

Universität Hohenheim  
Institut für Betriebswirtschaftslehre  
Lehrstuhl für Bankwirtschaft  
Prof. Dr. Hans-Peter Burghof



Seminararbeit zum Thema:

## **Methodology for Predicting Bankruptcy**

Eingereicht am 17.05.2010

# Inhaltsverzeichnis

<b>1.</b>	<b>Einleitung</b> .....	1
<b>2.</b>	<b>Erste Modelle</b> .....	1
2.1	Dichotomische Einteilung nach Beaver.....	1
2.2	Z-Score nach Altman .....	3
2.2.1	Modell .....	3
2.2.2	Empirische Validität des Z-Score .....	4
<b>3.</b>	<b>Artificial Neural Networks</b> .....	6
3.1	Hintergrund .....	6
3.1.1	Idee .....	6
3.1.2	Aufbau .....	6
3.2	Konstruktion, Training und Testing eines ANN .....	8
3.2.1	Konstruktion.....	8
3.2.2	Training.....	9
3.2.3	Testing.....	12
3.3	Empirische Validität von ANNs.....	13
<b>4.</b>	<b>Fazit</b> .....	155
I.	Abbildungen.....	II
II.	Literaturverzeichnis .....	IV

## 1. Einleitung

Mit dem Aufkommen der ersten Kennzahlenanalysen um 1900 war noch nicht abzusehen, dass diese im folgenden Jahrhundert und darüber hinaus Basis zahlreicher Forschungsarbeiten und Untersuchungen werden sollten. Die erste bekannte Kennzahl stellt dabei die current ratio dar, welche sich als Verhältnis von Umlaufvermögen zu kurzfristigen Verbindlichkeiten errechnet und als Indiz für die Kreditwürdigkeit galt.<sup>1</sup> Schon bald erkannte man den möglichen Nutzen, den man aus der genaueren Untersuchung von Kennzahlen im Hinblick auf die zukünftige Zahlungsfähigkeit von Firmen ziehen konnte, weshalb alsbald erste Versuche unternommen wurden, Modelle zu entwickeln, die gerade so eine Prognose ermöglichen sollten.

Diese Arbeit befasst sich mit einer Auswahl von solchen Modellen. Es wird dabei sowohl auf sehr frühe Ansätze eingegangen als auch ein aktuelles „state-of-the-art“ Konzept beleuchtet. Begonnen wird mit der Betrachtung der Arbeit von Beaver, welcher sich genauer mit der Aussagekraft einzelner Kennzahlen im Hinblick auf Bankrottprognosen befasst hat. Im Anschluss wird auf das populäre Z-Score-Modell von Altman eingegangen, welches die Summe unterschiedlich gewichteter Kennzahlen als Basis für die Vorhersage von Insolvenzen benutzt. Der Fokus dieser Arbeit liegt jedoch auf einem aktuellen Ansatz, bei welchem ein künstliches Neuronales Netzwerk benutzt wird, um eine Unterscheidung zwischen künftig solventen und künftig insolventen Firmen zu ermöglichen. Dabei wird zunächst auf den allgemeinen Aufbau sowie die hintergründige Idee eingegangen, das inhärente mathematische Konstrukt erläutert sowie ein Überblick darüber gegeben, welche empirische Validität ein solches System besitzt. Das übergeordnete Ziel dieser Herangehensweise ist dabei, einen historischen Überblick über Insolvenzprognosemodelle zu liefern, deren Möglichkeiten und Grenzen aufzuzeigen sowie auf Tendenzen für zukünftige Forschung hinzudeuten.

## 2. Erste Modelle

### 2.1 Dichotomische Einteilung nach Beaver

1966 war William H. Beaver einer der Vorreiter auf dem Gebiet der Kennzahlenanalyse. Sein Anspruch dabei war es, eine empirische Untersuchung über den Nutzen von aus der Buchhaltung stammendem Zahlenmaterial durchzuführen um damit eine Basis für zukünftige Forschung auf diesem Gebiet zu erschaffen. Dazu betrachtete er im Zeitraum von 1954 bis 1964 insgesamt 79 Unternehmen die insolvent gegangen sind und von denen die Jahresabschlüsse von wenigstens einem Jahr vor der Pleite verfügbar waren. Zu jedem dieser Unter-

---

<sup>1</sup> vgl. Beaver (1966), S. 71

nehmen wurde nun ein weiteres - nicht insolventes - Unternehmen der gleichen Branche und ähnlicher Größe hinzugezogen.<sup>2</sup> Für diese Stichprobe errechnete Beaver nun je 30 Kennzahlen, die ihm als Datenbasis für die Erschaffung seines Modells dienen sollten (s. Abb. 6). Diese Ratios wurden anschließend in 6 Gruppen mit ähnlichen Kennzahlen unterteilt und je eine Kennziffer pro Gruppe ausgewählt von der angenommen werden konnte, dass sie den größtmöglichen Informationsbeitrag liefert. Diese 6 Ratios wurden nun dahingehend untersucht, dass der kritische Wert ermittelt wurde, ab dem ein zukünftiger Bankrott der Firma am wahrscheinlichsten ist. Anschließend wurde anhand dieser Erkenntnis für jede Unternehmung eine Klassifikation in „bankrott“ oder „nicht bankrott“ vorgenommen und abgeglichen, ob die jeweilige Firma tatsächlich den vorhergesagten Status erlangt hat. Es stellte sich dabei heraus, dass „the ability to predict failure is strongest in the cash-flow to total-debt ratio“<sup>3</sup>. Mittels dieser Kennzahl war es möglich, im Jahr vor der Firmenpleite 87% der Unternehmen korrekt einzuteilen – fünf Jahre vorab betrug die Genauigkeit immerhin noch 78%.<sup>4</sup> Dieses Ergebnis ist nicht zuletzt deshalb so überraschend, als dass lediglich eine einzige Kennzahl ausreichte um bereits Jahre vor der tatsächlichen Insolvenz eben diese mit bemerkenswert hoher Präzision vorherzusagen. Fiel diese Kennzahl unter 0,02 und stieg anschließend nicht mehr über 0,08, so war der zukünftige Bankrott der Firma als höchstwahrscheinlich anzusehen.<sup>5</sup> Trotz der hohen Genauigkeit dieses Modells ist Beavers Studie nur eine unter zahlreichen Arbeiten, die sich mit dem Prognosepotenzial einzelner Ratios beschäftigen und von denen fast jede eine andere Kennzahl als am aussagekräftigsten identifiziert. „In general, ratios measuring profitability, liquidity, and solvency prevailed as the most significant indicators. The order of their importance is not clear since almost every study cited a different ratio as being the most effective indication of impending problems.“<sup>6</sup> Die empirische Validität dieses sehr anspruchsvollen, da besonders simplen, Modells ist daher nicht gewährleistet. Aus eben diesem Grund finden sich auch kaum Arbeiten, die sich mit diesem sehr speziell gehaltenen Modell auseinandersetzen. Dieser Umstand ist auch nicht zuletzt darauf zurückzuführen, dass bereits kurz darauf eine andere Methode entwickelt wurde, welche den Anspruch erhob, allgemeingültiger zu sein. Nichtsdestotrotz sollte Beaver als einer der Pioniere auf dem Forschungsgebiet der Bankrottprognose gewürdigt werden, denn die grundlegende Herangehensweise, nämlich Ratios als Insolvenzindikatoren zu verwenden, fließt bis heute in zahlreiche Arbeiten auf diesem Gebiet mit ein.

