

2 Methodisch-konzeptionelle Grundlagen zu den eingesetzten Arbeitstechniken und Hypothesen zu Übernahmen und Fusionen

2.1 Arbeitstechniken der Forschungsarbeit

2.1.1 Grundlegende hypothesenprüfende Techniken der multivariaten Analyse

2.1.1.1 Einführung in die Techniken der hypothesenprüfenden multivariaten Analyse

Eine Zweiteilung der Techniken der multivariaten Analyse in grundlegende und komplexe Verfahren ist nicht eindeutig, kann aber vorgenommen werden.¹ Zu den grundlegenden hypothesenprüfenden Techniken der multivariaten Analyse zählen:²

- Regressionsanalyse
- Zeitreihenanalyse
- Varianzanalyse
- Diskriminanzanalyse
- logistische Regression
- Kreuztabellierung und Kontingenzanalysen

¹ Vgl. BACKHAUS et al. (2011), S. 12.

² Vgl. BACKHAUS et al. (2011), S. 12.

Abhängig von dem Skalenniveau¹ der unabhängigen und der abhängigen Variablen² ist der Einsatz verschiedener multivariater Analysemethoden zur Überprüfung von Zusammenhängen möglich. Tabelle 2 zeigt eine Übersicht über die grundlegenden hypothesenprüfenden Techniken.³ Diese Klassifikation in hypothesenprüfenden Techniken kann keinen Anspruch auf Allgemeingültigkeit erheben, sondern nur die vorwiegenden Einsatzbereiche der Methoden kennzeichnen.⁴

¹ Je nach Art und Weise der Messbarkeit einer Variablen werden verschiedene Skalenniveaus unterschieden. Der Informationsgehalt und die Anwendbarkeit von Rechenoperationen sind abhängig vom Skalenniveau. Grundlegend kann man zwischen nicht-metrischen Skalen und metrischen Skalen unterscheiden. Zu den nicht-metrischen Skalen gehören die Nominalskala, welche nur eine qualitative Klassifikation von Eigenschaftsausprägungen vornimmt, und die Ordinalskala, welche die Aufstellung einer Rangordnung mithilfe von Rangwerten erlaubt. Somit stellt die Ordinalskala das nächst höhere Skalenniveau zur Nominalskala dar. Die metrischen Skalen können in die Intervallskala, die zusätzlich zu den Eigenschaften der nicht-metrischen Skalen einen Informationsgehalt in der Differenz zwischen den Ausprägung der Variablen enthält, und in die Ratioskala, welche im Vergleich zur Intervallskala zusätzlich einen Nullpunkt enthält, unterteilt werden. Vgl. BACKHAUS et al. (2008), S. 8 ff.; AUER et al. (2011), S. 6 ff.

² Die Differenzierung in unabhängige und abhängige Variablen erfolgt durch Regressionsanalysen. Hierfür ist die A-priori-Formulierung theoretisch fundierter Kausalzusammenhänge notwendig. Eine Einführung erfolgte im Rahmen der Kurzdarstellung der Hypothesen zu Übernahmen und Fusionen in Kapitel 1. Eine Detaillierung erfolgt in Kapitel 2.2 der vorliegenden Arbeit.

³ Die vorliegende Arbeit beschränkt sich auf die hypothesenprüfenden Techniken, da die kausalen Abhängigkeiten zwischen abhängigen Variablen und unabhängigen Variablen untersucht werden sollen. Über die Zusammenhänge existieren theoretische Überlegungen. Eine Betrachtung der strukturierten Verfahren ist daher nicht notwendig.

