

Schewe, Gerhard; Leker, Jens

Working Paper — Digitized Version

Empirische Insolvenzforschung: Ein Vergleich ausgewählter Instrumente

Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 401

Provided in Cooperation with:

Christian-Albrechts-University of Kiel, Institute of Business Administration

Suggested Citation: Schewe, Gerhard; Leker, Jens (1996) : Empirische Insolvenzforschung: Ein Vergleich ausgewählter Instrumente, Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 401, Universität Kiel, Institut für Betriebswirtschaftslehre, Kiel

This Version is available at:

<https://hdl.handle.net/10419/177303>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Nr. 401

**Empirische Insolvenzforschung:
Ein Vergleich ausgewählter Instrumente**

Gerhard Schewe, Jens Leker

Kiel 1996

A. Problemstellung

Die Überprüfung der Krisengefährdung einer Unternehmung ist für unternehmensexterne Personen meist nur aufgrund veröffentlichter Informationen möglich. Eine in diesem Zusammenhang regelmäßig betrachtete unternehmensspezifische Informationsquelle ist der veröffentlichte Jahresabschluß.¹⁾ Die vorliegende Untersuchung setzt hier an. Es werden drei moderne Instrumente der Insolvenzdiagnose miteinander verglichen, deren Aufgabe es ist, Unternehmen danach zu klassifizieren, ob sie als insolvenzgefährdet oder als nicht insolvenzgefährdet einzustufen sind. Dieser Vergleich greift bewußt nur auf Jahresabschluß-Informationen zurück, wohl wissend, daß diese das jeweilige Unternehmensgeschehen nur unvollständig wiedergeben.²⁾ Klassifikationsziel ist das frühzeitige und sichere Erkennen einer krisenhaften Unternehmenssituation, welche sich letztlich am nachhaltigsten in einer Insolvenzgefährdung zeigt.

Dieser Beitrag greift damit eine Fragestellung auf, die in jüngsten Veröffentlichungen verstärkt beachtet wurde. Es wird nicht mehr nur untersucht, welche Jahresabschluß-Kennzahlen möglichst gut zwischen insolvenzgefährdeten und nicht insolvenzgefährdeten Unternehmen trennen, sondern vielmehr mit welchem Instrumentarium eine solche Klassifikation durchzuführen ist.³⁾ Das inhaltliche Analyseziel scheint dabei von einem rein methodischen verdrängt zu werden.

Folgt man dieser Richtung, die sicherlich nicht unumstritten ist,⁴⁾ wird erkennbar, daß sich in der Literatur anscheinend ein Wechsel von der Methode der Diskriminanzanalyse zur Methode des künstlichen neuronalen Netzes vollzogen hat. Ein statistisches Verfahren wird durch eine Technik der künstlichen Intelligenz abgelöst. Folgt man dagegen einschlägigen statistischen Lehrbüchern,⁵⁾ so wäre ein Wechsel zu einem anderen, möglicherweise leistungsfähigeren statistischen Verfahren der konsequentere Weg gewesen.⁶⁾ Insbesondere in Betracht käme die logistische Regressionsanalyse, bei der die Anforderungen an das Datenmaterial weniger restriktiv sind als bei der multivariaten Diskriminanzanalyse. Insofern erscheint es uns notwendig, nicht nur die Instrumente der Diskriminanzanalyse und des künstlichen neuronalen Netzes miteinander zu vergleichen, sondern auch noch die logistische Regressionsanalyse in den Vergleich mit einzubeziehen.

B. Instrumente empirischer Insolvenzdiagnose

I. Insolvenzdiagnose mittels Jahresabschluß-Kennzahlen

Die ausschließliche Verwendung von Daten des Jahresabschlusses - genauer Angaben der Bilanz und der Gewinn- und Verlustrechnung eines Unternehmens⁷⁾ - als Grundlage für eine Insolvenzdiagnose, ist in der einschlägigen Literatur bereits ausführlich problematisiert worden. Allerdings scheint selbst in jüngsten Veröffentlichungen die zum Teil sehr kontrovers geführte Debatte immer noch nicht abgeschlossen.⁸⁾ Insofern soll auch nicht der Eindruck erweckt werden, daß mit der im folgenden durchgeführten Analyse die kritischen Einwände, die gegen ein entsprechendes Forschungsdesign vorgetragen werden, gänzlich außer Acht gelassen werden.

(1) Ein erster Einwand zielt auf die Qualität der Datenbasis "Jahresabschluß":⁹⁾

- Die Informationen des Jahresabschlusses geben die unternehmerische Handlungsweise nur unvollständig wieder. Es fehlen z.B. Informationen über das Produktprogramm, die Stellung im Wettbewerb oder aber die Eigenschaft einer Organisation, flexibel auf Veränderungen reagieren zu können. Die Jahresabschluß-Informationen liefern insofern nur Anhaltspunkte für die Leistungsfähigkeit eines Unternehmens.
- Die Informationen des Jahresabschlusses sind vergangenheitsbezogen. Sie beziehen sich auf einen Stichtag bzw. einen Zeitraum der Vergangenheit. Folglich muß geprüft werden, ob die Prämissen, unter denen der Jahresabschluß in der Vergangenheit aufgestellt wurde, auch noch zum Zeitpunkt der Klassifikation gelten.
- Die Daten des Jahresabschlusses sind bewertungsabhängig. Mit den Bewertungsspielräumen werden bestimmte Zwecke verfolgt. Ein zu positiver, wie auch ein zu negativer Ausweis bleibt nicht ohne Folgen für die Qualität der Unternehmensklassifikation.¹⁰⁾
- Der vorgelegte Jahresabschluß ist nicht aktuell. Er liegt in der Regel frühestens sechs Monate nach Abschluß des Geschäftsjahres dem interessierten Analytiker vor. Aussagen, die sich auf kurzfristig veränderbare Jahresabschluß-Positionen beziehen, sind damit wertlos.

Trotz dieser Einwände an der Qualität der Datenbasis muß berücksichtigt werden, daß es sich bei einem Jahresabschluß um ein Rechenwerk handelt, daß nach standardisierten Vorschriften aufgestellt wird. Es gibt kaum eine andere Informationsbasis, die es in qualitativ ähnlicher Weise erlauben würde, Unternehmen

großzahlig miteinander zu vergleichen.

(2) Eine weitere kritische Frage betrifft das Abgrenzungskriterium zur Bestimmung derjenigen Teilmenge von Jahresabschluß-Informationen, die zur Unternehmensklassifikation herangezogen wird. Die Insolvenzforschung bedient sich zu diesem Zweck bestimmter Jahresabschluß-Kennzahlen, die mehr oder minder komplexe Kombinationen einzelner Jahresabschluß-Positionen darstellen. Ein Blick in entsprechende Forschungsarbeiten zeigt, daß sich ein harter Kern von Kennzahlen identifizieren läßt,¹¹⁾ die Auskunft über eine Insolvenzgefährdung eines Unternehmens geben.

Kritisch bei der Bestimmung dieser Kennzahlen ist, daß sie i.d.R. ausschließlich aufgrund statistischer Signifikanztests ermittelt werden. Mit ihnen kann die Insolvenz nicht kausal erklärt werden.¹²⁾ Für diese Kennzahlen kann nur der Anspruch erhoben werden, daß ihre jeweiligen Ausprägungen sich in der Vergangenheit jeweils signifikant bei insolventen und nicht insolventen Unternehmen unterscheiden haben.

(3) Darüber hinaus ist es aufgrund der oben aufgezeigten Mängel in der Datenbasis sicherlich auch nicht zulässig, den Jahresabschluß-Kennzahlen eine starke Prognosekraft zu attestieren. In der vorliegenden Untersuchung wird den jeweiligen Jahresabschluß-Kennzahlen lediglich eine vergangenheitsbezogene Indikatorfunktion zugebilligt. Es gilt diejenigen Kennzahlen zu identifizieren, die ex post gut zwischen insolvenzgefährdeten und unauffälligen Unternehmen getrennt haben.

Die vorliegende Untersuchung zielt damit nicht auf eine weitere theoretische Fundierung der Insolvenzforschung ab. Es wird sich vielmehr einem praktischen Problem zugewandt: der Frage, welche Instrumente sind besonders geeignet, kritische Jahresabschluß-Kennzahlen als vergangenheitsbezogene Indikatoren (Frühwarnsignale) für eine drohende Insolvenzgefahr zu ermitteln.