---

<sup>2</sup> vgl. Beaver (1966), S. 72-73

<sup>3</sup> Beaver (1966), S. 85

<sup>4</sup> vgl. Beaver (1966), S. 85

<sup>5</sup> vgl. Beaver (1968), S. 187

<sup>6</sup> Altman (1968), S. 590

## 2.2 Z-Score nach Altman

### 2.2.1 Modell

Eines der ersten Mehrvariablenmodelle, um die Zahlungsunfähigkeit und somit den Bankrott einer Unternehmung vorherzusagen, entwickelte Edward I. Altman im Jahr 1968. In seiner Arbeit „Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy“ entwirft er ein von ihm als Z-Score bezeichnetes Maß, welches sich aus einer Summe unterschiedlich gewichteter Ratios zusammensetzt. Dabei wurden 66 Unternehmen, von denen 33 insolvent wurden, betrachtet. Von all diesen Firmen wurden 22 unterschiedliche Kennzahlen berechnet, die sich entweder in vorherigen Studien als besonders aussagekräftig dargestellt hatten und somit damals sehr populär waren oder die für die Studie potentiell von Relevanz waren. Außerdem wurden einige neuartige Ratios herangezogen, die Altman zum Zweck dieser Studie ebenfalls mit einbrachte. Aus diesem Pool wurden fünf besonders signifikante Kennzahlen ausgewählt, welche in den Fokus der Arbeit rücken sollten. Die Ratios wurden dabei jedoch nicht nach dem Aspekt der jeweils einzelnen Signifikanz für eine Prognose ausgewählt, sondern nach ihrer Bedeutung als Kennzahlengruppe für eine Vorhersage. Bei Altmans Modell handelt es sich mathematisch betrachtet um eine multivariate Diskriminanzanalyse. Dies ist eine statistische Technik, die dazu verwendet wird, Beobachtungen in vorab festgelegte Gruppen einzuteilen, wobei die Klassifikation anhand der Ausprägung der individuellen Merkmale eben dieser Beobachtung erfolgt.<sup>7</sup>

Die relevanten Kennzahlen in dem Modell von Altman sind:

$$X_1 = \text{Working Capital} / \text{Total Assets}$$

$$X_2 = \text{Retained Earnings} / \text{Total Assets}$$

$$X_3 = \text{EBIT} / \text{Total Assets}$$

$$X_4 = \text{Market Value Equity} / \text{Book Value of Total Debt}$$

$$X_5 = \text{Sales} / \text{Total Assets}$$

Diese Kennziffern werden in folgender Weise gewichtet und addiert, sodass sich der entsprechende Z-Score ergibt.

$$Z = 0,012 X_1 + 0,014 X_2 + 0,033 X_3 + 0,006 X_4 + 0,999 X_5$$

---

<sup>7</sup> vgl. Altman (1968), S. 591

Bei dieser Schreibweise der Formel ist zu beachten, dass die sich ergebenden Verhältniszahlen in absoluter Prozentschreibweise einzubringen sind (z.B. falls  $X_4 = 150\% = 1,5$  so wird  $X_4$  mit 150 eingebracht). Für den Fall, dass die mathematisch etablierte Schreibweise für die Prozentwerte als Dezimalzahl verwendet werden soll, muss die Formel in folgender Weise angepasst werden:

$$Z = 1,2X_1 + 1,4X_2 + 3,3X_3 + 0,6X_4 + 1,0X_5^8$$

Hierbei wurde sowohl die Variablenengewichtung für  $X_1$  bis  $X_4$  von der Schreibweise her angepasst, als auch die Gewichtung von  $X_5$  geringfügig gerundet. Die Sales / Total Assets Ratio ( $X_5$ ) geht jedoch weiterhin als absolute Zahl in die Gleichung ein, weshalb hier auch keine Veränderung des Gewichtungsfaktors vorgenommen werden musste. Mit Hilfe dieser Formel lässt sich nun der Z-Score ermitteln. Sobald dieser Wert über 2,99 liegt, lässt sich eine betrachtete Unternehmung in die Klasse „nicht bankrott“ einordnen, wohingegen bei einem Wert kleiner als 1,81 eine klare Zuordnung in „bankrott“ erfolgen kann. Aufgrund der Tatsache, dass es zwischen den Werten eine Grauzone gibt, in der keine eindeutige Klassifikation möglich ist, musste jedoch für eben diesen Ergebnisraum ein Median gefunden werden, der als kritischer Wert innerhalb dieses Bereiches dient.<sup>9</sup> Diesen ermittelte Altman indem er das Intervall mit der geringsten Anzahl an Falschklassifikationen innerhalb dieser Zone bestimmte. Dieses Intervall lag bei 2,67 bis 2,68, weshalb ein kritischer Z-Score von 2,675 als bestmöglicher Trennwert zwischen solventen und zukünftig insolventen Firmen verwendet wird.<sup>10</sup>

### 2.2.2 Empirische Validität des Z-Score

Altman selbst stellte bereits in seiner Arbeit fest, dass die Zuverlässigkeit seines Modells rapide nachlässt, je länger eine betrachtete Unternehmung vom tatsächlichen Bankrott entfernt ist. Ein Jahr vor der Insolvenz ließen sich mittels dem Z-Score 95% der Firmen korrekt klassifizieren. Dieser Wert ist aber nicht zuletzt deshalb mit Vorsicht zu genießen, da es sich bei der betrachteten Stichprobe um gerade die Unternehmen handelt, auf Basis derer die Zahlen für das Z-Score-Modell entstanden sind. Die hohe Trefferquote ist lediglich ein Indiz dafür, dass das Modell in sich konsistent ist und die Erarbeitung korrekt erfolgt ist. Doch selbst bei dieser Stichprobe verliert das Modell an Zuverlässigkeit, je weiter der Bankrott in der Zukunft liegt. Bereits zwei Jahre vor Zahlungsunfähigkeit können nur noch 72% der Unternehmen korrekt klassifiziert werden, vier Jahre vorher lediglich 29% (s. Abb. 7).<sup>11</sup> Weitaus interessanter ist die Betrachtung einer aktuelleren Stichprobe von Firmen, die nicht Teil der Datenbasis von Altman waren. Mit genau dieser Thematik haben sich Gerantonis et al. in

