

Konzeptionsgrundsätze erfolgreicher Scoring-Systeme

Das Fortschreiten der digitalen Möglichkeiten verändert den Verantwortungsbereich des Credit Managers nachhaltig. Die zunehmende Verfügbarkeit relevanter Kundeninformationen wird den Blick auf den Kunden weiter schärfen. Damit müssen zukünftig noch stärker die wesentlichen Informationen herausgefiltert, weiterverarbeitet und zusammengefasst werden. Scoring-Modelle können den Credit Manager dabei unterstützen und seine Effizienz steigern. Im Gegensatz zur traditionellen Bonitätsprüfung, die maßgeblich auf den persönlichen Erfahrungen des Credit Managers beruht und entsprechend ein hohes Maß an Subjektivität aufweist, handelt es sich beim Credit-Scoring um ein formalisiertes Verfahren, das eine objektivere Kreditentscheidung anstrebt. Die folgenden Ausführungen sind eine knappe Zusammenfassung der BvCM-Studie "Scoring-Modelle: 12 Fragen und Antworten", die die hier aufgeführten Konzepte mit vielen Beispielen und nachvollziehbaren Rechnungen im Detail vorstellt.

Definition und Grundlagen

Scoring-Modelle sind analytische Bewertungsverfahren. Sie sollen Hinweise auf die Wahrscheinlichkeit geben, mit der ein Geschäftspartner ein bestimmtes zukünftiges Verhalten zeigt. Scoring-Modelle fassen Merkmale, die in Punktwerten operationalisiert werden, zu einem Gesamt-Score zusammen. Dieser spiegelt letzten Endes die Häufigkeit wider, mit der ein Ereignis bei ähnlichen Merkmalsträgern in der Vergangenheit aufgetreten ist. Eine wesentliche Stärke des Ansatzes ist, dass er objektive Bewertungen liefert. Dies fördert die Akzeptanz der Urteile bei den Nutzern. Das Scoring findet dabei in mehren Schritten statt.

Im **ersten Schritt** muss das zu prognostizierende Verhalten genau beschrieben und der Prognosehorizont (z.B. 3 Monate oder 1 Jahr) festgelegt werden.

Im **zweiten Schritt** werden alle sinnvollen Merkmale (Rohdaten) zusammengestellt, verifizierbar sind.



Konzeptionsgrundsätze erfolgreicher Scoring-Systeme

Im dritten Schritt werden die Rohdaten in Punktwerte transformiert und deren Eignung für das Scoring-Modell zunächst univariat überprüft. Die Gewichtung der ausgewählten Merkmale ist dabei der wesentliche Faktor. Handelt es sich um ein optimales System, dann kann aus dem Scorewert direkt auf die Wahrscheinlichkeit für das im ersten Schritt definierte Verhalten geschlossen werden.

Im vierten Schritt werden für jedes Scoring-Objekt die Merkmalsausprägung gemessen, der dazugehörige Punktwert ermittelt und zu einem Gesamtscore zusammengefasst.

Ist nun jeder Kunde bewertet, werden in einem **fünften Schritt** Klassen gebildet. Die einzelnen Klassen fassen diejenigen Kunden zusammen, die ähnliche Wahrscheinlichkeiten für das zu prognostizierende Verhalten aufweisen.

Im **sechsten Schritt** (Evaluation) wird überprüft, ob die anhand der Scores entwickelten Klassen tatsächlich geeignet sind, um die Kunden im Hinblick auf ihr zukünftiges Verhalten zu unterscheiden.