II. Instrumente zur Klassifikation von Unternehmen

Im Rahmen der traditionellen Bilanzanalyse finden sich in der Literatur eine Vielzahl von Kennzahlen oder Kennzahlensystemen, die eine Beurteilung bzw. Klassifikation von Unternehmen hinsichtlich bestimmter Analyseziele ermöglichen sollen.¹³⁾ Hierauf soll im Rahmen dieses Beitrags nicht im einzelnen eingegangen werden. Vielmehr soll sich Verfahren der Klassifikation von Unternehmen zugewandt werden, die sich multivariater statistischer Analysetechniken bedienen oder aber dem Bereich der

künstlichen Intelligenz zuordnen lassen. Konkret werden in diesem Betrag drei Klassifikationsinstrumente miteinander verglichen: die multivariate Diskriminanzanalyse (MDA), die logistische Regressionsanalyse (LR) und die Klassifikation mittels eines künstlichen neuronalen Netzes (KNN).

Für einen Vergleich der drei betrachteten Klassifikationsinstrumente wird folgende Problemstellung herangezogen: Es gilt insolvenzgefährdete Unternehmen von solchen zu trennen, die als unauffällig zu bezeichnen sind. Die Informationsbasis aufgrund derer diese Trennung zu erfolgen hat, ist a priori unbestimmt. Es werden allerdings in der Regel Kennzahlen sein, die aus unterschiedlichen Positionen des Jahresabschlusses gebildet werden. Diese vielfältigen Kennzahlen, mit ihren bei den zu klassifizierenden Unternehmen unterschiedlichen Ausprägungen, stellen die unabhängigen Variablen des Klassifikationsproblems dar. Die abhängige Größe ist das Klassifikationsergebnis, für welche ein dichotomes Skalenniveau, mit den Ausprägungen "0" unauffällig und "1" insolvenzgefährdet, unterstellt wird.

1. Grundlegende Vorgehensweise der multivariaten Diskriminanzanalyse

Ziel der multivariaten Diskriminanzanalyse ist die Trennung einer Gesamtheit in disjunkte Teilgesamtheiten.¹⁴⁾ Im Zwei-Gruppenfall, der im Rahmen dieser Untersuchung von Interesse ist, gilt es eine Linearkombination unabhängiger Merkmale - in diesem Fall Jahresabschluß-Kennzahlen - zu finden, die eine statistisch signifikante Trennung der Untersuchungseinheiten in unauffällige und insolvenzgefährdete vornimmt. Da es das Ziel der Diskriminanzanalyse ist, möglichst optimal zwischen den Gruppen zu trennen, wird als Diskriminanzkriterium der Quotient aus der Streuung zwischen den Gruppen und der Streuung innerhalb der Gruppen gewählt, den es zu maximieren gilt. Aufgrund der Linearkombination der Kennzahlen ergibt sich dann ein Kennwert, der eine Zuordnung der Fälle zu den einzelnen Gruppen erlaubt. Aufgabe der multivariaten Diskriminanzanalyse ist es somit:

- diejenigen Kennzahlen zu identifizieren, die sich in einem multivariaten Design als trennfähig erweisen,
- die Gewichte zu bestimmen, mit denen die als trennscharf identifizierten Kennzahlen in die Linearkombination eingehen und
- die Bestimmung eines kritischen Trennwertes, der die Einordnung in die einzelnen Gruppen festlegt.

Die folgenden statistischen Voraussetzungen müssen erfüllt sein, damit eine multivariate Diskriminanzanalyse zu einem "optimalen" Klassifikationsergebnis führt:

- Die Kennzahlen, die zur Trennung der beiden Gruppen herangezogen werden, müssen in der Grundgesamtheit normalverteilt sein.
- Es besteht Varianzhomogenität, d.h. die Varianzen der Kennzahlen beider Gruppen sind in der Grundgesamtheit gleich groß.
- Die Kennzahlen müssen unabhängig voneinander sein, d.h. es besteht keine Multikollinearität.

Diese Voraussetzungen sind in der statistischen Methodenliteratur unstrittig. Da jedoch bei der Anwendung der multivariaten Diskriminanzanalyse diese Voraussetzungen oftmals nicht erfüllt sind, ist in der betriebswirtschaftlichen Literatur eine Kontroverse darüber entbrannt, ob die Nichterfüllung der Annahmeveraussetzungen die Durchführung einer multivariaten Diskriminanzanalyse generell verbietet oder nur eine "optimale" Klassifikation verhindert.¹⁵⁾ Im Rahmen dieses Beitrages soll keine Wertung der Kontroverse erfolgen. Sie ist auch nicht notwendig, da mit der logistischen Regression auch ein statistisches Klassifikationsinstrument verwendet wird, das die Aufrechterhaltung der Normalverteilungsannahme und die Annahme homogener Varianzen nicht erfordert und somit einen Vergleich der Klassifikationsergebnisse ermöglicht.

2. Grundlegende Vorgehensweise der logistischen Regressionsanalyse

Grundlage der logistischen Regressionsanalyse¹⁶⁾ ist die Annahme, daß eine zufällige, abhängige Variable Y (die Insolvenzgefährdung) existiert, die die Werte 1 (insolvenzgefährdet) oder 0 (nicht insolvenzgefährdet) annehmen kann. Im Vordergrund steht hier die Frage, mit welcher Wahrscheinlichkeit die Variable Y den Wert 1 annimmt, d.h., ein Unternehmen als insolvenzgefährdet klassifiziert wird. Weiter wird angenommen, daß Y von k unabhängigen Variablen X_k mit $k = 1 \dots K$ (den Jahresabschluß-Kennzahlen) abhängt. Damit läßt sich eine bedingte Wahrscheinlichkeit bestimmen, die folgenden Ausdruck annimmt:

$$P = P(Y = 1 \mid X_1 \dots X_K).$$

Während man bei der linearen OLS-Regression davon ausgeht, daß der Zusammenhang zwischen Y und X linear ist, wird jetzt unterstellt, daß die bedingte Wahrscheinlichkeit $P(Y=1 \mid X)$ einen logistischen Funktionsverlauf in Bezug auf den linearen Prädiktor des logistischen Regressionsmodells annimmt. Formal läßt sich dieser Funktionsverlauf folgendermaßen ausdrücken:

$$P(Y=1 | X) = e^Z / (1+e^{-Z})$$

$$\text{mit: } z = b_0 + b_1 X_1 + \dots + b_K X_K \text{ für } k = 1 \dots K$$

Wie bei der OLS-Regression wird unterstellt, daß die Daten einer Zufallsstichprobe der Größe N entstammen und die einzelnen Beobachtungen von Y voneinander statistisch unabhängig sind. Schließlich dürfen auch keine Beziehungen zwischen den Jahresabschluß-Kennzahlen auftreten, d.h. Redundanz muß ausgeschlossen werden.

Zur Schätzung der Parameter wird das logistische Regressionsmodell in die Form des natürlichen Logarithmus überführt. Die zu schätzende Gleichung ist damit:

$$\ln [p/(1-p)] = z = b_0 + b_1 X_1 + \dots + b_K X_K \text{ für } k = 1 \dots K$$

Die Parameter des logistischen Regressionsmodells werden mit Hilfe des Maximum-Likelihood-Schätzverfahrens¹⁷⁾ bestimmt. Damit dieser Schätzer in seiner Leistungsfähigkeit nicht beeinträchtigt ist, wird ein relativ großer Stichprobenumfang vorausgesetzt.¹⁸⁾

3. Grundlegende Vorgehensweise eines künstlichen neuronalen Netzes

Im Gegensatz zu den statistischen Instrumenten der Unternehmensklassifikation, wie der multivariaten Diskriminanzanalyse oder der logistischen Regression, ist unter einem künstlichen neuronalen Netz nicht ein bestimmtes Instrument zu verstehen. Es gilt eine Vielzahl unterschiedlichster Erscheinungsformen zu unterscheiden.¹⁹⁾ Je nach Aufgabenstellung werden unterschiedliche künstliche neuronale Netze entwickelt. Sie bestehen jeweils aus einer Vielzahl netzartig miteinander verbundener und parallel arbeitender Informationsverarbeitungseinheiten, den sog. Neuronen. Die Neuronen nehmen Signale auf und multiplizieren sie mit einem Gewichtungsfaktor, um so das Eingangssignal zu berechnen, welches den Aktivitätszustand des Neurons festlegt. Anschließend gibt das Neuron ein Signal ab. Dieses Signal wird über eine Outputfunktion berechnet, die linear, sigmoid oder ein Schwellwert sein kann. Die Neuronen werden je nach Art der Verbindung der Neuronen zu unterschiedlichen Netztypen zusammengefaßt. Man unterscheidet dabei Schichten- (feed forward) von Klumpenmodellen (feed back). Im Gegensatz zum Schichtenmodell sind beim Klumpenmodell auch Rückbeziehungen zwischen den Neuronen zugelassen. Für das hier behandelte Klassifikationsproblem ist die Aufnahme von Rückbeziehungen nicht sinnvoll, d.h. die Neuronen werden hier in einem Schichtenmodell gruppiert.