⁴ Vgl. BACKHAUS et al. (2008), S. 19.

		unabhängige Variable	
		metrisches Skalenniveau	nominales Skalenniveau
abhängige Variable	metrisches Skalenniveau	Regressionsanalyse	Varianzanalyse
		Zeitreihenanalyse	Regression mit Dummies
	nominales Skalenniveau	Diskriminanzanalyse	Kontingenzanalyse
		logistische Regression	

Tabelle 2: Grundlegende hypothesenprüfende Techniken der multivariate Analysemethoden¹

Die unabhängigen Variablen, die als Finanzkennzahlen in Kapitel 1 eingeführt wurden und in Kapitel 2.2 in den Tabellen 4-7 der vorliegenden Arbeit konkret dargestellt werden, sind metrisch skaliert. Die abhängige Variable, die unternehmensspezifische Übernahme- und Fusionswahrscheinlichkeit von potenziellen Akquisezielen, ist nominal skaliert. Durch diese Vorgabe für die Skalenniveaus der Variablen² beschränkt sich die Anwendbarkeit der grundlegenden hypothesenprüfenden Techniken der multivariaten Analysemethoden auf die vorliegende Arbeit.

Grundlegende hypothesenprüfende multivariaten Analysemethoden, die gemäß Tabelle 2 in der vorliegenden Arbeit Anwendung finden könnten, sind:³

1. die Diskriminanzanalyse und
2. die logistische Regression.

Bedingt durch die geringe wissenschaftliche Bedeutung der Diskriminanzanalyse bei der Entwicklung von Prognosemodellen, die eine systematische und kontextbezogene Übernahme- und Fusionswahrscheinlichkeit von Akquisezielen mithilfe von hypothesenbasierten Finanzkennzahlen ermöglichen (vgl. Tabelle 1), wird diese in der vorliegenden Forschungsarbeit nicht berücksichtigt.

¹ Vgl. BACKHAUS et al. (2008), S. 12.

² Wenn keine Differenzierung zwischen unabhängigen und abhängigen Variablen erfolgt, sind beide Arten von Variablen gemeint.

³ Die Notation und die Aufbaustruktur orientieren sich in diesem Kapitel an BACKHAUS et al. (2011a), S. 56 ff.

2.1.1.2 Vorgehensmodell zur logistischen Regression

Die logistische Regression gehört zu den in Kapitel 2.1.1.1 vorgestellten klassischen multivariaten Analysemethoden. Diese Methode hat das Ziel, die Einflussgrößen zu ermitteln, die Akquiseziele und Nicht-Akquiseziele am besten unterscheidet.¹ Im Vergleich zu der linearen Regressionsanalyse ist bei der logistischen Regressionsanalyse nicht nur die Eingruppierung sondern das Ableiten einer Eintrittswahrscheinlichkeit das Ziel.²

Die Vorgehensweise bei der logistischen Regressionsanalyse kann in fünf Schritte untergliedert werden.³

1. Modellformulierung
2. Schätzung der logistischen Regressionsfunktion
3. Interpretation der Regressionskoeffizienten
4. Prüfung des Gesamtmodells
5. Prüfung der Merkmalsvariablen

2.1.1.2.1 Modellformulierung

Bei der logistischen Regression stehen für die Modellformulierung die in Kapitel 1 eingeführten Hypothesen zu Übernahmen und Fusionen mit den dazugehörigen Finanzkennzahlen im Vordergrund. Allerdings werden keine Je-desto-Hypothesen unmittelbar zwischen den unabhängigen und abhängigen Variablen formuliert, sondern zwischen den unabhängigen Variablen und der Eintrittswahrscheinlichkeit für das Ereignis Akquiseziel = 1.⁴ Die Wirkungsbeziehungen, die in den Hypothesen aufgestellt worden sind, haben keinen linearen Charakter, denn durch die logistische Funktion wird eine nicht-lineare, um genauer zu sein eine S-förmige, Wahrscheinlichkeitsverteilung angenommen.⁵ Bei der Betrachtung

¹ Vgl. SCHLITTGEN (2009), S. 203.

² Vgl. BACKHAUS et al. (2011a), S. 251.

³ Vgl. KRAFFT (1999), S. 249; BACKHAUS et al. (2011a), S. 250 ff.

⁴ Vgl. SCHLITTGEN (2009), S. 203 ff.

⁵ Vgl. KRAFFT (1999), S. 243.

einer abhängigen Variable mit nur zwei Ausprägungen wird eine binär-logistische Regression durchgeführt.

