

Wettbewerbsfähige Ratingmodelle trotz unvollständiger Daten

Die Verwendung von Rating- und Scoring-Modellen hat sich in weiten Teilen der Kreditwirtschaft durchgesetzt. Ratings unterstützen Kreditentscheidung und Kreditfolgebearbeitung. Darüber hinaus erlangen Bonitätsratings mit der Umsetzung der Solvabilitätsverordnung (SolV) und der MaRisk eine zentrale Rolle im Kreditrisikomanagement, im Risikoreporting, in der Portfoliosteuerung und der regulatorischen Eigenkapitalberechnung.

Die gestiegene Bedeutung von Bonitätseinstufungen und der Quantifizierung des absoluten Risikoniveaus in Form der PD-Schätzung geht einher mit erhöhten Anforderungen an die Qualität von Rating-Systemen. Diese Anforderungen spiegeln sich in den Mindestanforderungen an die Nutzung des IRBA in der SolV wider.

Scorekarten als Industriestandard

Entsprechend dieser Entwicklung im Kreditgeschäft wie auch im Aufsichtsrecht sind Ratingverfahren in den letzten Jahren ständig verfeinert worden. Im Retail-Banking stellen statistisch entwickelte Scorekarten zur Unterstützung der Kreditentscheidung mittlerweile einen Industriestandard dar. Aufgrund der regulatorischen Anforderungen nehmen Banken die Bonitätseinstufung ihres Kreditbestandes mittlerweile über Verhaltensscorekarten oder Bestandsscore-Modelle vor.

Auch im Corporate-Segment werden Expertensysteme zunehmend durch statistisch entwickelte Ratingmodelle verdrängt. Obwohl in diesem Bereich die Experteneinschätzung weiterhin eine wichtige Rolle bei der Bonitätsprüfung einnimmt, bildet mittlerweile in den meisten Banken ein statistisch entwickeltes Ratingmodell die Basis für die Bonitätsprognose durch den Kreditanalysten.

Die Trennkraft und Prognosegüte von statistischen Modellen hängt dabei in starkem Maße vom Umfang, von der Qualität, sowie der Repräsentativität der zur Verfügung stehenden Entwicklungsdaten ab. Statistisch entwickelte Scoringmodelle haben nicht zuletzt deshalb zunächst im Retailgeschäft ihre Anwendung gefunden, weil hier aufgrund der Vielzahl von Einzelgeschäften umfangreiche Daten zur Verfügung stehen. Mit dem Aufbau von Rating-Systemen und -datenbanken wächst der Datenumfang jedoch auch im Corporate-Bereich ständig.

Ein wichtiger Baustein zur Entwicklung von trennscharfen und genauen Ratingmodellen ist in beiden Bereichen – im Kreditscoring wie im Unternehmens-Rating – der adäquate Umgang mit fehlenden Daten. Der Begriff „fehlende Daten“ oder besser „fehlende Werte“ (Missing Va-

lues) bezeichnet nicht etwa das komplette Fehlen von Ratingdaten zu einem Geschäft oder Kreditnehmer. Vielmehr spricht man von fehlenden Werten, wenn einzelne Informationen, die für die Bonitätseinstufung eine zentrale Rolle spielen, nicht verfügbar sind. Beispiele für Größen, die fehlende Werte aufweisen können, sind etwa Bilanzkennzahlen, Selbstauskünfte und Büroauskünfte.

Umgang mit fehlenden Daten

Einerseits kann das Auftreten von fehlenden Werten in den Datenbeständen ein Zeichen von schlechter Datenqualität sein. Andererseits treten jedoch häufig auch Konstellationen auf, in denen das Fehlen eines Wertes fachlich korrekt ist. Dies kann beispielsweise vorkommen, wenn die Erhebung von Daten zur Bonitätseinstufung im Kreditprozess nicht einheitlich geschehen soll. So kann beispielsweise der Kreditprozess vorsehen, dass Büroauskünfte erst ab einem bestimmten Gesamtengagement eingeholt werden. In diesem Fall wird man im Portfolio Kreditnehmer mit Büroauskunft und Kreditnehmer mit richtigerweise fehlender Büroauskunft beobachten. Möglicherweise gibt es den zusätzlichen Fall, dass der Versuch unternommen wurde, eine Büroauskunft einzuholen, der Kreditnehmer aber beim Auskunftsbüro nicht bekannt war. In diesem Fall hat man einen weiteren fehlenden Wert, der jedoch fachlich eine zusätzliche Information trägt.