Ein solches Schichtenmodell besteht in der Regel aus einer Eingabeschicht, einer Ausgabeschicht und einer versteckten Schicht. Für den Fall, daß die versteckte Schicht zwei oder mehr Neuronen umfaßt, ist man in der Lage, jeden stetigen Zusammenhang zwischen einem Eingabeneuron und einem Ausgabeneuron abzubilden.

Durch die wiederholte Präsentation von Informationsmustern über die Eingangsneuronen, ist das künstliche neuronale Netz in der Lage, wichtige von unwichtigen Informationen im Hinblick auf den Output zu trennen. Je nach Art des Lernvorgangs spricht man von einem überwachten Lernen oder von einem assoziativen Lernen. Künstliche neuronale Netze neigen, wegen ihrer Eigenschaft, nicht lineare Funktionen beliebiger Art approximieren zu können, stärker als lineare statistische Verfahren dazu, in den Trainingsdaten nicht nur die wirklichen Abhängigkeiten, sondern auch jede Art von Scheinzusammenhang ("Rauschen") zu erkennen. Anhand einer "Kontrollgruppe" - die mit der Holdout-Stichprobe nicht identisch ist - versucht man das "Overfitting" zu kontrollieren. Der Lernvorgang wird abgebrochen, wenn bei der Kontrollgruppe der Zuordnungsfehler zu steigen beginnt, das Netz also anfängt "Unsinn" zu lernen.²⁰⁾

4. Vergleichsuntersuchungen zum Instrumenteneinsatz

Wird ein Vergleich von Klassifikationsinstrumenten durchgeführt, so interessiert in erster Linie die Güte des Klassifikationsergebnisses. Dies gilt unabhängig vom angewandten Klassifikationsverfahren. Die Güte der Klassifikation findet ihren Ausdruck im Prozentsatz der falsch klassifizierten Unternehmen. Dabei sind zwei Fehler voneinander zu unterscheiden: Zum einen kann der Fall auftreten, daß ein insolvenzgefährdetes Unternehmen als unauffällig eingestuft wird. Dieser Fehler wird in der Literatur allgemein als Alpha-Fehler bezeichnet. Zum anderen kann sich die Fehlbeurteilung auf ein unauffälliges Unternehmen beziehen, welches fälschlich als insolvenzgefährdet klassifiziert wird - hier spricht man vom Beta-Fehler. Beide Fehler sind unterschiedlich in ihrer Konsequenz für den Beurteilenden. Findet die Beurteilung aus Sicht eines Kreditinstitutes statt, so führt der Alpha-Fehler zu Forderungsausfällen und der Beta-Fehler dazu, daß mögliche Gewinne entgehen. Da beide Fehler nicht unabhängig voneinander sind, müssen geschäftspolitische Entscheidungen darüber getroffen werden, welche Größenordnungen für beide Fehlerarten als akzeptabel zu bezeichnen sind.²¹⁾

In der Literatur finden sich bisher nur vereinzelte Untersuchungen, die Instrumente zur

Unternehmensklassifikation hinsichtlich ihrer Leistungsfähigkeit miteinander vergleichen. Sie sollen an dieser Stelle kurz dargestellt werden:

(1) Ein erster Vergleich der Leistungsfähigkeit der multivariaten Diskriminanzanalyse und der Klassifikation mit Hilfe eines künstlichen neuronalen Netzes wurde von Odom und Sharda²²⁾ durchgeführt. Sie analysieren die Jahresabschlüsse von 65 insolventen und 64 vergleichbaren nicht insolventen Unternehmen. Dabei zeigte sich, daß der Einsatz des künstlichen neuronalen Netzes im Vergleich zur multivariaten Diskriminanzanalyse zu besseren Ergebnissen führte. So lagen der Alpha- und Beta-Fehler bei der Holdout-Stichprobe für die multivariate Diskriminanzanalyse bei 15,7 % und 30,6 %, beim künstlichen neuronalen Netz dagegen lediglich bei 18,5 % und 17,9 %. Die Überlegenheit des künstlichen neuronalen Netzes wird von den Autoren darauf zurückgeführt, daß der mehrlagige Netzaufbau im Gegensatz zur multivariaten Diskriminanzanalyse erlaubt, auch nicht lineare Beziehungen abzubilden. Kritisch muß jedoch bei dieser Untersuchung festgestellt werden, daß Odom und Sharda den beiden Analyseinstrumenten zur Klassifikation fünf Kennzahlen vorgegeben haben. Die Verfahren waren somit nicht in der Lage, ihre Leistungsfähigkeit im Bereich der Kennzahlenselektion unter Beweis zu stellen. Darüber hinaus wurden die Jahresabschlüsse des Jahres t-1 in Bezug auf die Insolvenz betrachtet, ein sicherlich höchst problematischer Zeitpunkt.

(2) In einer weiteren Untersuchung haben Tam und Kiang²³⁾ neben der multivariaten Diskriminanzanalyse und dem künstlichen neuronalen Netz auch noch die logistische Regressionsanalyse in einen Vergleich einbezogen. Allerdings wurden bei der logistischen Regressionsanalyse nicht die Jahresabschluß-Kennzahlen als Prädiktoren gewählt, sondern Faktorwerte. Diese Faktorwerte ergaben sich aus einer Faktorenanalyse über 19 a priori festgelegte Kennzahlen - eine theoretisch höchst strittige Vorgehensweise.²⁴⁾ Für das Jahr t-2 in Bezug auf den Insolvenzzeitpunkt ergaben sich folgende Ergebnisse für die Holdout-Stichprobe:

- Die multivariate Diskriminanzanalyse führte bei 8 Kennzahlen zu einem Alpha-Fehler von 35 % (Beta-Fehler 0 %).
- Die logistische Regressionsanalyse führte bei 9 Kennzahlen zu einem Alpha-Fehler von 30 % (Beta-Fehler 0 %).
- Das künstliche neuronale Netz führte bei 19 Kennzahlen zu einem Alpha-Fehler von 2,5 % (Beta-Fehler 20 %).

Auch hier schneidet das künstliche neuronale Netz besser ab. Es muß jedoch berücksichtigt werden, daß Tam und Kiang nur eine sehr kleine Holdout-Stichprobe zur Validierung der Klassifikation herangezogen haben. Darüber hinaus konnten die

statistischen Verfahren ihre Selektionsleistung nicht ausspielen, da ihnen die 19 Kennzahlen des künstlichen neuronalen Netzes vorgegeben wurden.

(3) Ein weiterer Vergleich wurde von Rehkugler und Poddig²⁵⁾ vorgenommen. In ihrer Untersuchung wurden 900 Jahresabschlüsse (drei aufeinander folgender Jahre) von 150 solventen und 150 insolvente Unternehmen mit Hilfe von 45 Kennzahlen analysiert. Die Autoren entwickeln unterschiedliche mehrlagige künstliche neuronale Netze. Dabei scheint sich mit der Anzahl der Eingabeneuronen auch die Klassifikationsgüte zu erhöhen. Im besten Fall (künstliches neuronales Netz mit 11 Kennzahlen) wird eine Gesamttrefferquote von 89,0 % erreicht. Im Vergleich hierzu kommt die multivariate Diskriminanzanalyse (drei Kennzahlen) auf 85,7%. Auch diese Untersuchung zeigt wieder die Überlegenheit des künstlichen neuronalen Netzes. Leider wird bei Rehkugler und Poddig nur die Gesamtklassifikationsgüte der Einzelinstrumente ausgewiesen und nicht der Alpha- und Beta-Fehler. Nichtsdestoweniger scheint sich auch hier die Möglichkeit der Berücksichtigung nicht-linearer Zusammenhänge positiv auf die Klassifikationsgüte auszuwirken, wobei die Autoren jedoch einschränkend darauf hinweisen, daß "durch die Nutzung von Nichtlinearitäten kein Quantensprung bezüglich der Klassifikationsgüte zu erzielen ist."²⁶⁾

(4) Schließlich sei noch auf den Vergleich der multivariaten Diskriminanzanalyse und des künstlichen neuronalen Netzes in der Untersuchung von Erxleben et al.²⁷⁾ eingegangen. Sie macht sehr gut deutlich, daß die Leistungsfähigkeit eines künstlichen neuronalen Netzes - selbst wenn es vom Aufbau her in der Lage ist, nicht-lineare Zusammenhänge abzubilden - entscheidend von der Leistungsfähigkeit der bereitgestellten Hardware abhängt. Zwar konnten die Autoren auf einen sehr großen Datenbestand zurückgreifen (3539 Jahresabschlüsse), die Tatsache jedoch, daß ihnen lediglich ein PC mit einem Intel-80386-Prozessor zur Verfügung stand, führte dazu, daß dem künstlichen neuronalen Netz diejenigen Kennzahlen vorgegeben wurden, die sich im Rahmen einer multivariaten Diskriminanzanalyse als trennscharf erwiesen haben. Das künstliche neuronale Netz wies entsprechend kaum bessere Klassifikationsgüten auf als die multivariate Diskriminanzanalyse. Bei einem konstanten Beta-Fehler von 20,84 % im Zeitpunkt t-3 erreichte das künstliche neuronale Netz einen Alpha-Fehler von 33,3 % und die multivariate Diskriminanzanalyse einen Alpha-Fehler von 33,73 %.