Funktionsweise von Scoring Modellen

Der Grundgedanke von Scoring-Modellen soll anhand einer vereinfachten Darstellung eines Bonitäts-Scorings vermittelt werden. Das zukünftige Ereignis wird folgendermaßen definiert: "Der Kunde kommt nicht fristgerecht – definiert bspw. als eine Überfälligkeit von max. 90 Tagen - und vollumfänglich seinen finanziellen Verpflichtungen nach". Der Prognosezeitraum wird auf 1 Jahr festgelegt. Aufgabe des Systems ist es, für jeden Kunden die Eintrittswahrscheinlichkeit für eben dieses Ereignis zu prognostizieren. Hierzu müssen kundenspezifische Merkmale gefunden werden, die Vorhersagekraft bzw. einen Vorlauf für dieses Ereignis aufweisen. Beispielhaft wählen wir hierzu 3 Datensätze aus und unterstellen folgende Wirkungszusammenhänge:

- Firmenalter (gemessen in Jahren seit der Gründung) als Indikator für die unternehmerische Erfahrung: Mit zunehmenden Firmenalter sinkt die Insolvenzwahrscheinlichkeit
- Eigenkapitalquote (gemessen als Eigenkapital in % der Bilanzsumme) als Indikator für die Risikotragfähigkeit: Mit steigender Eigenkapitalquote sinkt die Insolvenzwahrscheinlichkeit
- EBIT-Marge (gemessen als EBIT in % des Umsatzes) als Indikator für unternehmerischen Erfolg: Mit steigender EBIT-Marge sinkt die Insolvenzwahrscheinlichkeit.





Anschließend werden die Merkmale in Punktwerte in zwischen 1 (geringes Ausfallrisiko) und 5 (hohes Ausfallrisiko) umgewandelt. Damit sind die Voraussetzungen für die Berechnung eines Gesamt-Score geschaffen. Vorher muss allerdings noch die Gewichtung der einzelnen Punktwerte festgelegt werden. Kunden mit dem gleichen Gesamtscore werden dann als eine Klasse aufgefasst.

Relevante Daten für die Ausfallprognose

Ob ein Kunden seinen finanziellen Verpflichtungen nachkommt, hängt von den folgenden, häufig als "fünf C" bezeichneten, Faktoren ab (siehe hierzu Damodaran, 1997, S. 376 und Streuele, 2009, S. 99 ff):

- Character: Der Wille des Kunden, die Forderung zu bezahlen
- Capacity: Die Fähigkeit des Kunden, genügend Cashflows für die Begleichung der Forderung zu erarbeiten.
- Capital: Die finanziellen Ressourcen des Kunden
- Colateral: Die beim Kunden vorhanden Sicherheiten
- Conditions: Das wirtschaftliche Umfeld des Unternehmens

Um diese fünf C im B2B-Geschäft zu messen, stehen unterschiedliche Informationsquellen zur Verfügung. Die Credit Manager ziehen traditionell – insbesondere bei Neukunden – **externe Daten** zur Urteilfindung zu Rate. Hier bieten sich neben (1) Firmengrunddaten wie Unternehmensgröße, Rechtsform und Unternehmenssitz hauptsächlich (2) Jahresabschlussdaten an. Zusätzlich können externe Daten auch in Form von (3) Bankauskünften, (4) Selbstauskünften der Kunden – etwa mit Steuerunterlagen – und (5) Ratings spezialisierter Informationsanbieter wie Auskunfteien eingeholt werden.

Daneben spielen **Daten aus internen Quellen** eine sehr wichtige Rolle. Diese sind im Gegensatz zu Finanzkennzahlen ohne zeitliche Verzögerung sowie exklusiv verfügbar. Einen hohen Stellenwert hat dabei das (1) historische Zahlungsverhalten. Wichtige Hinweise liefern daneben die (2) vergangenen Mahnaktivitäten – und die entsprechenden Reaktionen des Unternehmens – sowie (3) die Limitenausnutzungen. Zu guter Letzt sind (4) Informationen der operativen Einheiten hilfreich.

Häufig werden in Scoring-Modellen auch **subjektive Einschätzungen** verarbeitet, die insbesondere die beiden C´s *Character und Conditions* beurteilen sollen. Dabei können grundsätzlich drei verschiedene Teilbereiche unterschieden werden: (1) Markt und Wettbewerb, (2) Strategie und (3) Management.