Das Beispiel zeigt, dass fehlende Werte nicht zwangsläufig durch Verbesserungen im Ratingprozess oder in der Datenhaltung vermieden werden können. Sie können vielmehr korrekter Bestandteil eines Ratingmodells sein. Folglich sind fehlende Werte bei Anwendung und Entwicklung eines Ratingmodells zu berücksichtigen.

Dr. Bernd Appasamy, Dr. Eike Bick, Dr. Jochen Meyer, alle d-fine GmbH, Frankfurt am Main, und Holger Flick, DaimlerChrysler-Bank AG, Stuttgart

Für die Unterstützung von Kreditentscheidungen der Banken und eine tragfähige Bonitätseinstufung ihrer Kunden sind in der Praxis oft Informationen nötig, die aus irgendwelchen Gründen nicht verfügbar sind. Im Umgang mit diesen Datenlücken greifen die Autoren auf geeignete Schätzwerte zurück, die sich ihrerseits wieder mit einfacheren oder komplexeren Rechenmethoden entwickeln lassen. Nach einer unerlässlichen statistischen Analyse der Charakteristik der fehlenden Werte halten sie für die Modellentwicklung, etwa von Bilanzratings, den Einsatz von ausgefeilten statistischen Methoden für gerechtfertigt. Für die Bonitätsprognose bevorzugen sie hingegen pragmatische, aber gleichwohl effiziente Ersetzungsverfahren. (Red.)

Zeitschrift für das gesamte Kreditwesen



Verlag und Redaktion:

Verlag Fritz Knapp GmbH
Aschaffener Str. 19, 60599 Frankfurt,
Postfach 11 11 51, 60046 Frankfurt.

Telefon: (0 69) 97 08 33 - 0, Telefax: (0 69) 7 07 84 00
E-Mail: red.zfgk@kreditwesen.de
Internet: www.kreditwesen.de

Herausgeber: Klaus-Friedrich Otto

Chefredaktion: Dr. Berthold Morschhäuser,
Philipp Otto

Redaktion: Swantje Benkelberg, Lars Haugwitz,
Alexander Hofmann, Frankfurt am Main

Redaktionssekretariat: Gisela Nuhn

Die mit Namen versehenen Beiträge geben nicht immer die Meinung der Redaktion wieder. Bei unverlangt eingesandten Manuskripten ist anzugeben, ob dieser oder ein ähnlicher Beitrag bereits einer anderen Zeitschrift angeboten worden ist. Beiträge werden nur zur Alleinveröffentlichung angenommen.

Die Zeitschrift und alle in ihr enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig.

Manuskripte: Mit der Annahme eines Manuskripts zur Veröffentlichung erwirbt der Verlag vom Autor das ausschließliche Verlagsrecht sowie das Recht zur Einspeicherung in eine Datenbank und zur weiteren Vervielfältigung zu gewerblichen Zwecken in jedem technisch möglichen Verfahren. Die vollständige Fassung der Redaktionsrichtlinien finden Sie unter www.kreditwesen.de.

Verlagsleitung: Uwe Cappel

Anzeigenleitung:

Ralf Werner, Tel. (0 69) 97 08 33-43

Anzeigendisposition:

Alexandra Knab, Tel. (0 69) 97 08 33-33,

sämtl. Frankfurt am Main, Aschaffener Str. 19.

Zurzeit ist die Anzeigenpreisliste Nr. 48 vom 1.1.2006 gültig.

Zitierweise: KREDITWESEN

Erscheinungsweise: am 1. und 15. jeden Monats.

Bezugsbedingungen: Abonnementspreise incl. MwSt. und Versandkosten: jährlich € 348,96, bei Abonnements-Teilzahlung: 1/2-jährlich € 179,40, 1/4-jährlich € 92,22. Ausland: jährlich € 358,56. Preis des Einzelheftes € 17,90 (zuzügl. Versandkosten).

Bestandteil des Abonnements ist eine 4-mal jährlich erscheinende Ergänzungsausgabe (Supplement) „Technik – IT für Finanzdienstleister“.