---

<sup>8</sup> Altman (2000), S. 13

<sup>9</sup> vgl. Altman (1968), S. 606

<sup>10</sup> vgl. Altman (1968), S. 607

<sup>11</sup> vgl. Altman (1968), S. 604

ihrem Paper „Can Altman Z-Score Models Predict Business Failures in Greece?“ von 2009 beschäftigt. Dazu wurde eine Grundgesamtheit von 373 Unternehmen, welche an der Athener Börse gehandelt werden bzw. wurden, im Zeitraum von 1999 - 2006 untersucht. 44 dieser Unternehmen gingen im Laufe der betrachteten Periode Bankrott bzw. wurden permanent vom Handel an der Börse ausgeschlossen. Mittels der Z-Score Formel von Altman 1993 ergaben sich folgende Ergebnisse. Bei ausschließlicher Betrachtung der insolvent gegangenen Firmen war die Prognosegenauigkeit des Z-Score ein Jahr vor der Unternehmenspleite 66%. Somit wurden von den insgesamt 44 Firmenpleiten 29 im Jahr vor der Zahlungsunfähigkeit korrekt klassifiziert. Die restlichen 15 Unternehmen unterlagen dem Typ-II-Fehler und wurden trotz drohender Insolvenz als liquide eingestuft. Bei den Unternehmen denen kein Bankrott drohte wurden ebenfalls ein Jahr vor der prognostizierten Insolvenz 54% korrekt eingeordnet. Besonders auffällig an dieser Studie ist jedoch, dass die Zuverlässigkeit der Einordnung, je nachdem ob es sich um zukünftig solvente oder insolvente Firmen handelt, mit längerem Betrachtungshorizont entweder zu- oder abnimmt. Das soll heißen, dass beispielsweise 4 Jahre vor dem Bankrott lediglich 20% der tatsächlich gefährdeten Firmen als solche identifiziert werden konnten, wohingegen die ungefährdeten Firmen 4 Jahre vorab zu 78% richtig klassifiziert worden sind.<sup>12</sup> Es ist somit festzuhalten, dass das Z-Score Modell eine allgemein gute Prognose zulässt. Insbesondere die Tatsache, dass langfristig solvente Unternehmungen als solche erkannt werden sowie die Tatsache, dass zumindest ein Jahr vorab über die Hälfte der bankrotten Firmen als eben solche identifiziert werden können ist keine schlechte Performance. (s. Abb. 8-10) In ähnlicher Art und Weise untersuchten Grice und Ingram den Z-Score mit einer Stichprobe von 972 Unternehmen aus den Jahren 1988 – 1991. Sie kamen dabei auf eine Gesamtgenauigkeit von 57,8%. Weiterhin stellen sie fest, dass die Gewichtungsfaktoren, die Altman errechnet hat, je nach betrachtetem Sample angepasst werden sollten, um eine höhere Genauigkeit zu erzielen.<sup>13</sup> Abschließend ist für die Evaluierung der empirischen Validität festzuhalten: „[The] results of recent studies that have used Altman’s model to estimate financial distress of sample firms should be interpreted cautiously. The ability of the model to accurately classify firms as being financially distressed is likely to differ considerably from that assumed by those employing the model.“<sup>14</sup>

---

<sup>12</sup> Gerantonis et al. (2009), S. 22-25

<sup>13</sup> vgl. Grice, Ingram (2001), S. 60

<sup>14</sup> Grice, Ingram (2001), S. 60

### 3. Artificial Neural Networks

#### 3.1 Hintergrund

##### 3.1.1 Idee

Die bisherigen Untersuchungen zur Vorhersage von Bankrott basierten neben den oben diskutierten Beispielen für uni- und multivariate Diskriminanzanalysen auch auf Logit- und Probit-Modellen<sup>15</sup>. Da die Validität und Effektivität dieser Methoden jedoch stark von einigen statistischen Annahmen wie Linearität, Unabhängigkeit und Normalverteilung der Ausgangsvariablen abhängt, war es von Nöten, eine neue Herangehensweise zu entwickeln, die von eben diesen Kriterien unabhängig ist.<sup>16</sup> Um 1990 wurde begonnen, in die Richtung künstlicher Neuroner Netzwerke (Artificial Neural Networks, im folgenden ANN) zu forschen, wobei die Arbeit von Odom und Sharda<sup>17</sup> eine der ersten war, die sich mit der Implementierung von ANNs zur Prognose von Firmenpleiten befasste.<sup>18</sup> ANNs sind flexible, nicht-parametrische Modellierungsinstrumente, die komplexe Funktionen abzubilden vermögen.<sup>19</sup> An dieser Stelle ist der Fokus zunächst auf die Begrifflichkeit der nicht-parametrischen Betrachtung zu lenken. Nicht-parametrische Statistik bedeutet nicht etwa, dass es keine Parameter gibt, sondern vielmehr handelt es sich um das Pendant zur parametrischen Statistik, welche durch zahlreiche restriktive Annahmen geprägt ist. Nicht-parametrische Statistik ist also eine Art Sammelbegriff, der die statistischen Methoden zusammenfasst, deren Ergebnisse unter weniger restriktiven Annahmen valide sind.<sup>20</sup> Die Idee hinter einem künstlichen Neuronalen Netzwerk ist es, einer Maschine beizubringen, logisch zu schlussfolgern, zu lernen und einfache Aufgaben so zu lösen, wie es ein Mensch könnte. Dabei soll es sich jedoch nicht um ein programmiertes System per se handeln, welches eine speziell vorgegebene Reihenfolge an Rechenschritten vollführt, um sodann ein Ergebnis zu fabrizieren, sondern es soll vielmehr eigene Regeln und Formeln zur Problemlösung auf Basis einiger Beispiele entwickeln.<sup>21</sup>

##### 3.1.2 Aufbau

Ein ANN ist ein Geflecht aus zahlreichen Knotenpunkten, den so genannten Nodes, welche auf mehreren Ebenen angeordnet und miteinander verbunden sind. Bei den Verknüpfungssträngen zwischen den Nodes gibt es je nach Typ des ANN Unterschiede. Die derzeit popu-

---

<sup>15</sup> Amemya (1981), S. 1499: „[...] the word probit is used to denote a QR (Qualitative Response, Anm. des Autors) model where the probability function is a standard normal distribution function. Similarly, the word logit [...] means a QR model with a logistic distribution.“

<sup>16</sup> vgl. Zhang et al. (1997), S. 1

<sup>17</sup> vgl. Odom, Sharda (1990)

<sup>18</sup> vgl. Atiya (2001), S. 930

<sup>19</sup> vgl. Zhang et al. (1997), S. 17

<sup>20</sup> vgl. Gibbons (1993), S. 1

<sup>21</sup> vgl. Udo (1993), S. 378



lärste Form ist das multi-layer perceptron (MLP). Wie der Name vermuten lässt, handelt es sich bei diesem Netzwerk um eines von mehrlagiger Gestalt, wobei für den Zweck der Bankrottprognose idealerweise drei Lagen zu wählen sind. Ein weiteres Merkmal dieser Netzwerkarchitektur ist die Verknüpfung zwischen den Nodes sowie die Richtung des Informationsflusses. Da es sich bei dem MLP um ein feedforward-Netzwerk handelt, fließen die Informationen nur in eine Richtung. Dabei wird das MLP in der untersten Schicht, der Input-Layer, mit Daten versorgt, welche dann an die mittlere Hidden-Layer weitergegeben und dort verarbeitet werden. Im Anschluss an dieses Procedere erfolgt die Ausgabe des Ergebnisses über die dritten, so genannten Output-Layer. Das Ergebnis ist so geartet, dass eine Klassifikation in eine von zwei Gruppen vorgenommen wird. In dem betrachteten Fall der Bankrottprognose sind dies folglich „bankrott“ oder „nicht bankrott“.<sup>22</sup>