Auswahl der relevanten Merkmale

Grundsätzlich soll der Einfluss der Merkmale auf das zu prognostizierende Ereignis eindeutig ökonomisch fundiert sein. Die einzelnen Merkmale sollen daneben unabhängige Informationen liefern. Zudem soll jedes Merkmal zur Unterscheidung zwischen den guten und den schlechten Ereignissen beitragen. Die Beurteilung erfolgt in Form einer so genannten univariaten Analyse. Damit die Einzelbetrachtung systematisch durchgeführt werden kann, sind im Idealfall für jede Merkmalsklasse Schlechtereignisse verfügbar.

Gewichtung der ausgewählten Merkmale

Der statistische Ansatz greift unter anderem auf das Verfahren der Diskriminanzanalyse zurück. Deren Ziel ist eine möglichst gute Trennung einer abhängigen Variablen Y (z.B. solventen von insolventen Kunden) mithilfe einer Funktion aus mehreren unabhängigen Variablen X (z.B. Unternehmensalter, EKQ, Umsatzrendite, Standort, usw.). In den letzten Jahren hat sich zunehmend die logistische Regression als Gewichtungsverfahren etabliert. Mit diesem Verfahren wird die Wahrscheinlichkeit ermittelt, mit der der betrachtete Merkmalsträger seinen Zahlungsverpflichtungen nicht nachkommt . Der zentrale Vorteil des statistischen Ansatzes liegt in seiner objektiven Berechnung der Gewichte. Das führt zu einer hohen Nachvollziehbarkeit der Ergebnisse. Wissenschaftlich belegt ist zudem seine hohe Trennschärfe im Vergleich zum Expertenansatz, bei dem die Gewichte subjektiv bestimmt werden und die Anforderungen an die verwendeten Daten deutlich niedriger sind. Allerdings stellt sich für Unternehmen nicht die Frage nach dem zu wählenden Ansatz. Die bescheidene Datenverfügbarkeit verhindert gerade bei der Neuentwicklung meist das statistisch fundierte Vorgehen. Dieses ist häufig erst zu einem späteren Zeitpunkt möglich – und drückt entsprechend auch einen gewissen Reifegrad des gesamten Scoring-Systems aus.

Gütekriterien eines Scoring-Systems

Einem Scoring-System kann dann eine hinreichende Prognosegüte attestiert werden, wenn gute Scores mit einer niedrigen Ausfallwahrscheinlichkeit einhergehen und umgekehrt. In diesem Sinne ist die Prognosegüte ist umso größer, je größer die Differenz zwischen dem Score-Durchschnitt (Mittelwert) der ausgefallenen Unternehmen und dem Score-Durchschnitt der nicht-ausgefallenen Unternehmen. Diese Kennzahl wird als **Trenndistanz** bezeichnet. Je größer der Wert der Kennzahl, desto besser ist das Scoring.

Die am weitesten verbreitete Kennzahl ist der **Gini-Koeffizient**. Sein Wertebereich liegt wie oben angedeutet zwischen 0% und 100%. Ab welchem Gini-Koeffizienten



ein Ratingsystem nun gute Qualität aufweist, kann nur in der relativen Betrachtung beantwortet werden. Im B2B-Bereich liegen sehr gute Credit-Scoring-Modelle im Bereich von 55-70% (Siehe Schulte-Mattler et. al., 2004, S. 50). Nach Erfahrungen von Munch (2009, S. 212) sind im Geschäftsfeld B2C Scoring-Modelle Gini-Koeffizienten zwischen 45% und 50% realistisch.