2.1.1.2.2 Schätzung im Rahmen der logistischen Regression

Zur Schätzung der Modellparameter wird die Maximum-Likelihood-Methode mit dem Ziel angewandt, die Regressionskoeffizienten so zu bestimmen, dass die Wahrscheinlichkeit, die Erhebungsdaten zu erhalten, maximiert wird.¹ Um ein betrachtetes Unternehmen k als Akquizeziel oder Nicht-Akquizeziel zu klassifizieren, sollte das Ergebnis der Parameterschätzung entweder die Wahrscheinlichkeit $p_k (y_k = 1)$ für Akquizeziel oder $p_k (y_k = 0)$ für Nicht-Akquizeziel erbringen. Dies wird wie folgt formuliert:²

$$y_k = \begin{cases} 1 & \text{falls } z_k > 0 \\ 0 & \text{falls } z_k \leq 0 \end{cases} \quad (1.1)$$

wobei:³

$$z_k = b_0 + \sum_{j=1}^J b_h * x_{hk} + u_k \quad (1.2)$$

mit

b_0 = konstantes Glied ($b_0 \in \mathbb{R}$)

b_h = Regressionskoeffizienten für die unabhängige Variable h ($b_h \in \mathbb{R} \forall h = 1, 2, \dots, H$)

x_{hk} = Ausprägung der unabhängigen Variable h für das Unternehmen k ($x_{hk} \in \mathbb{R} \forall h = 1, 2, \dots, H; k = 1, 2, \dots, K$)

y_k = Ausprägung der abhängigen Variable für das Unternehmen k ($k = 1, 2, \dots, K$)

u_k = Residualgröße für das Unternehmen k ($u_k \in \mathbb{R} \forall k = 1, 2, \dots, K$)

z_k = Logit für das Unternehmen k ($z_k \in \mathbb{R} \forall k = 1, 2, \dots, K$)

H = Anzahl der unabhängigen Variablen

K = Anzahl der Unternehmen

¹ Vgl. BEST et al. (2010), S. 836.

² Vgl. BEST et al. (2010), S. 834 f.; BACKHAUS et al. (2011a), S. 254 f.

³ Für die Bestimmung der Wahrscheinlichkeit p für $y_k = 1$ oder $y_k = 0$ wird unterstellt, dass eine nicht empirisch beobachtbare Variable z_k (auch Logit genannt) die binäre Ausprägung von y_k annehmen kann. Vgl. BEST et al. (2010), S. 834 f.; BACKHAUS et al. (2011a), S. 254 f.

Die Variable z_k , die durch die Linearkombination der unabhängigen Variablen x_{hk} erzeugt wird, stellt die Verbindung zwischen der binären abhängigen Variable y_k und den unabhängigen Variablen dar.¹ Um eine Wahrscheinlichkeitsaussage treffen zu können, wird in der logistischen Regression die logistische Funktion verwendet.²

$$p(y_k) = \frac{e^{z_k}}{1 + e^{z_k}} \text{ bzw. } p(y_k) = \frac{1}{1 + e^{-z_k}} \quad (1.3)$$

mit

$p(y_k)$ = Eintrittswahrscheinlichkeit für das Ereignis y_k ($p \in \mathbb{R} \mid 0 \leq p \leq 1$)

$$\forall y_k = \begin{cases} 1 & \text{falls } z_k > 0 \\ 0 & \text{falls } z_k \leq 0 \end{cases}$$

e = 2,71828183 (Eulersche Zahl)

z_k = Logit für das Unternehmen k ($z_k \in \mathbb{R} \forall k = 1, 2, \dots, K$)