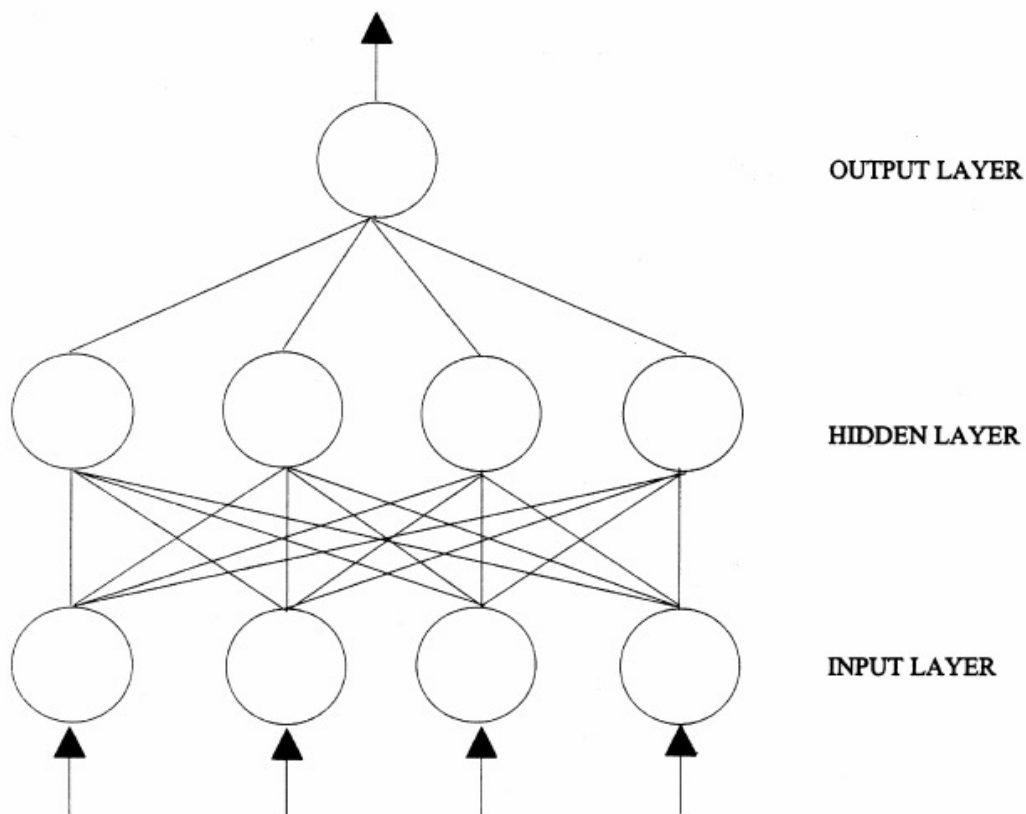


Abb. 1: Aufbau eines MLP<sup>23</sup>

<sup>22</sup> vgl. Zhang et al. (1997), S. 17

<sup>23</sup> Zhang et al. (1997), S. 18

## 3.2 Konstruktion, Training und Testing eines ANN

### 3.2.1 Konstruktion

Um ein ANN in den betriebsfähigen Zustand zu versetzen, ist es notwendig, das Netzwerk so zu programmieren, dass es der gewünschten Aufgabe gerecht werden kann. Dieser Vorgang der Implementierung der grundlegenden Architektur und Funktionsweise wird auch als „Training“ bezeichnet. Während dieser Phase werden Beispieldaten über die Input-Layer in das ANN eingegeben, wobei es während der Trainingsphase geboten ist, das jeweilig erwünschte Ergebnis der Berechnung schon vorab zu kennen. Die Eingangsdaten werden dann von der Input-Layer an die Hidden-Layer weitergegeben, wo sie an jedem Knotenpunkt gewichtet und kumuliert werden. Die gewichtete Summe bildet dann den neuen Wert, der mittels einer geeigneten Transferfunktion an die nächsthöhere Schicht, in diesem Fall die Output-Layer weitergegeben wird. Nach Abschluss der Berechnungen in der letzten Schicht wird ein Ergebnis ausgegeben, welches sich im Beispiel der Bankrottprognose als 0 oder 1 darstellt und somit eine Klassifikation in „insolvent“ und „solvent“ vornimmt. Ziel der Trainingsphase muss es nun sein, Unterschiede zwischen dem vorab bekannten, erwünschten Ergebnis und der durch Berechnung im ANN erhaltenen Einteilung zu minimieren.<sup>24</sup> Dieser Prozess soll im Folgenden auf Basis der Arbeit von Rumelhart „Learning representations by back-propagating errors“ mathematisch erläutert werden.<sup>25</sup> Es wird angenommen, dass es sich bei der Inputvariablen um einen Vektor aus  $n$  Attributen handelt. Für die Input-Layer ist dies üblicherweise eine gewisse Anzahl verschiedener Ratios.

$x_j = (x_1, x_2, \dots, x_n)$  ... Gesamtinput für Node  $j$

$y_i$  ... Output von Node  $i$

$w_{ji}$  ... Verknüpfungsgewichtung zwischen den Nodes

Der gesamte Input  $x_j$  für einen Knotenpunkt  $j$  lässt sich als lineare Funktion der Outputs  $y_i$  der Knotenpunkte, die mit Node  $j$  verknüpft und mit  $w_{ji}$  gewichtet sind, darstellen als:

$$x_j = \sum_i y_i w_{ji}$$

Dieses Ergebnis wird sodann mittels einer geeigneten Transferfunktion (auch Aktivierungsfunktion) an die nächsthöhere Ebene weitergegeben. Regelmäßig wird hierzu eine Sigmoidfunktion implementiert, welche wie folgt lautet:

---

<sup>24</sup> vgl. Zhang et al. (1997), S. 17-18

<sup>25</sup> vgl. Rumelhart et al. (1986), S.533-535

$$y_j = \frac{1}{1 + e^{-x_j}}$$

Hintergrund für die Verwendung eben dieser Formel ist, dass die Sigmoidfunktion jeden Inputwert  $x_j$  aus dem Intervall  $(-\infty, \infty)$  in einen im Intervall  $(0, 1)$  liegenden Wert umwandelt, wobei bei einem Wert  $x_j$  von genau 0 der Wendepunkt der Sigmoidfunktion mit einem Output von  $y_j = 0,5$  erreicht ist.<sup>26</sup> Diese Tatsache ist besonders in der Verarbeitung innerhalb der Output-Layer sinnvoll und erwünscht, da als Ergebnis einer Berechnung meist eine 1 oder 0 für die Einteilung in eine bestimmte Klasse stehen soll. Bemerkenswert ist hierbei noch, dass das Ergebnis lediglich bei relativ kleinen Ausprägungen von  $x_j$ , nämlich für das Intervall  $(-5, 5)$ , einen Wert von 0 bis 0,5 für  $x_j$   $(-5, 0)$  bzw. 0,5 bis 1 für  $x_j$   $(0, 5)$  annimmt. Diese, auch logistische Funktion genannte, Sigmoidfunktion führt aufgrund seines Ergebnisintervalls daher außerdem zu einer Normalisierung des Datensatzes.<sup>27</sup>