III. Fazit: Qualitätsindikatoren für die Güte der einzelnen Instrumente

Sicherlich unstrittig ist, daß die Klassifikationsgüte des jeweiligen Instrumentes als Erfolgskriterium anzusehen ist. An der Höhe des Alpha- und Beta-Fehlers läßt sich ablesen, wie gut es den einzelnen Instrumenten gelingt, insolvente Unternehmen auch als solche zu klassifizieren, bzw. umgekehrt, unauffällige Unternehmen entsprechend zu erkennen. Diesem Erfolgskriterium wird auch in der Praxis die größte Bedeutung zugemessen. Für einen Vergleich erscheint es allerdings angebracht, weitere Qualitätsindikatoren heranzuziehen.

Um die Akzeptanz eines Klassifikationsergebnisses beim Anwender zu gewährleisten, ist es in der Regel notwendig, daß das Ergebnis vom Systementwickler intuitiv nachvollzogen werden kann. Die in die Klassifikation eingegangenen Kennzahlen sollten nicht im Widerspruch zu klassischen Kennzahlinterpretationen stehen. Dabei ist vielfach auch von Interesse, wie sensitiv das Klassifikationsergebnis auf etwaige Änderungen der Konstellation der unabhängigen Variablen reagiert. Zudem sollte die Konstruktion der identifizierten trennscharfen Kennzahlen und deren Verknüpfung plausibel sein.

Schließlich müssen auch die Anwendungsvoraussetzungen der Klassifikationsinstrumente miteinander verglichen werden. Die Voraussetzungen beziehen sich dabei auf unterschiedliche Aspekte:

- Zum einen die Frage, welchen Anforderungen die Analysedaten genügen müssen.
- Zum anderen das Problem des Vorhandenseins der notwendigen Kapazitäten - zeitlich, materiell, personell - für die Anwendung des jeweiligen Klassifikationsinstruments.

C. Empirisches Design der Vergleichsstudie

I. Untersuchungsstichprobe

Für die vorliegende Untersuchung stand eine Stichprobe zur Verfügung, deren Daten nicht selbst erhoben wurden. Sie wurde freundlicherweise von einem großen deutschen Kreditinstitut zur Verfügung gestellt. Somit war es möglich, die einzelnen Klassifikationsinstrumente anhand einer repräsentativen Stichprobe von Firmenkunden, also einem realitätsnahen Anwendungsfall, zu testen. Gleichzeitig blieb den Autoren die Eingabe von knapp 15000 Einzelinformationen, die notwendig sind, um die

ausgewählten Kennzahlen zu bilden, erspart.

Die Unternehmen der Stichprobe entsprechen dem typischen Kreditengagement einer Großbank im Firmenkundengeschäft:

- Das Einzelkreditvolumen liegt über DM 100.000, so daß es nach § 18 KWG für den Kreditnehmer notwendig war, seine wirtschaftlichen Verhältnisse offenzulegen. Es wurden keine Kreditnehmer betrachtet, bei denen es zu einer Vergabe von Großkrediten im Sinne des § 13 Abs. 1 KWG kam.
- Es handelte sich bei den einbezogenen Unternehmen überwiegend um mittelständische Unternehmen (Umsatz zwischen DM 8 Mio. und DM 32 Mio.) des verarbeitenden Gewerbes.

Die Operationalisierung der Unternehmensinsolvenz orientierte sich nicht nur an der Antragstellung auf Konkurs oder Vergleich. Von einer Insolvenz wird bereits dann gesprochen, wenn eine der nachfolgend genannten "erheblichen Leistungsstörungen" während der Kreditlaufzeit aufgetreten ist:

- Einberufung eines außergerichtlichen Moratoriums (Stundungsvergleich).
- Haftandrohung zur Abgabe einer eidesstattlichen Versicherung.
- Einleitung einer Sicherheitenverwertung.
- Einzelwertberichtigung der Forderung.

Aus Sicht der Bank ist damit eine solche Leistungsstörung immer gleichzusetzen mit der Übernahme des Kreditengagements auf ein sogenanntes Abwicklerkonto. Die Wahl der Leistungsstörung während der Kreditlaufzeit führt - im Vergleich zum tatsächlichen Insolvenzzeitpunkt - zu einer Vorverlagerung der zu diagnostizierenden Unternehmenskrise.

Schließlich gilt es auch noch den Zeitpunkt der Untersuchung festzulegen, da dieser, wie empirische Untersuchungen gezeigt haben,²⁸⁾ nicht ohne Einfluß auf das Klassifikationsergebnis ist. Als Zeitraum für die Insolvenzklassifikation werden die Jahre 1992 bis 1994 gewählt. Von denjenigen Unternehmen, die in diesem Zeitraum das Insolvenzkriterium erfüllen, werden die Jahresabschlüsse drei Jahre vor dem Auftreten der oben genannten Leistungsstörungen als Informationsgrundlage für die vergleichende Unternehmensklassifikation herangezogen. Dieser Zeitpunkt erweist sich als zweckmäßig, da er bei erfolgreicher Diagnose noch dazu führt, daß die identifizierten kritischen Jahresabschluß-Kennzahlen die Funktion eines Frühindikators besitzen. Es besteht ausreichende Reaktionszeit für eine "sich-selbst-zerstörende Prophezeiung".

Entsprechend dieser Vorgehensweise wurden 116 insolvente Unternehmen in die Stichprobe aufgenommen. Ergänzend wurden dann noch 116 Vergleichsfälle ausgewählt, die das Insolvenz Kriterium im Untersuchungszeitraum nicht erfüllen. Diese Auswahl erfolgte wiederum durch das Kreditinstitut, wobei als Vergleichskriterien die Merkmale

- Jahr des Jahresabschlusses,
- Umsatzvolumen,
- Rechtsform und
- Branchenzugehörigkeit

gewählt wurden.

Eine derartige Auswahl ist sicherlich problematisch. Neben der Frage, ob diese vier Kriterien genügen, eine Vergleichbarkeit zu sichern, ist insbesondere das Problem der Operationalisierung der "Nicht-Insolvenz" von Bedeutung. Die Klassifikation als unauffällig besagt nämlich nicht, daß es sich bei diesen Unternehmen auch wirklich um solvente, d.h. "gute" Unternehmen handelt. Die unauffälligen Unternehmen sind durchschnittliche Unternehmen. Unter ihnen verbergen sich natürlich auch die insolventen Unternehmen der Zukunft, also solche Unternehmen bei denen sich die Insolvenz Kriterien noch nicht gezeigt haben, diese jedoch in vier, fünf Jahren oder später auftreten. Folglich wird auch die Klassifikationsgüte der einzelnen Instrumente nicht optimal sein, da es nicht zur Gegenüberstellung ausschließlich "guter" Unternehmen mit "schlechten" Unternehmen kommt.

Die Gesamtstichprobe umfaßt somit die Jahresabschlußdaten des Jahres $t-3$ in Bezug auf den Insolvenzzeitpunkt von 116 Unternehmungspaaren. Diese Stichprobe wurde dann zufällig in zwei jeweils gleich große Teilstichproben unterteilt. Die sog. Analysestichprobe umfaßt damit 58 Unternehmenspaare. Mit ihrer Hilfe werden die statistischen Klassifikations-Modelle entwickelt. Die verbliebenen 58 Unternehmenspaare bilden die sog. Holdout-Stichprobe. Sie dient der Prüfung der Klassifikationsgüte der einzelnen Modelle.