Durch die Kombination der Formeln 1.2 und 1.3 entsteht die logistische Regressionsgleichung.³

$$p(y_k) = \left(\frac{1}{1 + e^{-z_k}} \right)^{y_k} * \left(1 - \frac{1}{1 + e^{-z_k}} \right)^{1-y_k}, \text{ für } y_k = \begin{cases} 1 & \text{falls } z_k > 0 \\ 0 & \text{falls } z_k \leq 0 \end{cases} \quad (1.4)$$

mit

$p(y_k)$ = Eintrittswahrscheinlichkeit für das Ereignis y_k ($p \in \mathbb{R} \mid 0 \leq p \leq 1 \forall y_k$)

e = 2,71828183 (Eulersche Zahl)

y_k = Ausprägung der abhängigen Variable ($k = 1, 2, \dots, K$)

z_k = Logit für das Unternehmen k ($z_k \in \mathbb{R} \forall k = 1, 2, \dots, K$)

K = Zahl der Unternehmen

Die logistische Regressionsgleichung ermittelt den Wahrscheinlichkeitswert p für das Eintreten des Ereignisses $y_k = 1$. Bei der Klassifikation eines Unternehmens als Akquiseziel

¹ Vgl. BEST et al. (2010), S. 834 f.; BACKHAUS et al. (2011a), S. 255.

² Vgl. BEST et al. (2010), S. 835; BACKHAUS et al. (2011a), S. 255.

³ Da y_k nur die Werte 0 und 1 annehmen kann, ist immer nur ein Teil der Formel relevant. Bei $y_k = 1$ ist die erste Hälfte des rechten Terms und bei $y_k = 0$ die zweite Hälfte relevant. Die Funktion kann Werte zwischen 0 und 1 annehmen. Vgl. BEST et al. (2010), S. 835 ff.

oder Nicht-Akquiseziel ist daher eine Klassifikationsvorschrift festzulegen. I. d. R. wird der Wahrscheinlichkeitswert von 0,5 verwendet und für $p(y_k) \geq 0,5$ eine Klassifikation zum Ereignis $y_k = 1$ für Akquiseziel und für $p(y_k) < 0,5$ zu $y_k = 0$ für Nicht-Akquiseziel vorge-
nommen.¹ Der Wahrscheinlichkeitsansatz für unabhängige Ereignisse (Beobachtungswerte der abhängigen Variable) besagt, dass die Wahrscheinlichkeit des gleichzeitigen Eintretens der Ereignisse durch Multiplikation der Einzelereignisse ermittelt werden kann.² Der Wahrscheinlichkeitsansatz für unabhängige Ereignisse wird durch die folgende Likelihood-Funktion beschrieben und muss folglich maximiert werden:³

$$L = \prod_{k=1}^K \left(\frac{1}{1 + e^{-z_k}} \right)^{y_k} * \left(1 - \frac{1}{1 + e^{-z_k}} \right)^{1-y_k} \rightarrow \max! \quad (1.5)$$

Durch das Logarithmieren der Gleichung (1.5) wird eine Vereinfachung des Maximierungsproblems erreicht. Das Logarithmieren der Gleichung (1.5) führt z.B. zur Transformierung der Produkte in Summen ohne das Maximum der Funktion zu beeinflussen.⁴ Die sich hieraus ergebende LogLikelihood-Funktion sieht wie folgt aus:⁵

$$LL = \sum_{k=1}^K \left[y_k * \ln \left(\frac{1}{1 + e^{-z_k}} \right) \right] + \left[(1 - y_k) * \ln \left(1 - \frac{1}{1 + e^{-z_k}} \right) \right] \rightarrow \max \quad (1.6)$$

Zur Maximierung der LogLikelihood-Funktion werden in einem iterativen Verfahren Schätzwerte für die Regressionskoeffizienten b_h , auch Logit-Koeffizienten genannt, verwendet.⁶

Für die Iteration kann der Newton-Raphson-Algorithmus verwendet werden.¹

¹ Vgl. BACKHAUS et al. (2011a), S. 259. Definition für genau 0,5 erfolgt durch den Autor der vorliegenden Arbeit.

² Vgl. BACKHAUS et al. (2011a), S. 259.

³ Vgl. KRAFFT (1999), S. 242; BEST et al. (2010), S. 837; BACKHAUS et al. (2011a), S. 259.

⁴ Vgl. BEST et al. (2010), S. 837.