Verbundabonnement mit der Zeitschrift »bank und markt«: € 549,12, bei Abonnements-Teilzahlung: 1/2-jährlich € 286,80, 1/4-jährlich € 150,06. Ausland: jährlich € 558,72.

Studenten: 50% Ermäßigung (auf Grundpreis).

Der Bezugszeitraum gilt jeweils für ein Jahr. Er verlängert sich automatisch um ein weiteres Jahr, wenn nicht einen Monat vor Ablauf dieses Zeitraumes eine schriftliche Abbestellung vorliegt. Bestellungen direkt an den Verlag oder an den Buchhandel.

Bei Nichterscheinen ohne Verschulden des Verlags oder infolge höherer Gewalt entfallen alle Ansprüche.

Bankverbindungen: Landesbank Hessen-Thüringen – Girozentrale – 10 555 001 (BLZ 500 500 00), Frankfurt am Main.

Druck: Kern & Birner GmbH + Co., Werrastraße 4, 60486 Frankfurt am Main.

ISSN 0341-4019



Damit ergeben sich konkret zwei Fragen: Wie geht man bei der statistischen Modellentwicklung mit fehlenden Werten in den vorliegenden Daten um? Wie wird im Einzelfall trotz fehlender Werte eine angemessene Bonitätsprognose berechnet?

Sicherheitszuschläge

Beide Fragen werden zumindest teilweise durch die Mindestanforderungen der SolV beantwortet. So sind in der Modellentwicklung bei Datenknappheit entsprechende Sicherheitszuschläge zu berücksichtigen. Bei der Bonitätsprognose gilt: je weniger Informationen vorliegen, desto konservativer muss die Einstufung sein. ⁽¹⁾ §112 (2), §128 (1), §128 (6)). Eine rein konservative Herangehensweise führt allerdings zu einer systematischen Überschätzung der Ausfallwahrscheinlichkeit. Aus methodischer Sicht und vor dem Hintergrund einer adäquaten Risikosteuerung ist es hingegen wünschenswert, zu einer statistisch abgesicherten Bewertung der Ausfallwahrscheinlichkeit zu gelangen, die der Struktur der fehlenden Informationen Rechnung trägt.

In der Statistik sind Verfahren zum Umgang mit fehlenden Werten seit mehreren Jahren bekannt und erprobt.²⁾ Sie finden beispielsweise Anwendung in den Politikwissenschaften bei Wahlprognosen oder auch in klinischen Studien zur Einführung neuer Medikamente. Die bei Banken benutzten Statistiksoftware-Pakete (etwa SAS, S-Plus) unterstützen die entsprechenden Methoden.³⁾ In der Ratingentwicklung sind diese Verfahren jedoch bisher weitgehend unberücksichtigt geblieben. Die Grundidee beim Umgang mit fehlenden Werten ist, diese mit geeigneten Schätzwerten zu befüllen. Die so vervollständigten Datensätze können Grundlage einer statistischen Modellentwicklung werden, die das vollständige Vorliegen der Datensätze voraussetzen. Das Befüllen der fehlenden Werte wird in der Literatur als „Imputing“ bezeichnet.

Struktur der fehlenden Werte

Die Auswahl eines geeigneten Imputing-Verfahrens setzt eine Analyse der Struktur der fehlenden Werte voraus. Hier werden die folgenden Fälle unterschieden²⁾:

Völlig zufällig fehlende Werte: Bei völlig zufällig fehlenden Werten hängt die

Struktur der fehlenden Werte nicht von den Daten ab. Völlig zufällig fehlende Werte treten oft im Fall technischer Probleme auf – wenn beispielsweise ein Datenbanklauf zur Überspielung von Kreditnehmerbilanzdaten aus technischen Gründen abbricht, besteht meist kein Zusammenhang zwischen der Existenz fehlender Werte und ökonomischen Eigenschaften (wie Branche oder Umsatz) des Kreditnehmers. In der Literatur wird dieser Fall als „MCAR“ (Missing Completely at Random) bezeichnet.

Zufällig fehlende Werte: Bei zufällig fehlenden Werten kann die Struktur der fehlenden Werte von den beobachteten Daten abhängen. Solch eine Fehlstruktur findet man beispielsweise in Firmenkunden-Daten von Ratingssystemen, in denen detaillierte Bilanzkennzahlen nur dann erhoben werden, wenn das Gesamtbligo eines Kunden einen bestimmten Wert übersteigt. In diesem Fall spricht man von einem „zufälligen Fehlen der Kennzahlen“. Sie fehlen aber eben nicht völlig zufällig im obigen Sinne, da ja abhängig vom Obligo die Bilanzen eingeholt werden. Dieser zweite Fall wird auch als „MAR“ (Missing at Random) bezeichnet.