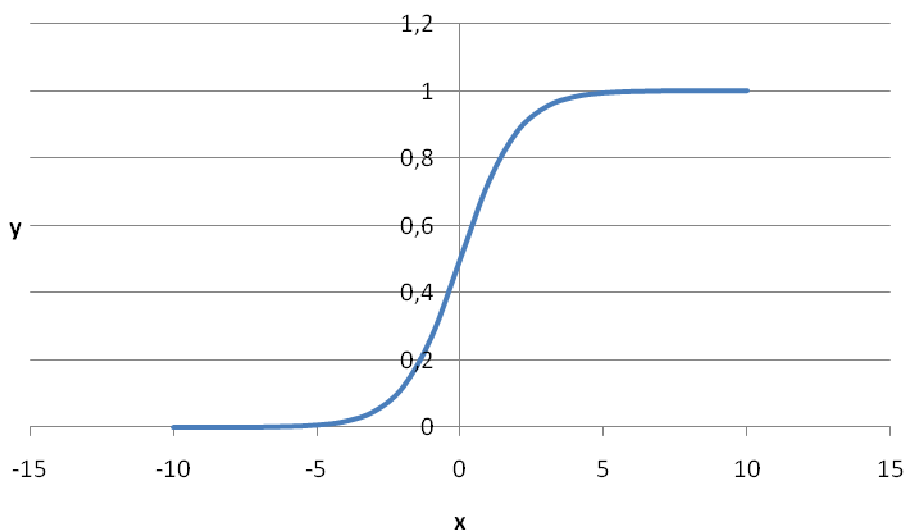


Abb. 2: Standard Sigmoidfunktion

### 3.2.2 Training

Das designierte Ziel für das eigentliche Training des ANN muss es nun sein, die Verknüpfungsgewichte  $w_{ji}$  so zu schätzen, dass der Output des ANN dem vorab bekannten Ergebnis des Klassifizierungsproblems entspricht, bzw. sich diesem zumindest stark annähert. Dazu bedarf es einem Maß, welches die Differenz zwischen tatsächlichem und gewünschtem Output misst. Dazu verwendet Rumelhart den Gesamtfehler  $E$ :

<sup>26</sup> vgl. Bishop (1995), S. 83

<sup>27</sup> vgl. Zhang et al. (1997b), S. 49

$$E = \frac{1}{2} \sum_c \sum_j (y_{j,c} - d_{j,c})$$

$y$  ... tatsächlicher Output

$d$  ... erwünschter Output

$c$  ... Set (Input-Output-Paar)

$j$  ... Output Node

Um  $E$  zu minimieren müssen die partiellen Ableitungen bezüglich jedes Gewichtungsfaktors berechnet werden. Dies geschieht über die Summierung der partiellen Ableitungen eines jeden Input-Output-Sets. Die partielle Ableitung von  $E$  nach  $y$  für einen bestimmten Fall  $c$  (der Index  $c$  wird ab hier nicht mehr mitgeführt) lautet:

$$\frac{\partial E}{\partial y_j} = y_j - d_j$$

Über die Kettenregel lässt sich nun  $\partial E / \partial x_j$  berechnen:

$$\frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} * \frac{dy_j}{dx_j}$$

Über die Ableitung der Sigmoidfunktion nach  $x_j$  (ausgedrückt als rekursive Funktion der Stammfunktion  $y_j$ ) kann man nun  $dy_j/dx_j$  ermitteln und einsetzen:

$$\left[ \frac{dy_j}{dx_j} = \frac{1}{1 + e^{-x_j}} * \left( 1 - \frac{1}{1 + e^{-x_j}} \right) = y_j(1 - y_j) \right]$$

$$\frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} * y_j(1 - y_j)$$

Diese Formel zeigt auf, wie sich der Gesamtinput  $x$  der Output-Node auf den Fehler  $E$  auswirkt. Um nun aber einen Bezug zu den Gewichtungsfaktoren herstellen zu können, bedarf es weiterhin einer Betrachtung eben dieser. Für den Gewichtungsfaktor  $w_{ji}$  von Knoten  $i$  zu Knoten  $j$  gestaltet sich die Ableitung der Fehlerfunktion wie folgt:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial x_j} * \frac{\partial x_j}{\partial w_{ji}} = \frac{\partial E}{\partial x_j} * y_i$$

und als Output der  $i$ -ten Node ist der Beitrag zu  $\partial E / \partial y_i$  der sich aus dem Einfluss von  $i$  auf  $j$  ergibt gerade:

$$\frac{\partial E}{\partial x_j} * \frac{\partial x_j}{\partial y_i} = \frac{\partial E}{\partial x_j} * w_{ji}$$

Betrachtet man also alle Verbindungen, die von Node  $i$  ausgehen, erhält man

$$\frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} * w_{ji}$$

Die eben aufgezeigten Zusammenhänge sollen verdeutlichen, wie man, ausgehend von  $\partial E / \partial y$  der letzten Layer,  $\partial E / \partial y$  für jeden Knotenpunkt der mittleren Schicht errechnen kann. Dieser Vorgang kann zur Berechnung der untersten Layer wiederholt werden, wobei parallel stets die Berechnung von  $\partial E / \partial w$  erfolgt. Bisher ist somit geklärt worden, wie sich die Zusammenhänge in einem ANN darstellen und wie einzelne Parameter in den verschiedenen Lagen ermittelt werden können. Es fehlt nun noch der eigentliche Optimierungsprozess für das ANN, nämlich die Justierung der Übertragungsgewichte. Ein Weg dabei wäre, die Gewichtungsfaktoren nach jedem Durchlauf eines Sets gemäß dem Gewichtungsfehler  $\partial E / \partial w$  anzupassen. Alternativ dazu kann jedoch auch  $\partial E / \partial w$  über alle  $c$  kumuliert werden um erst im Anschluss an den Durchlauf aller Sets eine Anpassung der Gewichtung vorzunehmen. Dabei wird eine Form des Gradientenabstiegsverfahrens angewandt, bei der die einzelnen Gewichtungsfaktoren proportional zum Gesamtfehler  $\partial E / \partial w$  geändert werden.

$$\Delta w = -\varepsilon \frac{\partial E}{\partial w}$$

Bei dieser Methode handelt es sich um die einfachste Möglichkeit der Gewichtungsfaktorenanpassung. Eine andere Methode, die wesentlich schneller abläuft, ist die, dass zusätzlich zu der einfachen Änderung der Gewichtungsfaktoren ein Momentum-Term eingefügt wird, der dazu führt, dass Minima in der Fehlerfunktion schneller erreicht werden. Die Formel lautet sodann folgendermaßen:

$$\Delta w(t) = -\varepsilon \frac{\partial E}{\partial w(t)} + \alpha \Delta w(t - 1)$$

mit

$\varepsilon$  ... Lernrate

$t$  ... Steigerungsparameter, der mit jedem (kompletten) Durchlauf um 1 erhöht wird