II. Klassifikationsinstrumente

Die multivariaten statistischen Klassifikationsmodelle wurden mit Hilfe der Standard-Software SPSS entwickelt. Dabei wurden jeweils die Voraussetzungen geprüft, die von den einzelnen Verfahren hinsichtlich des Datenmaterials gefordert werden. Es zeigte sich insbesondere, daß einige Jahresabschluß-Kennzahlen nicht normalverteilt waren und auch die Varianzhomogenitäts-Annahme nicht gesichert war. Schließlich führten

Korrelationen bzw. Redundanzen zwischen der Jahresabschluß-Kennzahlen dazu, daß einige Kennzahlen zur Modellbildung nicht herangezogen werden konnten. Die Güte des multivariaten Diskriminanzmodells wie auch des logistischen Regessionsmodells wurde mit den einschlägigen statistischen Gütekriterien überprüft. Die Grundlage für die Entwicklung des multivariaten Diskriminanzmodells und des logistischen Regessionsmodells bilden 31 Jahresabschluß-Kennzahlen. Die Auswahl der Kennzahlen orientiert sich an vergleichbaren Studien der Insolvenzforschung.²⁹⁾

In der vorliegenden Untersuchung wurde ein künstliches neuronales Netz verwandt, das dem sogenannten "Backpropagation-Algorithmus" folgt. Dieses Netz wurde unter Leitung von Baetge am Institut für Revisionswesen der Universität Münster entwickelt und wird in den jeweiligen Veröffentlichungen als "BP 14" bezeichnet.³⁰⁾

Da das Netz "BP14" nicht auf der Basis unserer Stichprobe entwickelt worden ist, erscheint es notwendig, kurz die von Baetge gewählte Vorgehensweise für die Netzentwicklung darzulegen:

Als Lernstichprobe wurden die Jahresabschlüsse dreier aufeinander folgender Jahre von 131 insolvente Unternehmen und 131 unauffällige Vergleichsunternehmen gewählt. Das Jahr der Insolvenz ist der zeitliche Anker der Stichprobe. Die drei Jahresabschlüsse, die vor diesem Jahr lagen, bilden die Informationsgrundlage für die insolventen und die vergleichbaren unauffälligen Unternehmen. Im Gegensatz zu den Stichproben für die Diskriminanzanalyse und logistische Regression werden die Kennzahlen nicht danach unterschieden, aus welchem Jahr sie stammen. Das künstliche neuronale Netz sieht jeden Jahresabschluß als neue Fall an, auch wenn dieser von demselben Unternehmen, nur aus einem anderen Jahr stammt. Für jede Untersuchungseinheit werden 209 Jahresabschluß-Kennzahlen ermittelt.³¹⁾ Sie bilden die Informationsmuster anhand derer das Netz trainiert wird.

Die Auswahl der "trennscharfen" Kennzahlen erfolgt über sogenannte "Pruning-Algorithmen".³²⁾ Es wurden zwei Algorithmen angewandt: ein gewichtsorientiertes Pruning und ein relevanzorientiertes Pruning.

Beim gewichtsorientierten Pruning werden diejenigen Verbindungen mit den niedrigsten Verbindungsgewichten abgeschnitten. Sobald bei einem Eingabeneuron sämtliche Verbindungen zu der versteckten Schicht eliminiert sind, wird die Kennzahl selbst eliminiert. Dieses gewichtsorientierte Pruning wird solange fortgeführt, bis sich der Beta-Fehler erhöht. Im Ergebnis konnte durch ein gewichtsorientiertes Pruning die Anzahl der Eingabeneuronen von 209 Kennzahlen auf 50 Kennzahlen verringert

werden.

Diese 50 Kennzahlen bilden den Ausgangspunkt für das relevanzorientierte Pruning. Dabei werden nicht mehr einzelne Verbindungen eliminiert, sondern einzelne Eingabeneuronen. Das Relevanzkriterium ist dabei die Stärke des Einflusses auf den Beta-Fehler bei einem fixierten Alpha-Fehler. Die Relevanz eines Neurons ist damit die Beta-Fehler-Differenz zwischen einem Netz ohne dieses Neuron im Vergleich zu einem Netz welches noch alle Neuronen zu Beginn des relevanzorientierten Prunings umfaßt. Ist diese Differenz negativ, d.h der Beta-Fehler ohne das Neuron geringer als mit dem Neuron, so wird das entsprechende Neuron abgeschnitten. Mit Hilfe dieses Pruning-Algorithmus gelang eine Verkleinerung der Zahl der Eingabeneuronen auf 14.³³⁾

Das so optimierte künstliche neuronale Netz zur Unternehmensklassifikation hat damit die folgende Struktur:

Abb. 1: Struktur des KNN zur Unternehmensklassifikation

- bitte hier einfügen -

Im Unterschied zu den statistischen Verfahren umfaßt das künstliche neuronale Netz auch redundante Kennzahlen. So unterscheiden sich beispielsweise die Rentabilitätskennzahlen K2 und K3 nur durch die Zählervariation um den Posten "Zuführung zu den Pensionsrückstellungen". In der vorangegangenen Abbildung ist die Zusammenfassung der Einzelkennzahlen zu Kennzahlenblöcken eine Konsequenz ihrer Redundanz.

D. Untersuchungsergebnisse

I. Vergleich der Klassifikationsleistung

Der zentrale Qualitätsindikator für die Leistungsfähigkeit der einzelnen Klassifikationsmodelle ist die Höhe der Trefferquote, also der Anteil von Hundert derjenigen Fälle, die richtig als "insolvenzgefährdet" bzw. "nicht insolvenzgefährdet" eingestuft wurden. Die folgende Tabelle zeigt die entsprechenden Ergebnisse:

Tab. 1: Trefferquoten der Klassifikationsmodelle bei der Holdout-Stichprobe

- bitte hier einfügen -

Generell fällt auf, daß alle drei Modelle eine gute Trefferquote erreichen. Von der

Literatur wird allgemein ein Gesamtfehler von kleiner als 25% als gut bezeichnet. Wendet man sich dem Vergleich der einzelnen Instrumente zu, so zeigen sich folgende Unterschiede:

Es besticht sicherlich die überaus gute Trefferquote der logistischen Regression. Insbesondere den niedrigen Alpha-Fehler bei gleichzeitig niedrigem Beta-Fehler gilt es hervorzuheben. Die logistische Regression erzielt damit nicht nur eine Klassifikationsgüte, die im Rahmen der vorliegenden Untersuchung beachtlich ist, sondern auch im Vergleich zu anderen Untersuchungen - die wie der vorliegende Datensatz - Jahresabschluß-Daten des Zeitpunktes "drei Jahre vor der Insolvenz" betrachten. Ähnlich gute Ergebnisse berichten nur Kerling und Poddig, die für ihre besten künstlichen neuronalen Netze eine Gesamttrefferquote von 85,0 % bzw. 84,0% ausweisen.³⁴⁾ Die logistische Regressionsanalyse scheint damit einen wirklichen Beitrag zur Verbesserung der Klassifikationsgüte zu leisten.

Der Unterschied zwischen den beiden anderen Instrumenten (MDA und KNN) fällt demgegenüber nicht stark ins Gewicht. Der Unterschied im Alpha-Fehler beträgt 2,6 % und im Beta-Fehler 4,3 % jeweils zugunsten der multivariaten Diskriminanzanalyse. Allerdings muß bei dem Vergleich der Ergebnisse des künstlichen neuronalen Netzes mit den statistischen Verfahren berücksichtigt werden, daß die Analysestichproben, anhand derer die Modellcharakteristika spezifiziert wurden, unterschiedlich waren.

Die multivariaten statistischen Klassifikationsmodelle wurden anhand einer Stichprobe entwickelt, deren Elemente sämtlich Kunden eines bestimmten Kreditinstitutes waren.³⁵⁾ Insofern ist der Verdacht nicht unbegründet, daß die einzelnen Klassifikationsmodelle mit einem gewissen "Institutsbias" behaftet sind. Dies hat zur Folge, daß das künstliche neuronale Netz benachteiligt ist, da es nicht in der Lage war, diesen "Institutsbias"³⁶⁾ zu erlernen. Umso erstaunlicher ist die vergleichsweise sehr gute Trefferquote des künstlichen neuronalen Netzes zu beurteilen.³⁷⁾ Sie läßt darauf schließen, daß sich das "BP 14" durch ein sehr gute Allgemeingültigkeit auszeichnet.

Vergleicht man die beiden statistischen Verfahren, so scheinen die weniger restriktiven Anforderungen an das Datenmaterial und die Möglichkeit der Annahme eines logistischen Funktionsverlaufs zwischen der Insolvenzwahrscheinlichkeit und dem linearen Prädiktor der Jahresabschluß-Kennzahlen der Grund dafür zu sein, daß die logistische Regressionsanalyse zu deutlich besseren Ergebnissen führt als die multivariate Diskriminanzanalyse.

II. Vergleich der Ergebnisakzeptanz

Die Akzeptanz der Ergebnisse beim Anwender eines Klassifikationsmodells wird in erste Linie von der Plausibilität der als trennscharf identifizierten Jahresabschluß-Kennzahlen und der Möglichkeit ihrer Interpretation im Hinblick auf die Insolvenzgefährdung bestimmt. Die folgende Tabelle gibt diejenigen Kennzahlen wieder, die als trennscharf von den einzelnen Klassifikationsmodelle identifiziert worden sind.