Ablehnungsverhalten in der Praxis

Die für den Entscheidungsträger in der Praxis entscheidende Frage lautet: Soll dem Kunden mit einem Score von XY ein Kredit gewährt werden oder nicht? Oder konkreter: Was ist der exakte Schwellenwert des Scoring-Systems, ab dem einem Kunden der Kredit verweigert werden sollte? Im Idealfall trennt ein Schwellenwert die Kunden ex ante in die Gruppe der guten Risiken und die Gruppe der schlechten Risiken. In der Praxis fallen allerdings auch Kunden mit verhältnismäßig guten Scores aus. Auf der anderen Seite kommen Kunden mit verhältnismäßig schlechten Scores ihren Zahlungsverpflichtungen nach. Für die Wahl des Schwellenwertes (sogenannter Cut-Off) gilt entsprechend:

- Wird der Schwellenwert sehr hoch angesetzt (werden also auch Kreditnehmer mit schwächeren Ratings akzeptiert), dann fallen viele Kunden aus, die als kreditwürdig klassifiziert werden. Wir sprechen hier vom sogenannten Fehler erster Art (Alpha Fehler). Die Kosten für diesen Fehler sind umso höher, je größer die Kreditbeträge und je tiefer die Umsatzmargen des Unternehmens sind.
- Wird dagegen der Schwellenwert sehr tief angesetzt, dann erhalten viele Kunden keinen Kredit, obwohl sie in der Lage wären ihren Verpflichtungen vollumfänglich nachzukommen. Wir sprechen hier vom sogenannten Fehler zweiter Art (Beta-Fehler). Hier fallen Kosten in Form entgangener Gewinne (so genannte Opportunitätskosten) und verlorener Marktanteile an, weil auch ex post gute Risiken keine Kredite erhalten würden.

Legt man einen sehr ambitionierten Schwellenwert fest, dann minimiert man zwar die Ausfälle (= Fehler erster Art), unterlässt auf der anderen Seite aber lukrative Geschäftsbeziehungen mit im Nachhinein solventen Kunden (= Beta-Fehler). Wie kann vor diesem Hintergrund ein optimaler Schwellenwert ermittelt werden? Die Beantwortung dieser Frage beginnt zunächst mit der Berechnung der durchschnittlichen Kosten für ein nicht-abgelehntes, schlechtes Risiko (= Kosten pro "Fehler erster Art –Kunde") und den durchschnittlichen Kosten für einen "Fehler zweiter Art – Kunden". Die Kosten für einen Fehler zweiter Art Kunden könnten approximativ berechnet werden als der gesamte Rohertrag aller Kunden geteilt durch die Anzahl der Nicht-Ausfälle. Das Verhältnis dieser beiden Kostenarten ermöglicht es, einen gewichteten Klassifikationsfehler zu berechnen.





Stabilität des Scoring Systems

Eine wichtige Anforderung an ein erfolgreiches Scoring-System ist dessen Stabilität. Sie zeichnet sich durch eine nachhaltige Modellierung des Zusammenhanges zwischen den Risikomerkmalen und der Bonität aus. Andernfalls verschlechtert sich die Prognosegüte sich im Zeitablauf signifikant (siehe Deutsche Bundesbank, 2003, S. 64). Stabilität kann mit so genannten **Stabilitätsindizes** (StI) gemessen werden. Was sind die Gründe für die Veränderung der Stabilität? Eine Antwort hierauf können sogenannte **Verschiebungsindizes** (ShI) geben. Dieses Instrument der Veränderungsanalyse ermittelt die Verschiebungsrichtung und die Größe des Einflusses auf den Score.

Wartung von Scoring Modellen

Um den langfristigen Erfolg von Scoring-Modellen zu gewährleisten, muss eine regelmäßige Kontrolle und Überprüfung implementiert werden. Ist nämlich die Aussagekraft des Systems gestört, dann kann dem anwendenden Unternehmen erheblicher Schaden entstehen. Modellanpassungen sind dann angezeigt, wenn sich **interne Faktoren** ändern. Verschlechtert sich bspw. die Datenverfügbarkeit wegen Umstellungen der IT-Systeme, dann leidet die Qualität des Scoring-Modells. Daneben können **externe Faktoren** die Qualität eines Scoring-Modells negativ tangieren, die nicht im unmittelbaren Einflussbereich des Unternehmens liegen. Hierzu zählen insbesondere Änderungen der wirtschaftlichen Rahmenbedingungen.