⁵ Vgl. CRAMER (2003), S. 34 f.; BEST et al. (2010), S. 837; BACKHAUS et al. (2011a), S. 259.

⁶ Vgl. BEST et al. (2010), S. 837. Die Regressionskoeffizienten b_h sind gemäß Formel 1.2 Bestandteile von $-z_k$.

2.1.1.2.3 Interpretation der Regressionskoeffizienten

Bei der Betrachtung der Likelihood-Funktion (1.5) und der LogLikelihood-Funktion (1.6) wird deutlich, dass die Interpretation der Regressionskoeffizienten (b_h) nicht ohne weiteres geschehen kann, auch wenn die Interpretation der Logits (z_k) analog zur Interpretation der OLS-Regression erfolgt.² Denn anders als bei der OLS-Regression werden die Logits (z_k) logarithmiert (siehe Funktionen 1.5 und 1.6). Zum einen nehmen die unabhängigen Variablen (b_h) als Exponenten der e-Funktion nur indirekt Einfluss auf die Eintrittswahrscheinlichkeit (p) für das Ereignis $y_k = 1$ und zum anderen ist der Einfluss durch die Anwendung der logistischen Funktion nicht linear.³ Eine Interpretationserleichterung kann durch die Betrachtung der *Odds* erreicht werden. Die *Odds* beschreiben das Wahrscheinlichkeitsverhältnis zwischen Akquiseziel und Nicht-Akquiseziel und können wie folgt ermittelt werden:⁴

$$Odds(y_k = 1) = \frac{p(y_k = 1)}{1 - p(y_k = 1)} \quad (1.7)$$

Die Formel 1.7 spiegelt die Wahrscheinlichkeit $p(y_k = 1)$ im Vergleich zum Wahrscheinlichkeit $p(y_k = 0)$ wider. Durch die Umformulierung der e-Funktion wird deutlich, dass die *Odds* sich entsprechend der e-Funktion mit den Exponenten z entwickeln:⁵

¹ Im ersten Schritt wird ein Schätzwert für die Logit-Koeffizienten angenommen. Im zweiten Schritt wird mit dem neu gewonnenen Logit-Koeffizienten die Wahrscheinlichkeit von $p(y_k = 1)$ bestimmt. Im Schritt drei wird dann der LogLikelihood-Wert bestimmt. Im Schritt vier werden die Schritte zwei und drei für alle Beobachtungsfälle durchgeführt. Das Ergebnis ist eine erste Gesamt-LogLikelihood-Funktion. Im Schritt fünf werden die Schritte zwei bis vier mit anderen Werten für b_h wiederholt. Die unterschiedlichen Gesamt-LogLikelihood-Funktionen werden anschließend verglichen und die Regressionskoeffizienten so lange verändert, bis keine Steigerung der Gesamt-LogLikelihood-Funktion möglich ist. Vgl. KRAFFT (1999), S. 242; BACKHAUS et al. (2011a), S. 260.

² Die Interpretation der OLS-Regression und der Logits erfolgt analog zueinander. Das konstante Glied b_0 gibt den y-Achsenabschnitt an und die Regressionskoeffizienten b_h geben die Steigung der Geraden an. Vgl. BEST et al. (2010), S. 831; BACKHAUS et al. (2011a), S. 262.

³ Vgl. BACKHAUS et al. (2011a), S. 263 f.

⁴ Vgl. BACKHAUS et al. (2011a), S. 264 f.; CRAMER (2003), S. 13.