Nicht zufällig fehlende Werte: Die MAR-Annahme (und damit auch MCAR) ist nicht erfüllt, wenn die Fehlstruktur von den fehlenden Werten selbst abhängt. Würden im vorangegangenen Beispiel Bilanzen in Abhängigkeit von dem Unternehmensumsatz eingeholt, so würde die Kennzahl „Unternehmensumsatz“ nicht zufällig fehlen. Vielmehr würde der Unternehmensumsatz immer dann fehlen, wenn er einen gewissen Wert unterschreitet.

Nach der Analyse der Fehlstruktur eines Datensatzes muss ein angemessenes Verfahren ausgewählt werden, um mit den fehlenden Werten umzugehen.

Einfache und anspruchsvollere Ansätze

Die einfachsten Ansätze bestehen darin, beispielsweise alle Beobachtungen mit fehlenden Werten zu löschen (Listwise Deletion) oder etwa die fehlenden Werte einer Variablen durch Median oder Mittelwert der in den restlichen Datensätzen beobachteten Verteilung dieser Variablen zu ersetzen. Diese Vorgehensweisen führen jedoch allenfalls in Spezialfällen zu befriedigenden Ergebnissen.



Eine weitere Möglichkeit ist, die fehlenden Werte als Ausprägung einer kategorischen Variablen zu behandeln. Ist die Ausgangsvariable jedoch eine metrische Variable (zum Beispiel eine Bilanzkennzahl) oder eine quasi-metrische Variable (zum Beispiel das Alter eines Kunden), so ist mit einer Kategorisierung prinzipiell ein Verlust an Information verbunden. Um bei diesem Ansatz die Verminderung der Trennschärfe zu minimieren, ist ein äußerst sorgfältiges Vorgehen bei der Festlegung der Kategorisierung notwendig.

Modellentwicklung und Bonitätsprognose

Moderne statistische Methoden bieten darüber hinaus verschiedene Möglichkeiten, auch mit höheren Anteilen fehlender Daten umzugehen. So stellt der EM-Algorithmus (Expectation-Maximization-Algorithmus) kombiniert mit einem stochastischen Verfahren wie MCMC (Markov Chain Monte Carlo) eine leistungsfähige Methode für die multivariate Ersetzung (Multivariate Imputing) fehlender Werte zur Verfügung. Im Unterschied zu den einfachen Ersetzungsverfahren erzeugt man hier mehrere Varianten des Originaldatensatzes. In diesen Varianten sind die fehlenden Werte dann jeweils durch stochastisch gezogene Werte ersetzt. Die gewünschten Analysen (zum Beispiel eine logistische Regression zur Ermittlung einer Scorefunk-

tion) werden dann auf allen Varianten durchgeführt. Die Ergebnisse, wie beispielsweise Regressionskoeffizienten, erhält man dann durch die Mittelung über die Ergebnisse für alle einzelnen Varianten.

In der konkreten Anwendung auf Scoring- und Ratingmodelle lassen sich multivariate Ersetzungsverfahren für die eingangs gestellten Fragen der Modellentwicklung und der Bonitätsprognose unterschiedlich gut anwenden.

In der Modellentwicklung stellen multivariate Ersetzungsverfahren ein wirkungsvolles Werkzeug zur Behandlung von fehlenden Werten dar. Diese Verfahren besitzen den Vorteil, dass sie besonders geeignet sind, metrische oder quasi-metrische Risikomerkmale wie Bilanzdaten oder Bonitätsindizes in der Modellentwicklung zu berücksichtigen. Der Grund hierfür liegt darin, dass durch das mehrfache Imputing der fehlenden Werte die Wahrscheinlichkeitsverteilungen der zugrunde liegenden Risikofaktoren respektiert werden. Zudem erlaubt die Wiederholung von Analysen auf Varianten des Originaldatensatzes eine Abschätzung der statistischen Fehler, die durch die fehlenden Daten verursacht werden.