$\alpha$  ... Momentum

$\alpha$  liegt im Intervall  $(0,1)$  und bestimmt den Einfluss der vorangegangenen Iteration auf die aktuelle. Das so genannte Momentum dient dabei dem Zweck, die Gewichtungsfaktorenänderung in unregelmäßigen Bereichen der Fehlerfunktion abzubremesen sowie in langen, flachen Bereichen eine Beschleunigung herbeizuführen. Rasche Richtungswechsel in der Gewichtungsfaktorenjustierung werden dadurch vermieden.<sup>28</sup> Die Lernrate  $\epsilon$  bestimmt die Geschwindigkeit der Veränderung der Gewichtungen und ist negativ, da sich die Gewichte entgegengesetzt dem Fehler entwickeln sollen, um selbigen zu minimieren. Denkbar bei der Implementierung von  $\epsilon$  ist auch eine adaptive Herangehensweise in dem Sinne, dass  $\epsilon$  zunächst etwas größer gewählt wird, um sich dem Minimum rasch zu nähern und nach einer gewissen Anzahl von Trainingsläufen verringert wird, um das Minimum exakt zu lokalisieren. Mittels dieser Herangehensweise lassen sich die fehlerminimierenden Gewichte  $w$  ermitteln.

### 3.2.3 Testing

Sobald ein MLP vollständig konstruiert und trainiert ist, kann es theoretisch für die erwünschten Prognosezwecke eingesetzt werden. Allerdings läuft man Gefahr, dass das Netzwerk Fehler aufweist und nicht so funktioniert, wie ursprünglich erhofft. Um dem vorzubeugen, ist es daher ratsam, jedes ANN vor tatsächlicher Inbetriebnahme einer Test-Phase zu unterziehen. Dabei soll in dieser Arbeit lediglich auf Tests bezüglich der Prognosegenauigkeit eingegangen werden. Eingabeinterface-, Funktionalitäts- sowie Strukturtests sollten jedoch nach der Programmierung eines ANN ebenfalls durchgeführt werden. Typischerweise erfolgt das Testing mittels eines kleinen Datensets, von dem – ähnlich dem Training-Set – die gewünschten Ergebnisse vorab bekannt sind. Wichtig ist hierbei jedoch, dass es sich bei den Test-Samples nicht um einen Teil des Training-Samples handelt, da ja eben auf dessen Basis das Netzwerk konzipiert wurde. Typischerweise werden von der Gesamtheit der verfügbaren Beispieldaten 20-25% für die Testphase verwendet.<sup>29</sup> Tendenziell ist davon auszugehen, dass die Prognosegenauigkeit bei der Betrachtung des Test-Samples im Vergleich zu den für das Training verwendeten Beispielen abnimmt, was jedoch nicht immer der Fall sein muss. Zumeist wird zudem noch eine Klassifikation in Fehler vom Typ I und Typ II vorgenommen, da die Folgen einer Falschklassifikation sich je nachdem, von welchem Typus sie sind, in ihren (bei Bankrottprognose insbesondere finanziellen) Auswirkungen stark unterscheiden können. Folgende Tabelle soll beispielhaft die Unterschiede zwischen Test- und Trainings-Sample-Ergebnissen darstellen (1180 Datensätze; MLP; 27 Input-Nodes; 12 Hidden-Nodes; 1 Output-Node; adaptives  $\epsilon$  von 0,3 - 0,01;  $\alpha = 0,3$ ; 1868 Trainingsdurchläufe):

#### Fehlerquote des ANN

---

<sup>28</sup> vgl. Moreira, Fiesler (1995), S. 2

<sup>29</sup> vgl. Pullum et al. (2007), S.22

Training-Sample			Testing-Sample		
Gesamt	Typ I	Typ II	Gesamt	Typ I	Typ II
10,805%	4,025%	17,585%	12,712%	7,627%	17,797%

Abb. 3: Unterschiede Training- und Testing-Sample<sup>30</sup>

### 3.3 Empirische Validität von ANNs<sup>31</sup>

Typischerweise wird bei der Anwendung von ANNs eine Auswahl verschiedener Kennzahlen als Input verwendet. Je nachdem welche Ratios als Ausgangsbasis verwendet werden, kann sich die Aussagekraft eines Neuronalen Netzes ändern. Atiya untersucht in seiner Arbeit bspw. die Prognosegenauigkeit eines MLP unter Verwendung von Finanzkennzahlen sowie einem Mix aus Finanz- und (aktienbasierten) Eigenkapitalkennzahlen sowie die Zuverlässigkeit eines ANN in verschiedenen Zeitspannen vor dem Bankrott. Als Finanzkennzahlen verwendet er hierbei Book Value / Total Assets, Cashflow / Total Assets, Rate of Change of Cashflow / Share, Gross Operating Income / Total Assets sowie den Return on Assets. Die gemischten Ratios aus Finanz- und Eigenkapitalkennzahlen sind Book Value / Total Assets, Cashflow / Total Assets, Price / Cashflow Ratio, Rate of Change of Stock Price, Rate of Change of Cashflow / Share und Stock Price Volatility. In den Ausgangsdaten befinden sich Angaben zu den Firmen im Zeitraum von einem Monat bis zu drei Jahren vor der Pleite. In der Summe handelt es sich bei der Datenbasis um 716 solvente und 195 insolvente Firmen, wobei für jede insolvente Unternehmung für zwei bis drei Zeiträume vor Bankrott die jeweiligen Kennzahlen ermittelt wurden, was dazu führt, dass für die illiquiden Firmen insgesamt 444 Datensätze verfügbar waren, was zu einer Datenbasis von 1160 Sets führt. Von dieser Gesamtheit wurden 491 Datensätze für das Training verwendet („in Sample“) und die restlichen 669 („out of Sample“) dienen der Verifizierung des Netzwerkes. Die folgenden beiden Tabellen sollen nun verdeutlichen, wie exakt ein ANN in der Lage ist, Bankrott zu prognostizieren.

Time to Default	# Correct (in Sample)	# in Sample	% Correct (in Sample)	# Correct (out of Sample)	# out of Sample	% Correct (out of Sample)
6 mnth or less	34	38	89.47	59	65	90.77
6 to 12 mnth	47	51	92.16	47	54	87.04
12 to 18 mnth	31	37	83.78	51	63	80.95
18 to 24 mnth	31	37	83.78	21	32	65.63
more than 24 mnth	14	25	56.00	24	42	57.14
Total Defaulted	157	188	83.51	202	256	78.91
Solvent	258	303	85.15	343	413	83.05
Total	415	491	84.52	545	669	81.46

Abb. 4: Ergebnisse unter Verwendung von Finanzkennzahlen<sup>32</sup>

Time to Default	# Correct (in Sample)	# in Sample	% Correct (in Sample)	# Correct (out of Sample)	# out of Sample	% Correct (out of Sample)
6 mnth or less	35	38	92.11	56	65	86.15
6 to 12 mnth	43	51	84.31	44	54	81.48
12 to 18 mnth	33	37	89.19	47	63	74.60
18 to 24 mnth	33	37	89.19	25	32	78.13
more than 24 mnth	19	25	76.00	28	42	66.67
Total Defaulted	163	188	86.70	200	256	78.13
Solvent	276	303	91.09	372	413	90.07
Total	439	491	89.41	572	669	85.50