⁵ Vgl. BEST et al. (2010), S. 831 f.; BACKHAUS et al. (2011a), S. 264 f.; CRAMER (2003), S. 13.

$$p(y_k = 1) = \frac{e^{z_k}}{1 + e^{z_k}} \quad (1.8)$$

$$\Rightarrow Odds : \frac{p(y_k = 1)}{1 - p(y_k = 1)} = e^{z_k} \quad (1.9)$$

2.1.1.2.4 Prüfung des Gesamtmodells

Die Prüfung des Gesamtmodells beinhaltet die folgenden zwei Teilbereiche:

1. Güte des Regressionsansatzes und
2. Ausreißerdiagnostik.

Bei der Beurteilung des ersten Teilbereiches der Güte des Regressionsansatzes insgesamt soll die Frage beantwortet werden, wie gut die unabhängigen Variablen in ihrer Gesamtheit zur Trennung der Ausprägungen von y_k (1 = Akquizeziel und 0 = Nicht-Akquizeziel) beitragen. Hierfür können folgende Gütekriterien betrachtet werden:¹

1. Gütekriterien auf der Basis der LogLikelihood-Funktion
2. Pseudo-R-Quadrat-Statistiken
3. Klassifikationsergebnisse

Zur Beurteilung der Güte des Regressionsansatzes mittels der LogLikelihood-Funktion können zwei Verfahren angewendet werden.²

Im ersten Verfahren wird die Devianz analysiert. Die Devianz kann inhaltlich mit der Fehlerquadratsumme der linearen Regressionsanalyse verglichen werden.³ Das minus Zweifache des logarithmierten Likelihoods $[-2LL$ (LL gemäß 1.6)] ergibt die Devianz, diese Größe folgt approximativ einer Chi-Quadrat-Verteilung mit $(K-J-1)$ Freiheitsgraden. Die Devianz kann zusätzlich als Gütemaß zur Überprüfung des Modellfits herangezogen werden. Bei einem Likelihood (L gemäß 1.5) von 1 [entspricht einer Devianz $(-2LL)$ von 0] wird ein perfekter Modellfit ausgewiesen.⁴ Ausgehend von der Devianz erfolgt der Hypothesen-

¹ Vgl. KRAFFT (1999), S. 244 ff.

² Vgl. BROSIUS (2013), S. 615 f.

³ Vgl. BACKHAUS et al. (2011a), S. 267.

⁴ Vgl. KRAFFT (1997), S. 245.

test. Hierfür wird eine Alternativhypothese formuliert, die aussagt, dass das Modell keine perfekte Anpassung besitzt. Weist die Devianz im Vergleich zu dem Chi-Quadrat-Wert einen geringeren Wert aus, kann die Alternativhypothese abgelehnt werden. Das bedeutet, dass die Nullhypothese, die eine perfekte Anpassung des Modells unterstellt, angenommen wird. Dadurch, dass keine nach Gruppen differenzierte Betrachtung erfolgt, hat die Analyse der Devianz den Nachteil, dass der Einfluss der Verteilungen der Beobachtungen auf die Ausprägungen von y_k nicht berücksichtigt wird und dies zu falschen Erkenntnissen führen kann.¹

Der Likelihood-Ratio-Test versucht dieses Problem zu beheben, indem alle Regressionskoeffizienten auf null gesetzt werden und die Devianz des konstanten Terms betrachtet wird (Nullmodell). Im Anschluss wird diese Devianz des Nullmodells mit der Devianz des Gesamtmodells verglichen. Ist die absolute Differenz zwischen den Devianzen klein, so ist der Beitrag der unabhängigen Variablen zur Unterscheidung der Ausprägungen von y_k gering.² Durch die Anwendung des Nullmodells wird der Effekt der Gruppengröße neutralisiert.³ Ausgehend von der absoluten Differenz zwischen dem Nullmodell und dem vollständigen Modell erfolgt der Hypothesentest. Ist die absolute Differenz größer als der Wert der Chi-Quadrat-Tabelle⁴, ist die Nullhypothese abzulehnen und die Alternativhypothese anzunehmen.

Die Pseudo-R-Quadrat-Statistiken sind inhaltlich vergleichbar mit dem Bestimmtheitsmaß R^2 der linearen Regressionsanalyse.⁵ Bei den Pseudo-R-Quadrat-Statistiken wird für die Beurteilung der Güte des Regressionsansatzes insgesamt auf das Verhältnis zwischen dem LogLikelihood (LL gemäß 1.6) des Nullmodells LL_0 und dem LogLikelihood (LL gemäß 1.6) des vollständigen Modells LL_V zurückgegriffen.⁶ Folgende Pseudo-R-Quadrat-Statistiken werden berechnet:⁷

¹ Vgl. BACKHAUS et al. (2011a), S. 268.