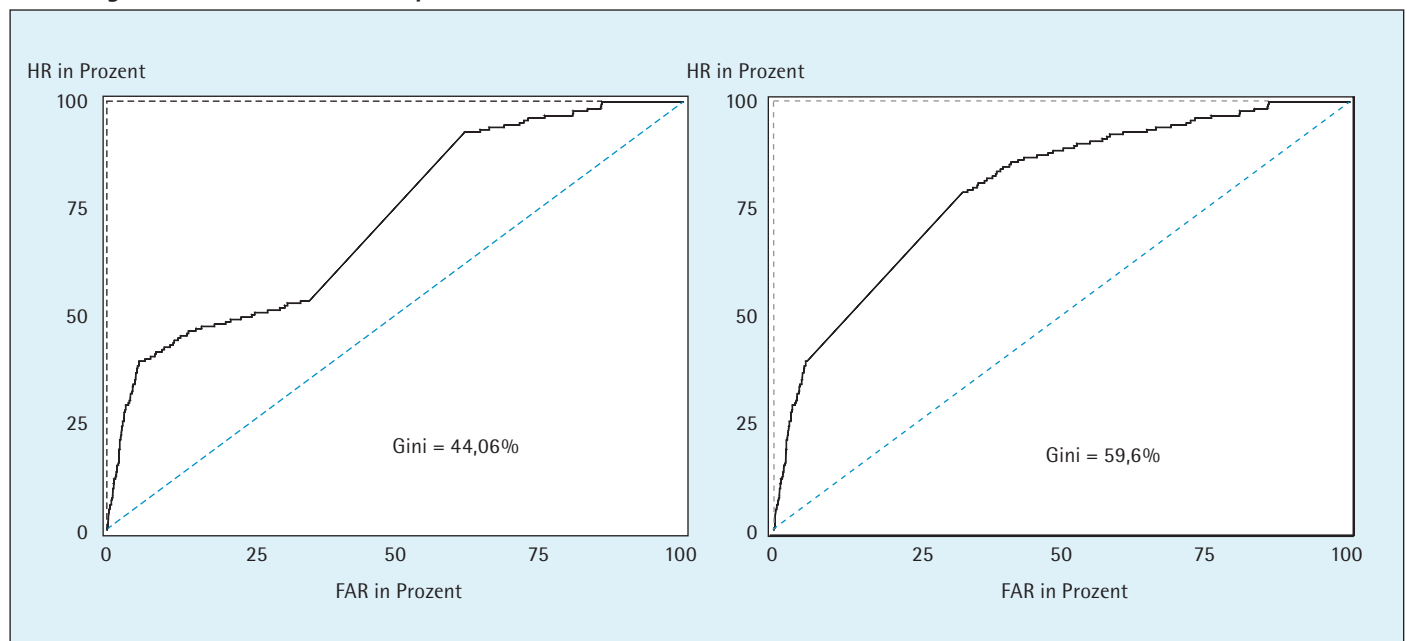
Bei der Bonitätsprognose, das heißt in der Anwendung des Ratingmodells auf einen einzelnen Vertrag oder Kunden haben mul-

tivariate Ersetzungsverfahren jedoch den entscheidenden Nachteil, dass die Bonitätsprognose nicht mehr deterministisch von den Eingangsgrößen abhängt. Vielmehr ist der Zusammenhang zwischen Risikofaktoren und Bonitätsprognose durch eine stochastische Ersetzung fehlender Werte gegeben. Damit ist die regulatorisch geforderte Nachvollziehbarkeit eines Ratingergebnisses nicht mehr gegeben. In diesem Fall sind also einfachere Verfahren gefragt.

Gefahr des Trennkraftverlustes

Die nahe liegende Ersetzung fehlender Werte durch den Mittelwert oder den Median der beobachteten Werte erweist sich in der Praxis als unvorteilhaft. Insbesondere bei einer MAR-Fehlstruktur der Daten ist dieser Ansatz ungünstig. Bereits wenn der Anteil der Beobachtungen mit fehlenden Werten fünf Prozent übersteigt, kann dieses Vorgehen zu einer erheblichen Verminderung der Vorhersagekraft des Risikofaktors führen. Handelt es sich um einen wichtigen Risikotreiber, so ergibt sich damit immer auch eine Verminderung der Trennkraft des Gesamtmodells. Der Grund hierfür ist darin zu suchen, dass insbesondere unter MAR die Ausfallwahrscheinlichkeit von Beobachtungen mit fehlenden Werten mitunter deutlich vom Portfolio-mittel abweicht. Entsprechend werden bei einer Zuweisung des Mittelwerts oder Me-

Abbildung: ROC-Kurve der Gesamtkapitalrentabilität – 25 Prozent fehlende Werte wurden ersetzt



dians die Beobachtungen mit fehlenden Werten mit einem gemessen an der tatsächlichen Ausfallwahrscheinlichkeit entweder zu hohen oder zu niedrigen Score versehen.