Abb. 5: Ergebnisse unter Verwendung von Finanz- und Eigenkapitalkennzahlen<sup>33</sup>

Diese Gegenüberstellung zeigt relativ eindeutig, dass nicht nur die korrekte Implementierung des MLP wichtig ist, sondern dass es auch auf die Art der Datenbasis ankommt, wie sicher die Prognose eines ANN letztendlich ist. Die hier zitierte Studie macht deutlich, wie sinnvoll es ist, sich nicht nur auf Daten aus der Buchhaltung zu stützen, wenn Bankrottprognosen betrieben werden sollen. Das MLP, welches mit dem gemischten Datensatz ausgestattet wurde, erzielt eine Gesamtperformance von 85,5%, was im Vergleich zu dem finanzkennzahlenbasierten ANN mit 81,46% rund 4 Prozentpunkte mehr ist. Gerade auch in der langen Frist wird der Vorteil der gemischten Datenbasis deutlich (bspw. 18-24 Monate vor Bankrott: 65,63% vs. 78,13% Trefferquote). Die höhere Genauigkeit bei dem gemischten Datensatz wird von Atiya auf die schnellere Verarbeitung von Informationen sowie das Einfließen von Erwartungshaltungen über bestimmte Unternehmen auf den Finanzmärkten zurückgeführt. Andere Studien zeigen auf, wie mittels verschiedener Methoden die Ausgangsdaten verbessert werden können, um dadurch eine noch höhere Genauigkeit zu erzielen. Die dabei wohl bekanntesten Methoden sind Boosting und Bagging. Unter Bagging versteht man eine Aggregationsmethode, bei der verschiedene Klassifikatoren kombiniert und kreiert werden. Beim Boosting wird ein zusammengesetzter Klassifikator konstruiert. Dies geschieht über ein sequentielles Training der Datenbasis, bei der die Gewichtung der einzelnen Klassifikatoren gemäß ihrem Beitrag zur Falschklassifikation sukzessive erhöht wird.<sup>34</sup> Auf die genaue Funktionsweise dieser Methoden wird hier jedoch nicht näher eingegangen. Es soll damit viel-

<sup>32</sup> Atiya (2001), S. 933

<sup>33</sup> Atiya (2001), S. 933

<sup>34</sup> vgl. Kim et al. (2009), S. 3374



mehr aufgezeigt werden, dass die Thematik der künstlichen Neuronalen Netzwerke Gegenstand zahlreicher Forschungsarbeiten ist, die sich mit der Optimierung eben dieser Netze befassen, denn je nach angewandter Methode können die Prognosefehler noch stärker reduziert werden.<sup>35</sup>

#### 4. Fazit

Die vorliegende Arbeit hat die Problematik der Bankrottprognose mittels verschiedener Modelle aufgezeigt. Dabei ist deutlich geworden, dass die Komplexität der Herangehensweise der einzelnen Autoren im Zeitablauf stark zugenommen hat, was jedoch im Umkehrschluss auch dazu führte, dass die Modelle genauer, allgemeingültiger und vor allem weniger stör anfällig geworden sind. Erste Ansätze entwickelte Beaver 1966, indem er mit Hilfe einer einzigen Kennzahl eine ex ante Klassifikation von Unternehmen in künftig solvent respektive insolvent vorzunehmen versucht. Mit einer Zuverlässigkeit von bis zu 87% gelang es ihm dabei, zumindest die Unternehmen korrekt zu klassifizieren, welche Teil seiner Ausgangsdatenbasis waren. Da jedoch parallel dazu auch zahlreiche Arbeiten verfasst wurden, die jeweils eine andere Ratio als diejenige identifizierten, die sich am besten für Bankrottprognosezwecke eignet, war es alsbald an der Zeit, allgemeingültigere Herangehensweisen zu entwickeln. Ein Werk, welches bis heute eine gewisse Relevanz hat, ist die Arbeit von Altman. Dieser hat statt einem, fünf Ratios ermittelt, die unterschiedlich gewichtet und addiert wurden. Das Ergebnis dieser Kalkulation ist als Z-Score bekannt und gibt dem Anwender Aufschluss darüber, ob ein Unternehmen zukünftig zahlungsfähig sein wird oder nicht. Obwohl der Z-Score noch lange Zeit als das zuverlässigste Prognose-Tool galt, war auch klar, dass dieses Maß zwar als Indikator zu dienen vermochte, eine relativ sichere Vorhersage jedoch nicht zu erwarten war, da insbesondere die Gewichtung der einzelnen Ratios für jede Datenbasis hätte angepasst werden müssen, was lediglich a posteriori möglich gewesen wäre. Um dieses Problem zu umgehen, war es notwendig, eine Methode zu entwickeln, die für eine große Gesamtheit von vorhandenen Daten Gültigkeit besaß, was jedoch erst mit Hilfe von Computern möglich wurde. Den Grundstein für Konstruktionen aus dem Bereich künstlicher, intelligenter Modelle legte Rumelhart 1986, indem er einen Algorithmus für künstliche Neuronale Netzwerke erschuf, mit dem es möglich war, mittels Fehlerrückführung (Backpropagation) durch das System den Schätzfehler zu minimieren. Dank dieser Pionierarbeit auf dem Gebiet der künstlichen Intelligenz sowie zahlreichen Folgeuntersuchungen, ist es uns heute möglich, a priori eine relativ genaue Prognose über die Entwicklung von Unternehmen bezüglich ihrer Zahlungsfähigkeit abzugeben. Um ein absolut vollkommenes Sys-

---

<sup>35</sup> vgl. Kim et al. (2009), S. 3377

tem handelt es sich dabei jedoch trotzdem nicht, da stets ein Restfehler verbleibt. An der Minimierung dieses Fehlers wird jedoch rege geforscht, was zur Folge hat, dass die ANNs immer ausgefallener, besser und außerdem zunehmend komplexer werden. Besonders das Gebiet der ANNs bietet somit noch einiges Forschungspotenzial, weshalb in den nächsten Jahren wohl mit einigen Erweiterungen, Verbesserungen und Optimierungen zu rechnen ist. „The application of neural networks has been reported in many recent studies of bankruptcy prediction. However, the mechanism of neural networks in predicting bankruptcy or in general classification is not well understood. Without a clear understanding of how neural networks operate, it will be difficult to reap full potentials of this technique.“<sup>36</sup>