² Vgl. BROSIUS (2013), S. 616.

³ Vgl. BACKHAUS et al. (2011a), S. 268.

⁴ Vgl. ANLAGE A3, S. 219 f.

⁵ Vgl. BEST et al. (2010), S. 843.

⁶ Vgl. KRAFFT (1999), S. 246; BEST et al. (2010), S. 843.

⁷ Vgl. BEST et al. (2010), S. 843.

1) MCFADDENS – R^2

$$\text{MCFADDENS} - R^2 = 1 - \frac{LL_V}{LL_0} \quad (1.10)$$

mit

LL_0 = LogLikelihood des Nullmodells

LL_V = LogLikelihood des vollständigen Modells

Ab einem MCFADDENS – R^2 von 0,2 bis 0,4 wird von einer guten Modellanpassung gesprochen.¹

2) COX UND SNELL – R^2

$$\text{COX UND SNELL} - R^2 = 1 - \left[\frac{L_0}{L_V} \right]^{\frac{2}{K}} \quad (1.11)$$

mit

L_0 = Likelihood des Nullmodells

L_V = Likelihood des vollständigen Modells

K = Anzahl der Unternehmen

Der COX UND SNELL – R^2 kann nur Werte unter 1 annehmen.²

3) NAGELKERKE – R^2

$$\text{NAGELKERKE} - R^2 = \frac{\text{Cox \& Snell} - R^2}{R_{\max}^2} \quad (1.12)$$

mit

$$R_{\max}^2 = 1 - (L_0)^{\frac{2}{K}}$$

L_0 = Likelihood des Nullmodells

¹ Vgl. URBAN (1993), S. 62.

² Vgl. BEST et al. (2010), S. 843; BACKHAUS et al. (2011a), S. 271.

Da der NAGELKERKE – R^2 den Maximalwert von 1 erreichen kann und somit eine eindeutige inhaltliche Interpretation bezogen auf die Güte der Modellanpassung erlaubt, ist diesem Gütekriterium gegenüber dem COX & SNELL – R^2 Vorzug zu geben. Ein Wert von über 0,5 kann als sehr gut interpretiert werden.¹

Als ein weiteres Gütekriterium des Modells wird die Beurteilung der Klassifikationsergebnisse herangezogen. Hierbei werden die empirischen Beobachtungen, gekennzeichnet durch die Ausprägungen von y_k mit 1 für Akquiseziel und mit 0 für Nicht-Akquiseziel, mit den durch das Modell erzeugten Wahrscheinlichkeiten verglichen. Als Trennwert für die Klassifikation kann 0,5 verwendet werden:

$$y_k = \begin{cases} \text{Akquiseziel } y_k = 1, & \text{falls } p(y_k = 1) \geq 0,5 \\ \text{Nicht-Akquiseziel } y_k = 0, & \text{falls } p(y_k = 1) < 0,5 \end{cases} \quad (1.13)$$

Die Veranschaulichung der Klassifikationsergebnisse erfolgt in einer Klassifikationsmatrix (Tabelle 3).

tatsächliche Gruppenzugehörigkeit	prognostizierte Gruppenzugehörigkeit		
	Akquiseziel $y_k = 1$	Nicht-Akquiseziel $y_k = 0$	Prozentsatz (Trefferquote)
Akquiseziel $y_k = 1$	absolut	absolut	relativ
Nicht-Akquiseziel $y_k = 0$	absolut	absolut	relativ
Gesamtprozentsatz			relativ

Tabelle 3: Beurteilung des Klassifikationsergebnisses²

Zur Beurteilung der Trefferquote, der Gesamtprozentsatz der der korrekt klassifizierten Beobachtungen, der logistischen Regression ist ein Vergleich der Trefferquote mit derjenigen Trefferquote, die rein zufällig erreicht würde, notwendig. Bei zwei Gruppen mit gleicher Größe ist eine zufällige Trefferquote von 50% zu erwarten.³ Die formulierte Regressionsgleichung würde demnach nur einen Nutzen erbringen, wenn die errechnete Trefferquote über der zufälligen Trefferquote läge. Wird bei der Ermittlung der Trefferquote daselbe Datensample genutzt, welches auch für die Regressionsgleichung genutzt wurde,

¹ Vgl. BACKHAUS et al. (2011a), S. 270 f.