Eine bessere Ersetzungsvorschrift muss daher die unterschiedlichen a priori Ausfallwahrscheinlichkeiten (PD) von Beobachtungen mit und ohne fehlende Werte berücksichtigen. Als Beispiel sei ein Risikofaktor GKR „Gesamtkapitalrentabilität“ betrachtet, der in 25 Prozent aller Beobachtungen fehlt: Sind die a priori Ausfallwahrscheinlichkeiten der Beobachtungen mit vorhandener Gesamtkapitalrentabilität und die mit fehlender Gesamtkapitalrentabilität unterschiedlich, so führt eine Ersetzung der fehlenden Werte durch den Mittelwert oder den Median zu einem Trennkraftverlust des Risikofaktors.

Dieser Trennkraftverlust zeigt sich auch recht anschaulich in den Receiver Operating Characteristics (ROC-Kurve) [4,5]. Das Argument wird deutlich, wenn man sich den Zusammenhang zwischen Ausfallwahrscheinlichkeit und der Steigung der ROC-Kurve vergegenwärtigt. Betrachtet man als Beispiel die ROC-Kurve des Risikofaktors Gesamtkapitalrentabilität. In der ROC-Kurve wird zu einem gegebenen Cutoff C die Trefferquote HR(C) (hit rate) gegen die Fehlalarmquote FAR(C) (false alarm rate) aufgetragen (vergleiche [4,5]). Die Steigung der ROC-Kurve HR(FAR) an der Stelle, die einem Score s entspricht, ist dann Formel 1:

$$\begin{aligned} \left. \frac{\partial \text{HR}(\text{FAR}(\text{C}))}{\partial \text{FAR}} \right|_{\text{C=s}} &= \left. \frac{\partial \text{HR}}{\partial \text{C}} \frac{\partial \text{C}}{\partial \text{FAR}} \right|_{\text{C=s}} \\ &= \left. \frac{\partial \text{HR}}{\partial \text{C}} \left[\frac{\partial \text{FAR}}{\partial \text{C}} \right]^{-1} \right|_{\text{C=s}} \\ &= \frac{\phi^D(s)}{\phi^A(s)} \end{aligned}$$

In Worten: Die Steigung der ROC-Kurve bei einem gegebenem Score s ist nichts anderes als das Likelihood-Ratio bei diesem Score. Das Likelihood-Ratio ist gerade

durch das Verhältnis der Wahrscheinlichkeitsdichten $\phi^D(s)$ und $\phi^A(s)$ gegeben. Für die PD gilt dann

$$\text{PD}(s) = \frac{1}{1 + \frac{1 - \overline{\text{PD}}}{\text{PD}} \frac{\phi^A(s)}{\phi^D(s)}}$$

mit der mittleren Portfolio-Ausfallwahrscheinlichkeit $\overline{\text{PD}}$. Daraus folgt, dass ein direkter Zusammenhang zwischen Steigung der ROC-Kurve und der lokalen PD besteht.

Ersetzung von Bilanzkennzahlen

Werden jetzt die 25 Prozent fehlenden Werte der Gesamtkapitalrentabilität mit einem festen Wert (zum Beispiel GKR = 0,125 als Median der beobachteten Gesamtkapitalrentabilität) ersetzt, so erscheinen die ersetzten Werte in der ROC-Kurve als ein Geradenstück. Die Steigung dieses Abschnittes der Kurve entspricht der a priori PD aller Beobachtungen mit fehlenden (ersetzten) Werten. Gleichzeitig entspricht jeder der Werte der Gesamtkapitalrentabilität aus den vorhandenen Werten einer „lokalen“ PD. Um optimale Trennkraft zu erreichen, ist also der Ersetzungswert so zu wählen, dass die a priori PD zur lokalen PD passt.

