statistische Zusammenhänge, deutlicher noch gesagt: assoziative Zusammenhänge, die uns mit gewisser Wahrscheinlichkeit über die Bedeutung einzelner Faktoren informieren. Wenn wir uns allerdings eine Person auf einer der untersten Blätter des Entscheidungsbaums vorstellen (also mit einer komplexen Entscheidungsregeln), wäre sehr wohl auch denkbar, dass nicht diese multiplen Faktoren ihre Entscheidung zur Weiterempfehlung prägte, sondern ganz einfach das Glück über eine gelungene Operation alles andere dominierte.

Eine weitere Einschränkung bezieht sich auf die Globalanalyse, die zwar recht stabil und zuverlässig die „Weiterempfehlung des Krankenhauses“ voraussagt (schon aufgrund der großen Zahl von Fällen), aber nicht unbedingt für einzelne Krankenhäuser gelten muss. Erst krankenhausbezogene Analysen liefern spezifische Ansatzpunkte für Interventionen, deren Erfolg aber getrennt evaluiert werden müsste, da der „Entscheidungsbaum“ nicht mehr als ein – allerdings mächtiges – heuristisches Verfahren ist.

In jedem Fall empfanden wir es als positiv, dass die komfortabel-intuitive Bedienung des Programms einerseits, die graphische Darstellung der Ergebnisse andererseits einen umfangreichen Datensatz mit umfangreichen Informationen heuristisch aufschließt und für Maßnahmen der Qualitätssicherung erste Ansatzpunkte liefert.

Literatur

- [1] Glickman SW, Boulding W, Manary M, Staelin R, Roe MT, Wolosin RJ, Ohman EM, Peterson ED, Schulman KA.: Patient satisfaction and its relationship with clinical quality and inpatient mortality in acute myocardial infarction. *Circ Cardiovasc Qual Outcomes* 2010, 3 S. 188-195.
- [2] Rao M, Clarke A, Sanderson C, Hammersley R.: Patients' own assessments of quality of primary care compared with objective records based measures of technical quality of care: cross sectional study. *Br Med J* 2006, 333 S. 19.
[<http://www.bmj.com/content/333/7557/19.full.pdf?sid=22f40392-8471-4918-9bd2-102af3f623a3>]
- [3] Castle NG, Brown J, Hepner KA, Hays RD.: Review of the literature on survey instruments used to collect data on hospital patients' perceptions of care. *Health Serv Res* 2005, 40 S. 1996-2017.
- [4] Jenkinson C, Coulter A, Bruster S.: The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health Care* 2002, 14 S. 353-358.
- [5] Gaudard M, Ramsey P, Stephens M. Interactive Data Mining and Design of Experiments: the JMP Partition and Custom Design Platforms. North Haven Group. SAS (White Paper): March 2006
[http://www.jmp.com/software/whitepapers/pdfs/372455_interactive_datamining.pdf]
- [6] Mitchell TM: Machine Learning, Boston, Mass. McGraw-Hill, 1997.