Vorteile von Scoring Modellen

Scoring Modelle bieten eine Reihe grundsätzlicher Vorteile aus organisatorischer und kaufmännischer Sicht (siehe Hartmann und Laurich, 2014, S. 63ff):

- Kosten: Kostengünstige Informationsbeschaffung
- Effizienz: Geringer Aufwand und hohe Aussagekraft
- Technik: Einfache technische Zugangs- und Zugriffsmöglichkeit
- Prozessintegration: Unkomplizierte Prozessintegration und gute Kompatibilität zu internen Abläufen
- Verfügbarkeit: An Prozesserfordernissen anpassbare Frequenz
- Flexibilität: Hohe Flexibilität und Individualisierbarkeit bzgl. spezifischer Bedürfnisse
- Einheitlichkeit: Gleiches Vorgehen bei der Bewertung einzelner Kunden bzw. Kundensegmente
- Nachvollziehbarkeit: Gute Verständlichkeit des Zustandekommens der Ergebnisse
- Akzeptanz: In der Regel hohe Akzeptanz beim Anwender.







Probleme von Scoring Modellen

Natürlich weisen Scoring-Modelle auch eine Reihe von Problemen auf (siehe für einige der nachfolgenden Argumente Schröder und Taeger, 2014, S. 11ff):

- Die Kalibrierung trennscharfer Scoring Modelle beruht auf einem zeitintensiven Entwicklungsprozess, der die volle Unterstützung des Managements benötigt.
- Die Anforderungen an die Inputdaten im Hinblick auf Aktualität und Oualität sind hoch.
- Die Geschäftsfälle, die bei der Entwicklung des Scoring-Systems ausgewertet wurden, müssen eine strukturelle Ähnlichkeit mit den zukünftig zu beurteilenden Geschäften haben
- Entscheidend für den Erfolg eines Scoring-Systems ist die Auswahl unabhängiger Merkmale.
- Das klassische Anwendungsgebiet von Scoring-Modellen ist das Massengeschäft. Bei großen Einzelengagements ist von einer Entscheidung, die ausschließlich auf dem Modellergebnis beruht, abzuraten.
- Die Vorteile von Scoring-Systemen zeigen sich vollumfänglich nur bei einer Portfolio-Perspektive.
- Credit-Scores können die Kreditkonditionen für bestimmte Kundengruppen verschlechtern.

Literatur

Damodaran, A. (1997), Corporate Finance – Theory and Practice, New York, Wiley.

Deutsche Bundesbank (2003), Validierungsansätze für interne Rating Systeme, Monatsbericht 9/2003, 61 – 74. Hartmann und Laurich (2014), Methodenbasis des Credit Management - Teil 1, Lehrmaterialien zum Certified Credit Manager.

Hock (2016), Scoring Modelle: 12 Fragen und Antworten, White Paper des Bundesverbandes für Credit Manage-

Munch (2009), Wie steht es um Ihr Scoring? - Das Rating von Scoring-Systemen, Finanzierung - Leasing - Factoring 5/2009, 209 - 213.

Rieder (2004), Sein oder nicht sein – Workout-Scoring als Konzept zur Optimierung von Rückflussquoten, Risknews 04/05, 19 - 24.

Schröder und Träger (2014), Scoring im Fokus: Ökonomische Bedeutung und rechtliche Rahmenbedingungen im internationalen Vergleich, BIS Verlag der Carl von Ossietzky Universität Oldenburg.

Schulte-Mattler, Daun und Manns (2004), "Trennschärfemaße zur Validierung von internen Rating-Systemen", RATINGaktuell 06/2004, S. 46 bis 52.

Streule (2009), Lieferantenkreditmanagement - Gewährung und Management von Lieferantenkrediten in der Schweiz, Dissertation, Universität St. Gallen.