Tab. 2: Trennscharfe "Jahresabschluß-Kennzahlen" der Klassifikationsmodelle

- bitte hier einfügen -

Ein Vergleich der Kennzahlen der logistischen Regression mit denen der multivariaten Diskriminanzanalyse zeigt, daß diese im wesentlichen identisch sind. Ein Vergleich mit den Kennzahlen, die im Rahmen der Klassifikation durch das künstliche neuronale Netz herangezogen werden, zeigt, daß dort Kennzahlen, die ähnlich dem "Anteil FK-Zinsen" oder dem "Anteil Vorräte" sind, keine Berücksichtigung finden. Dafür werden eine Vielzahl von Kennzahlen berücksichtigt, die durch eine Zähler- oder Nennervariation einer anderen Kennzahl entwickelt worden sind. Wir haben versucht, dies durch die entsprechende Nummerierung (I,II,III) deutlich zu machen.

Für jede identifizierte Jahresabschluß-Kennzahl lassen sich Plausibilitätsüberlegungen hinsichtlich ihrer Ausprägung bei insolventen und nicht insolventen Unternehmen anstellen. Insofern kann man unterstellen, daß es sich um allgemein akzeptierte Kennzahlen handelt. Für die statistischen Verfahren sind wir darüber hinaus auch in der Lage, die Richtung und das relative Gewicht der Kennzahlen zu beurteilen. Die standardisierte Diskriminanzfunktion weist folgende Koeffizienten auf:³⁸⁾

$$Z_{\text{Stand.}} = -0,602 \cdot K1 + 0,899 \cdot K2 + 0,71 \cdot K3 + 0,61 \cdot K4 - 0,527 \cdot K5$$

Für das logistische Regressionsmodell lassen sich die folgenden partiellen Regressionskoeffizienten berechnen:

$$K1: r = -0,210$$

$$K2: r = +0,238$$

$$K3: r = +0,251$$

$$K4: r = +0,215$$

$$K5: r = -0,110$$

Die Vorzeichen der jeweils identifizierten trennscharfen Kennzahlen entsprechen den Plausibilitätsüberlegungen. Die relativ größte Trennschärfe besitzen bei beiden Instrumente die Jahresabschluß-Kennzahlen "Anteil FK-Zinsen" und "Anteil Vorräte". Es ist relativ einfach, die Ergebnisse im Hinblick auf die Insolvenzgefährdung zu interpretieren. Dies gilt insbesondere im Vergleich zu den Ergebnissen des künstlichen neuronalen Netzes. Hier ist nur eine Plausibilitätsprüfung der Kennzahlen möglich, d.h. des Umstandes, daß die Kennzahlenausprägung bei insolventen und nicht insolventen Unternehmen unterschiedlich ist. Die Richtung sowie die relative Stärke des Einflusses der einzelnen Kennzahlen sind dem Anwender eines solchen Klassifikationsmodells nicht bekannt.³⁹⁾ Folglich wird er die Klassifikationsergebnisse - zumindest dann, wenn er ihrer Interpretierbarkeit erhöhtes Gewicht beimißt - nicht in dem Maße akzeptieren, wie diejenigen der multivariaten statistischen Klassifikationsverfahren.

III. Vergleich der Anforderungsvoraussetzungen

Einen weiteren, wichtigen Aspekt stellen die unterschiedlichen Anforderungen an das zu klassifizierende Datenmaterial dar. Hier besitzt das künstliche neuronale Netz die größte Leistungsfähigkeit. Zum einen ist es in der Lage, eine große Anzahl von unterschiedlichen Informationen zu verarbeiten. Zum anderen stellt es die geringsten Anforderungen an das Datenmaterial. Dies wird insbesondere dann von Bedeutung sein, wenn neben den Kennzahlen des Jahresabschlusses auch andere Daten zur Klassifikation herangezogen werden, die kein kardinales Skalenniveau aufweisen. Demgegenüber weist die multivariate Diskriminanzanalyse mit ihrer Forderung nach Einhaltung der Normalverteilungs-Annahme und der Varianzhomogenitäts-Annahme sicherlich die höchsten Anforderungsvoraussetzungen auf. Die logistische Regressionsanalyse stellt hier lediglich die Forderung nach einer nicht gegebenen Redundanz der Kennzahlen und Autokorrelation. Sie ist darüber hinaus auch in der Lage, qualitative Urteile als Informationsgrundlage zu akzeptieren, soweit diese binäres Skalenniveau aufweisen.

Beurteilt man den Einsatz der einzelnen Instrumente abschließend noch vor dem Hintergrund der zur Verfügung stehenden Kapazität (zeitlich, personell, materiell), so ist das künstliche neuronale Netz augenblicklich sicherlich noch als das aufwendigste Instrument zu bezeichnen. Es ist jedoch anzunehmen, daß in der Zukunft dieser Nachteil zunehmend an Bedeutung verliert. In materieller und zeitlicher Hinsicht, sind mit der Entwicklung leistungsfähigerer Prozessoren und anwendungsfreundlicherer Software zur Entwicklung künstlicher neuronaler Netze in jüngster Zeit schon

erhebliche Fortschritte erzielt worden. Relativ aufwendig ist augenblicklich noch der Aufbau des künstlichen neuronalen Netzes. Hier gibt es eine Vielzahl unterschiedlicher Netztypen und Lernverfahren, ohne daß bereits gesicherte Erkenntnisse darüber vorliegen, welche Netzkonfiguration für welche Problemstellung am besten geeignet erscheint.⁴⁰⁾ Es ist anzunehmen, daß dieser Nachteil, ähnlich wie bei den statistischen Verfahren, mit der Zeit an Gewicht verlieren wird. Demgegenüber lassen sich die multivariaten statistischen Klassifikationsverfahren relativ einfach durchführen. Es stehen eine Vielzahl leistungsfähiger und einfach zu bedienender Statistik-Software-Produkte zur Verfügung.

E. Schlußbemerkungen

Die vorangegangenen Ausführungen haben gezeigt, daß die logistischen Regressionsanalyse sehr gut geeignet erscheint, Unternehmen aufgrund von Jahresabschluß-Kennzahlen zu klassifizieren. Allerdings muß auch hierbei berücksichtigt werden, daß die vorliegende Untersuchung lediglich Jahresabschluß-Kennzahlen als Informationsbasis heranzieht. Unabhängig von der Frage, ob diese Kennzahlen hinreichend gut geeignet erscheinen, eine entsprechende Klassifikation vorzunehmen, wäre es durchaus möglich, qualitative Daten ergänzend mit in die Informationsbasis aufzunehmen. Zu denken wäre hier beispielsweise an Variable, die die Zahlungsgepflogenheiten oder die Qualität des Managements abbilden. Derartige Informationen können im Rahmen eines logistischen Regressionsmodells nur noch dann verarbeitet werden, wenn sie kardinales oder binäres Skalenniveau aufweisen. Die Komplexität der Datenstrukturen ist jedoch auch von der logistischen Regressionsanalyse nur in beschränktem Maße zu verarbeiten. In diesem Bereich liegt eindeutig die Stärke der künstlichen neuronalen Netze. Erste Entwicklungen versuchen hier insbesondere qualitative Daten in die Klassifikation einzubeziehen.⁴¹⁾ Ein weiterer Entwicklungspfad geht in Richtung der Mustererkennung.

Eine Möglichkeit, wie dies durchzuführen ist, stellen aktuell Oestreicher und Piotrowski-Allert⁴²⁾ vor, die nicht nur Jahresabschluß-Kennzahlen sondern auch Kennzahlen zur Bilanzpolitik und zur Lageberichterstattung berücksichtigen. Sie entwickeln eine sog. "Kohonen-Feature-Map" zur Entdeckung von Ähnlichkeiten im Jahresabschluß unterschiedlicher Unternehmen. Es handelt sich dabei um ein selbstorganisierendes künstliches neuronales Netz, d.h. im Gegensatz zum oben angewandten Backpropagation-Netz werden diesem Netz keine Informationen über den zu ermittelnden Zielwert - beispielsweise das Eintreten einer Insolvenz - vorgegeben.

Die sich abzeichnende Tendenz zur Aufnahme von Informationen unterschiedlicher Qualität muß jedoch berücksichtigen, daß der Grenzertrag der Aufnahme einer zusätzlichen Information in das Klassifikationsmodell eines künstlichen neuronalen Netzes mit wachsender Informationsgrundlage höchstwahrscheinlich abnimmt. Es scheint hier angeraten zu prüfen, ob ein vergleichsweise einfach zu entwickelndes logistisches Regressionsmodell noch nachhaltig in der Klassifikationsgüte übertroffen werden kann.