Die Abbildung zeigt zwei ROC-Kurven für die Bilanzkennzahl Gesamtkapitalrentabilität mit der zugehörigen Trennschärfe (gemessen als Gini-Koeffizient [4,5]). Die Bilanzdaten wurden auf einem Portfolio großer und mittlerer internationaler Unternehmen beobachtet. Die Kurve im linken Graphen ergibt sich, indem 25 Prozent fehlender Werte durch den Median der nicht fehlenden Werte ersetzt wurde. Im rechten Graphen wurde bei der Ersetzung der fehlenden Werte die a priori PD adäquat berücksichtigt.

Eine weitere Anwendung der in der Abbildung demonstrierten Ersetzung fehlender Werte unter Berücksichtigung der a priori PD wurde von den Autoren im Rahmen der Weiterentwicklung von Scoringmodellen für das gewerbliche Retail-Automotive-Portfolio der Daimler-Chrysler-Bank umgesetzt. Für Bonitätseinschätzungen von Kreditnehmern in diesem Portfolio werden von der Bank häufig Bankauskünfte zur Bonitätsbeurteilung von Kreditnehmern eingeholt. Die Textbausteine von Bankauskünften werden automatisiert erfasst und durch ein Scoringmodell bewertet. Das Ergebnis ist eine Bankauskunftsnote mit De-

zimalwerten 1.0 bis 6.0, in Anlehnung an das Schulnotensystem.

Der Kreditprozess sieht vor, dass Bankauskünfte erst ab einem gewissen Risiko zur Kreditentscheidung eingeholt werden müssen. Dies führt dazu, dass im Portfolio nur für etwa 50 Prozent aller Kunden eine Bankauskunft vorliegt. Darüber hinaus ist es bei etwa drei bis vier Prozent der eingeholten Bankauskünfte nicht möglich eine Bankauskunftsnote zu berechnen. Für die Bankauskunft gibt es also zwei Arten von fehlenden Werten: a) Bankauskunft nicht eingeholt oder b) Bankauskunft eingeholt, aber nicht bewertbar. Analysen zeigen, dass die beobachteten Kreditnehmer mit fehlender Bankauskunft je nach Variante a) oder b) in der Tat unterschiedliche a priori Ausfallwahrscheinlichkeiten aufweisen.

Adäquate Ersetzung fehlender Bankauskünfte

Durch eine adäquate Ersetzung der fehlenden Bankauskünfte war es möglich, ein trennscharfes Scoringmodell für die Bewertung des Kreditbestandes zu entwickeln, obwohl die Bankauskunft in etwa 50 Prozent aller Beobachtungen fehlte. Dabei wies die Bankauskunft nach Ersetzung eine univariate Trennkraft von etwa Gini = 24 Prozent auf und verbesserte die Trennkraft des Gesamtmodells um etwa zehn Prozentpunkte.

Das dargestellte Beispiel zeigt deutlich, das bei fehlenden Daten eine genaue statistische Analyse der Charakteristik der fehlenden Werte unerlässlich ist. Ausgefeilte statistische Verfahren wie der EM-Algorithmus bieten viele Vorteile für die Modellentwicklung beispielsweise von Bilanzratings. Für die Bonitätsprognose hingegen sind pragmatische aber effiziente Ersetzungsverfahren anzuwenden.

Literatur

- ¹⁾ Entwurf der Solvabilitätsverordnung, 31. März 2006
- ²⁾ G. King, J. Honacker, A. Joseph, K. Scheve; Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation; American Political Science Review, March 2001, Vol. 95, No. 1
- ³⁾ J. Horton, R. Lipsitz; Multiple Imputation in Practice: Comparison of Software Packages for Regression Models With Missing Variables; The American Statistician, August 2001, Vol. 55, No. 3
- ⁴⁾ Basel Committee on Banking Supervision; Working Paper 14 - Studies on the validation of Rating System; Februar 2005
- ⁵⁾ J. Sobehart, S. Keenan, Measuring default accurately, Risk, März 2001

Redaktionshinweis

Dieser Ausgabe liegt das Sachverzeichnis des zweiten Halbjahres 2006 der ZfgK bei.

Zeitschriftenvertrieb – Hans Berger
(Telefon 0 69/97 08 33-32) oder Barbara
Himmelreich (Telefon 0 69/97 08 33-25)