---

<sup>36</sup> Zhang et al. (1997), S. 28

I. Abbildungen

*List of Ratios Tested<sup>a</sup>*

<p><b>GROUP I (CASH-FLOW RATIOS)</b></p> <ol style="list-style-type: none"> <li>1. Cash flow to sales</li> <li>2. Cash flow to total assets</li> <li>3. Cash flow to net worth</li> <li>4. Cash flow to total debt</li> </ol> <p><b>GROUP II (NET-INCOME RATIOS)</b></p> <ol style="list-style-type: none"> <li>1. Net income to sales</li> <li>2. Net income to total assets</li> <li>3. Net income to net worth</li> <li>4. Net income to total debt</li> </ol> <p><b>GROUP III (DEBT TO TOTAL-ASSET RATIOS)</b></p> <ol style="list-style-type: none"> <li>1. Current liabilities to total assets</li> <li>2. Long-term liabilities to total assets</li> <li>3. Current plus long-term liabilities to total assets</li> <li>4. Current plus long-term plus preferred stock to total assets</li> </ol> <p><b>GROUP IV (LIQUID-ASSET TO TOTAL-ASSET RATIOS)</b></p> <ol style="list-style-type: none"> <li>1. Cash to total assets</li> <li>2. Quick assets to total assets</li> <li>3. Current assets to total assets</li> <li>4. Working capital to total assets</li> </ol>	<p><b>GROUP V (LIQUID-ASSET TO CURRENT DEBT RATIOS)</b></p> <ol style="list-style-type: none"> <li>1. Cash to current liabilities</li> <li>2. Quick assets to current liabilities</li> <li>3. Current ratio (current assets to current liabilities)</li> </ol> <p><b>GROUP VI (TURNOVER RATIOS)</b></p> <ol style="list-style-type: none"> <li>1. Cash to sales</li> <li>2. Accounts receivable to sales</li> <li>3. Inventory to sales</li> <li>4. Quick assets to sales</li> <li>5. Current assets to sales</li> <li>6. Working capital to sales</li> <li>7. Net worth to sales</li> <li>8. Total assets to sales</li> <li>9. Cash interval (cash to fund expenditures for operations)</li> <li>10. Defensive interval (defensive assets to fund expenditures for operations)</li> <li>11. No-credit interval (defensive assets minus current liabilities to fund expenditures for operations)</li> </ol>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Abb. 6: 30 Ratios von Beaver<sup>37</sup>

**FIVE YEAR PREDICTIVE ACCURACY OF THE MDA MODEL  
(Initial Sample)**

Year Prior to Bankruptcy	Hits	Misses	Per cent Correct
1st n = 33	31	2	95
2nd n = 32	23	9	72
3rd n = 29	14	15	48
4th n = 28	8	20	29
5th n = 25	9	16	36

Abb. 7: Ergebnisse der Altman-Studie<sup>38</sup>

<sup>37</sup> Beaver (1966), S. 78

<sup>38</sup> Altman (1968), S. 604

<b>Bankrupted Companies</b>				
<b>Year</b>	<b>All</b>	<b>Correctly Classified</b>	<b>Type 2 error</b>	<b>% correct</b>
-4	44	9	35	20%
-3	44	17	27	39%
-2	44	23	21	52%
-1	44	29	15	66%

Abb. 8: Prognosegenauigkeit für bankrotte Firmen<sup>39</sup>

<b>Non Bankrupted Companies</b>				
<b>Year</b>	<b>All</b>	<b>Correctly Classified</b>	<b>Type 2 error</b>	<b>% correct</b>
-4	244	190	54	78%
-3	215	149	66	69%
-2	193	121	72	63%
-1	168	91	77	54%

Abb. 9: Prognosegenauigkeit für nicht-bankrotte Firmen<sup>40</sup>

<b>All Cases</b>				
<b>Year</b>	<b>All</b>	<b>Correctly Classified</b>	<b>Type 2 error</b>	<b>% correct</b>
-4	288	199	89	69%
-3	259	166	93	64%
-2	237	144	93	61%
-1	212	120	92	57%

Abb. 10: Prognosegenauigkeit für alle Firmen<sup>41</sup>

<sup>39</sup> Gerantonis et al. (2009), S. 24

<sup>40</sup> Gerantonis et al. (2009), S. 25

<sup>41</sup> Gerantonis et al. (2009), S. 25

## II. Literaturverzeichnis

*Alfaro, Esteban / García, Noelia / Gámez, Matías / Elizondo, David (2007): Bankruptcy Forecasting: An Empirical Comparison of AdaBoost and Neural Networks, in: Decision Support Systems, Vol. 45 (2008)*

*Altman, Edward I. (1968): Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, in: The Journal of Finance, Vol. 23, No. 4 (September, 1968)*

*Altman, Edward I. (2000): Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA® Models*

*Amemya, Takeshi (1981): Qualitative Response Models: A Survey, in: Journal of Economic Literature, Vol. 19, No. 4 (Dec., 1981)*

*Atiya, Amir F. (2001): Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Result, in: IEEE Transactions on Neural Networks, Vol. 12, No. 4 (Juli, 2001)*

*Beaver, William H. (1966): Financial Ratios as Predictors of Failure, in: Journal of Accounting Research, Vol. 4, Empirical Research in Accounting: Selected Studies 1966 (1966)*

*Beaver, William H. (1968): Market Prices, Financial Ratios, and the Prediction of Failure, in: Journal of Accounting Research, Vol. 6, No. 2 (Herbst, 1968)*

*Bishop, Christopher M. (1995): Neural Networks for Pattern Recognition, Oxford University Press (1995)*

*Gerantonis, Nikolas / Vergos, Konstantinos / Christopholus, Apostolos G. (2009): Can Altman Z-score Models Predict Business Failures in Greece? in: Research Journal of International Studies, Vol. 12 (Oktober, 2009)*

*Gibbons, Jean Dickinson (1993): Nonparametric Statistics: An Introduction, in: Sage University Papers Series, Quantitative Applications in the Social Sciences*

*Grice*, John Stephen / *Ingram*, Robert W. (2001): Tests of the Generalizability of Altman's Bankruptcy Prediction Model, in: *Journal of Business Research*, Vol. 54 (2001)

*Kim*, Myoung-Jong / *Kang*, Dae-Ki (2009): Ensemble with Neural Networks for Bankruptcy Prediction, in: *Expert Systems with Applications*, Vol. 37 (2010)

*Moreira*, M. / *Fiesler*, E. (1995): Neural Networks with Adaptive Learning Rate and Momentum Terms, in *Institut Dalle Molle D'Intelligence Artificielle Perceptive – Technical Report No. 95-04* (1995)

*Odom*, M. / *Sharda*, R. (1990): A Neural Network Model for Bankruptcy Prediction, in: *International Joint Conference on Neural Networks*, Vol. 2 (1990)

*Pullum*, Laura L. / *Taylor*, Brian J. / *Darrach*, Marjorie A. (2007): *Guidance for the Verification and Validation of Neural Networks*, Jon Wiley & Sons (2007)

*Rumelhart*, David E. / *Hinton*, Geoffrey E. / *Williams*, Ronald J. (1986): Learning Representations by Back-Propagating Errors, in: *Nature*, Vol. 323 (Oktober, 1986)

*Udo*, Godwin (1993): Neural Network Performance on the Bankruptcy Classification Problem, in: *Computers and Industrial Engineering* Vol. 25, No. 1-4 (1993)

*Zhang*, Guoqiang / *Hu*, Michael Y. / *Patuwo*, B. Eddy / *Indro*, Daniel C. (1997): Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-validation Analysis, in: *European Journal of Operational Research* Vol. 116 (1999)

*Zhang*, Guoqiang / *Hu*, Michael Y. / *Patuwo*, B. Eddy (1997b): Forecasting with Artificial Neural Networks: The State of the Art, in: *International Journal of Forecasting* Vol. 14 (1998)