Anmerkungen

- 1) Vgl. hierzu Bieg (1984) S. 498 f., Lachnit (1979) S. 1 f. und Rautenberg (1990) S. 258.
- 2) Vgl. hierzu ausführlich Schneider (1985), Gemünden (1988) und Burger (1994).
- 3) Vgl. z.B. Erxleben et al. (1992), Burger (1994), Baetge/Krause/Mertens (1994), Kumar/Rao/Soni (1995) und Markham/Ragsdale (1995).
- 4) Vgl. z.B. Schneider (1985) und Schneider (1989).
- 5) Vgl. z.B. Hosmer/Lemeshow (1989) S. 34.
- 6) Ein Grund könnte in den wenig überzeugenden Ergebnissen der ersten vergleichenden Anwendungen von verteilungsfreien Verfahren mit der multivariaten Diskriminanzanalyse liegen. Vgl. hierzu Gebhardt (1980) und Niehaus (1987).
- 7) In der Kommentarliteratur wird hierfür der Begriff "Zahlenwerk" gewählt. Vgl. Adler/Düring/Schmaltz (1987) § 264 HGB, Tz. 87.
- 8) Vgl. hierzu beispielhaft die Kontroverse zwischen Burger (1994) in Kritik zu Erxleben et al. (1992) und die Entgegnung von Baetge/Krause/Mertens (1994) in der Zeitschrift für Betriebswirtschaft.
- 9) Vgl. hierzu ausführlich Hauschildt (1996) S. 1 ff.
- 10) Zur Interdependenz von Jahresabschlußgestaltung und Jahresabschlußbeurteilung vgl. Küting (1996) S. 934 ff.
- 11) Vgl. Hauschildt/Leker (1995).
- 12) Vgl. hierzu ausführlich Gemünden (1988) S. 137 ff.
- 13) Vgl. hierzu die Übersicht bei Krehl (1988) S. 17 ff.
- 14) Vgl. allgemein zur Vorgehensweise der Diskriminanzanalyse Backhaus/Ericksen/Plinke/Weiber (1994) S. 90 ff.
- 15) Vgl. hierzu Niehaus (1987) S. 92 und 156, der zu dem Ergebnis kommt, daß die multivariate Diskriminanzanalyse im Gegensatz zu entsprechenden verteilungsfreien Klassifikationsverfahren zu besseren Klassifikationsergebnissen führt, obwohl die Normalverteilungsannahme nicht erfüllt ist. Vgl. ähnlich Baetge/Beuter/Feidicker (1992) S. 751 f. und Hüls (1995) S. 277 f. Kritisch sprechen sich hingegen Gemünden (1988) S. 146 und Burger (1994) S. 1167 aus.
- 16) Vgl. hierzu ausführlich Aldrich/Nelson (1984) S. 48 ff., DeMaris (1992) S. 42 ff. und Stegemann (1995) S. 4 ff.
- 17) Vgl. ausführlich zu den unterschiedlichen Formen der ML-Schätzung Stegemann (1995) S. 10 ff.
- 18) Vgl. hierzu Urban (1993) S. 13 der hier eine Zahl von 50 Untersuchungseinheiten nennt. Aldrich und Nelson (1984) S. 53 fordern für die Differenz von Stichprobenumfang und zu schätzenden Parametern mindestens eine Größenordnung von 100.
- 19) Zur Funktion von künstlichen neuronalen Netzen bei der Klassifikation von Unternehmen vgl. Krause (1993) S. 33 ff., Kerling/Poddig (1994) S. 427 ff., Burger (1995) S. 341 ff. und Rehkugler (1995) S. 250 ff. Vgl. ferner allgemein zum Forschungsgebiet des "Konnektionismus" Dorffner (1991).
- 20) Vgl. hierzu Krause (1993) S. 124 ff.

- 21) Vgl. z.B. Feidicker (1992) S. 212 f., der für ein Kreditversicherungsunternehmen ein kostenoptimales Verhältnis von Alpha- und Beta-Fehler von 1 : 5,26 ermittelt.
- 22) Vgl. Odom/Sharda (1990). Vgl. ferner auch die empirischen Vergleiche bei Coats/Fant (1991) S. 9 ff. und Altman/Marco/Varetto (1994) S. 518 ff.
- 23) Vgl. Tam/Kiang (1990).
- 24) Vgl. hierzu ausführlich Hüls (1995) S. 154 f.
- 25) Vgl. Rehkugler/Poddig (1992) sowie Kerling/Poddig (1994) S. 427 ff.
- 26) Kerling/Poddig (1994) S. 484.
- 27) Vgl. Erxleben/Baetge/Feidicker/Koch/Krause/Mertens (1992) S. 1237 ff.
- 28) Vgl. z.B. Niehaus (1987) S. 65 f.
- 29) Vgl. hierzu ausführlich Leker (1994).
- 30) Das Netz wurde mittels der Software SNNS der Universität Stuttgart entwickelt und läuft unter Unix auf einer Sun-Workstation. Vgl. ausführlich zur Beschreibung der Netzspezifika Baetge/Jerschensky/Herrmann/Stibi (1995) und Baetge/Hüls/Uthoff (1995).
- 31) Vgl. zur Operationalisierung der Kennzahlen Hüls (1995) S. 75 ff.
- 32) Für eine Darstellung der üblicherweise zur Anwendung kommenden Pruning-Algorithmen vgl. Rehkugler/Poddig (1992), S. 15-23.
- 33) Vgl. Baetge/Hüls/Uthoff (1995) S. 25, wonach ein relevanzorientiertes Pruning der versteckten Schicht keine Verbesserung der Klassifikationsgüte brachte.
- 34) Hierbei ist allerdings einschränkend zu bemerken, daß bei der Ermittlung des besten Netzes schon während der Trainingsphase die Ergebnisse der Holdout-Stichprobe berücksichtigt wurden. Eine Vorgehensweise, die, wie auch Kerling und Poddig einräumen, einen zu optimistischen Wert für die Holdout-Stichprobe ermittelt. Vgl. Kerling/Poddig (1994) S. 473 f.
- 35) Erschwerend kommt hinzu, daß auch die Auswahl der nicht insolventen Vergleichsfälle vom Kreditinstitut vorgenommen wurde.
- 36) Dieser Aspekt gewinnt in unserer Analyse insofern noch an Bedeutung, als wir unser Klassifikationsziel von der Insolvenz auf die stärker durch das jeweilige Institut geprägten "Leistungsstörungen" ausgedehnt haben.
- 37) Studien, die eine entsprechende Übertragung anderer Klassifikationsmodelle unternommen haben, berichten demgegenüber von deutlichen Verschlechterungen des Klassifikationsergebnisses, die bis zum völligen Versagen des Modells reichen. Vgl. z.B. Gemünden (1988) S. 148 ff.
- 38) Die abhängige Variable nimmt den Wert "1" bei "insolvenzgefährdet" und den Wert "0" bei "nicht insolvenzgefährdet" an.
- 39) Aktuelle Forschungsbemühungen versuchen dieser Kritik Rechnung zu tragen und entwickeln neue Verfahren, die eine bessere Interpretation der Klassifikation erlauben. Vgl. Baetge/Kruse/Uthoff (1996).
- 40) Vgl. zu einer Übersicht der Anwendung künstlicher neuronaler Netze auf betriebswirtschaftliche Fragestellungen Sharda (1994) S. 116 ff.
- 41) Vgl. Baetge/Kruse/Uthoff (1996) S. 280.
- 42) Vgl. Oestreicher/Piotrowski-Allert (1996) S. 335 ff.

Literaturverzeichnis

- Adler/Düring/Schmaltz (1987): Rechnungslegung und Prüfung der Unternehmen: Kommentar zum HGB, AktG, GmbHG, PublG nach den Vorschriften des Bilanzrichtlinien-Gesetzes, bearbeitet von Forster, K. H., Goerdeler, R., Lanfermann, J., Müller, H.P., Siepe, G., Stolberg, K. und Weirich, S., 5. Aufl., Stuttgart 1987.
- Aldrich, J.H. und Nelson, F.D. (1984): Linear Probability, Logit and Probit Models, Newbury Park (CA) 1984.
- Altman, E.I., Marco, G. und Varetto, F. (1994): Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (the Italian Experience), in: Journal of Banking and Finance 18 (1994), S. 505-529.
- Backhaus, K.; Erichson, B., Plinke, W. und Weiber, R. (1994): Multivariate Analysemethoden, 7. Aufl., Berlin 1994.

- Baetge, J., Beuter, H.B. und Feidicker, M. (1992): Kreditwürdigkeitsprüfung mit Diskriminanzanalyse, in: Die Wirtschaftsprüfung 45 (1992), S. 749-761.
- Baetge, J., Krause, C., Mertens, P., (1994): Zur Kritik an der Klassifikation von Unternehmen mit Neuronalen Netzen und Diskriminanzanalysen: Stellungnahme zum Beitrag von A. Burger, in: Zeitschrift für Betriebswirtschaft 64 (1994), S. 1181-1191.
- Baetge, J., Hüls, D. und Uthoff, C. (1995): Früherkennung der Unternehmenskrise, in: FJ WWU Münster (2/1995), S. 21-29.
- Baetge, J., Jerschensky, A., Herrmann, D., Stibi, B. (1995): Die Auswahl prüfungsbedürftiger Betriebe im Rahmen der steuerlichen Außenprüfung, in: Der Betrieb 48 (1995), S. 585-594.
- Baetge, J., Kruse A. und Uthoff, C. (1996): Bonitätsklassifikationen von Unternehmen mit Neuronalen Netzen, in: Wirtschaftsinformatik 38 (1996), S. 273-281.
- Bieg, H. (1984): Kann der Bankprüfer die Bonität gewerblicher Bankkreditnehmer beurteilen?, in: Zeitschrift für betriebswirtschaftliche Forschung 36 (1984), S. 495-512.
- Burger, A. (1994): Zur Klassifikation von Unternehmen mit neuronalen Netzen und Diskriminanzanalysen, in: Zeitschrift für Betriebswirtschaft 64 (1994), S. 1165-1179.
- Burger, A. (1995): Jahresabschlußanalyse, München 1995.
- Coats, P.K. und Fant, L.F.: (1991): A Neural Network Approach to Forecasting Financial Distress, in: The Journal of Business Forecasting 10 (1991), S. 9-12.
- DeMaris, A. (1992): Logit Modeling - Practical Applications, Newbury Park (CA) 1992.
- Dorffner, G. (1991): Konnektionismus: Von neuronalen Netzwerken zu einer "natürlichen" KI, Stuttgart 1991.
- Erxleben, K., Baetge, J., Feidicker, M., Koch, H., Krause, C. und Mertens, P. (1992): Klassifikation von Unternehmen - ein Vergleich von Neuronalen Netzen und Diskriminanzanalyse, in: Zeitschrift für Betriebswirtschaft 62 (1992), S. 1237-1262.
- Feidicker, M. (1992): Kreditwürdigkeitsprüfung: Entwicklung eines Bonitätsindikators, Düsseldorf 1992.
- Gebhardt, G. (1980): Insolvenzprognosen aus aktienrechtlichen Jahresabschlüssen, Wiesbaden 1980.
- Gemünden, H.-G. (1988): Defizite der empirischen Insolvenzforschung, in: Hauschildt, J. (Hrsg.), Krisendiagnose durch Bilanzanalyse, Köln 1988, S. 135-152.
- Hauschildt, J. (1996): Erfolgs-, Finanz- und Bilanzanalyse: Analyse der Vermögens-, Finanz- und Ertragslage von Kapital- und Personengesellschaften, 3. Aufl., Köln 1996.
- Hauschildt, J. und Leker, J. (1995): Bilanzanalyse unter dem Einfluß moderner Analyse- und Prognoseverfahren, in: Betriebswirtschaftliche Forschung und Praxis 47 (1995), S. 249-268.
- Hosmer, D.W. und Lemeshow, S. (1989): Applied Logistic Regression, New York 1989.
- Hüls, D. (1995): Früherkennung insolvenzgefährdeter Unternehmen, Düsseldorf 1995.
- Kumar, A., Rao, V.R. und Soni, H. (1995): An Empirical Comparison of Neural Network and Logistic Regression Models, in: Marketing Letters 6 (1995), S. 251-263.
- Küting, K. (1996): Das Spannungsverhältnis zwischen Bilanzpolitik und Bilanzanalyse: Zur Interdependenz von Jahresabschlußgestaltung und Jahresabschlußbeurteilung, in: Deutsches Steuerrecht 34 (1996), S. 934-944.
- Kerling, M. und Poddig, T. (1994): Klassifikation von Unternehmen mittels KNN, in: Rehkugler, H. und Zimmermann, H.G. (Hrsg.), Neuronale Netze in der Ökonomie, München 1994, S. 426-490.
- Krause, C. (1993): Kreditwürdigkeitsprüfung mit Neuronalen Netzen, Münster 1993.
- Krehl, H. (1988): Krisendiagnose durch klassische Bilanzkennzahlen?, in: Hauschildt, J. (Hrsg.), Krisendiagnose durch Bilanzanalyse, Köln 1988, S. 17-40.
- Lachnit, L. (1979): Systemorientierte Jahresabschlußanalyse, Wiesbaden 1979.
- Leker, J. (1994): Fraktionierende Frühdiagnose von Unternehmenskrisen anhand von Jahresabschlüssen, in: Zeitschrift für betriebswirtschaftliche Forschung 46 (1994), S. 732-750.
- Markham, I.S. und Ragsdale, C.T. (1995): Combining Neural Networks and Statistical Predictions to Solve the Classification Problem in Discriminant Analysis. in: Decision Sciences 26 (1995), S. 229-242.
- Niehaus, H.-J. (1987): Früherkennung von Unternehmenskrisen, Düsseldorf 1987.
- Odom, M.D. und Sharda, R. (1990): Neural Network Model for Bankruptcy Prediction, in: IEEE (Hrsg.), Proceedings of the International Joint Conference on Neural Networks, Band 2, San Diego 1990, S. 163-167.

- Oestreicher, A. und Piotrowski-Allert, S. (1996): Klassifikation und Beurteilung von Unternehmen mit Hilfe von Selbstorganisierenden Neuronalen Netzen, in: Zeitschrift für betriebswirtschaftliche Forschung 48 (1996), S. 335-371.
- Rautenberg, H.G. (1990): Externe Rechnungslegung: Maßgeblichkeitsprinzip und Finanzierung, in: Ahlert, D.; Franz, K.-P. und Göpl, H. (Hrsg.), Finanz- und Rechnungswesen als Führungsinstrument, Wiesbaden 1990, S. 252-268.
- Rehkugler, H. (1995): Neuronale Netze als Instrument der Jahresabschlußanalyse, in: Scheer, A.-W. (Hrsg.), Rechnungswesen und EDV: 16. Saarbrücker Arbeitstagung 1995, Heidelberg 1995, S. 248-263.
- Rehkugler, H. und Poddig, T. (1992): Klassifikation von Jahresabschlüssen mittels Multilayer-Perceptrons, Bamberger Betriebswirtschaftliche Beiträge Nr. 87, Bamberg 1992.
- Schneider, D. (1985): Die Warnung vor Frühwarn-Systemen, in: Der Betrieb 38 (1985), S. 1489-1494.
- Schneider, D. (1989): Erste Schritte zu einer Theorie der Bilanzanalyse, in: Die Wirtschaftsprüfung 42 (1989), S. 633-642.
- Sharda, R. (1994): Neural Networks for MS/OR Analyst: An Application Bibliography, in: Interfaces 24 (1994), S. 116-130.
- Stegemann, S. (1995): Logistische Regression und Diskrimination bei zufällig variierenden Parametern, Wuppertal 1995.
- Tam, K.Y. und Kiang, M. (1990): Predicting Bank Failures: A Neural Network Approach, in: Applied Artificial Intelligence 4 (1990), S. 283-307.
- Urban, D. (1993): Logit-Analyse: Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen, Stuttgart 1993.

Tab. 1: Trefferquoten der Klassifikationsmodelle bei der Holdout-Stichprobe

	Alpha-Fehler	Beta-Fehler	Gesamtfehler
MDA	12,1 %	25,0 %	18,5 %
LRA	10,3 %	21,2 %	15,7 %
KNN	14,7 %	29,3 %	22,0 %

Tab. 2: Trennscharfe "Jahresabschluß-Kennzahlen" der Klassifikationsmodelle

Kennzahl- Nummer	Jahresabschluß- Kennzahl	Plausibilität ^{*)}
MDA:		
K1	Gesamtkapitalrentabilität	I < S
K2	Anteil Fremdkapitalzinsen	I < S
K3	Anteil Vorräte	I > S
K4	Kundenziel	I > S
K5	Anteil langfristiges Kapital	I < S
LRA:		
K1	Gesamtkapitalrentabilität	I < S
K2	Anteil Fremdkapitalzinsen	I < S
K3	Anteil Vorräte	I > S
K4	Kundenziel	I > S
K5	Anlagendeckung	I < S
KNN:		
K1	Umsatzrentabilität	I < S
K2	Finanzwirtschaftliche Rentabilität I	I < S
K3	Finanzwirtschaftliche Rentabilität II	I < S
K4	Kehrwert des dynam. Verschuldungsgrad I	I < S
K5	Kehrwert des dynam. Verschuldungsgrad II	I < S
K6	Kapitalumschlag I	I > S
K7	Kapitalumschlag II	I > S
K8	Kapitalumschlag III	I > S
K9	Statischer Verschuldungsgrad I	I > S
K10	Statischer Verschuldungsgrad II	I > S
K11	Personalintensität	I > S
K12	Anlagendeckung I	I < S
K13	Anlagendeckung II	I < S
K14	Anlagendeckung III	I < S

^{*)} Plausibilität: Ausprägung der Kennzahl bei insolventen und solventen Unternehmen

Abb. 1.: Struktur des KNN zur Unternehmensklassifikation