² Vgl. BACKHAUS et al. (2011a), S. 273.

³ Hier ist zu berücksichtigen, dass bei ungleicher Gruppengröße die proportionale Zufallswahrscheinlichkeit gemäß der Formel $q^2 + (1-q)^2$, mit q als dem Anteil einer der Gruppen an der Gesamtzahl der Beobachtung, zu berechnen ist. Vgl. MORRISON (1969), S. 158.

wird die Trefferquote bei anderen Datensamples geringer ausfallen.¹ Eine um diesen Fehler bereinigte Trefferquote lässt sich ermitteln, indem die verfügbaren Daten in Daten zur Modellbildung und Daten zur Modellvalidierung zufällig aufgeteilt werden. Nun ist es möglich, mittels der Daten zur Modellbildung die Regressionsfunktion zu schätzen und durch die Modellvalidierungsdaten eine bereinigte Trefferquote zu berechnen.²

Neben der Analyse der Klassifikationsmatrix kann der PRESS's Q-Test zur Beurteilung der Klassifikationsergebnisse herangezogen werden. Bei dem PRESS's Q-Test wird eine Kreuzvalidierung von Klassifikationsergebnissen vorgenommen. Hierbei folgt die Prüfgröße einer Chi-Quadrat-Verteilung mit einem Freiheitsgrad:³

$$\text{PRESS's Q} = \frac{[K - (K * G * q)]^2}{K * (G - 1)} \quad (1.14)$$

mit

K = Anzahl der Unternehmen

G = Anzahl der Gruppen

q = Gesamtprozentsatz der korrekt klassifizierten Beobachtungen

Liegt PRESS's Q oberhalb des kritischen Chi-Quadrat-Wertes, sind die Klassifikationsergebnisse signifikant vom Zufall unterschiedlich.

Der zweite Teilbereich bei der Prüfung des Gesamtmodells besteht in der Ausreißerdiagnostik. Die schlechte Anpassung der erhobenen Daten an die sachlogischen Modellzusammenhänge kann grundsätzlich an folgenden Gründen liegen:⁴

1. Die unabhängigen Variablen beeinflussen das Zustandekommen der Ausprägung von y_k nicht.

¹ Vgl. KRAFFT (1999), S. 246 f.

² Dieses Vorgehen führt jedoch dazu, dass die Zuverlässigkeit der Regressionsgleichung mit zunehmendem Umfang der Kontrollstichprobe und bei abnehmender Größe der verfügbaren Daten sinkt. Um hier Abhilfe zu schaffen, kann die Jackknife-Methode angewandt werden. Vgl. MELVIN et al. (1977), S. 60 ff.

³ Vgl. BACKHAUS et al. (2011a), S. 274.

⁴ Vgl. BACKHAUS et al. (2011a), S. 277.

2. Es existiert eine Anzahl von Beobachtungen, die den vom Modell beschriebenen Wirkungszusammenhang nicht aufweisen und durch ihre besonderen Ausprägungen das Ergebnis verzerren.

Der erste Fall kann durch die Überprüfung der Hypothesen und ggf. durch Aufnahme weiterer Hypothesen behoben werden. Dies impliziert ggf. die Neuformulierung des Modells.

Der zweite Fall tritt auf, wenn zwischen der beobachteten Ausprägung von y_k und der durch die Regressionsgleichung geschätzten Wahrscheinlichkeit p für y_k große Diskrepanz herrscht.¹ Ob Ausreißer vorliegen, lässt sich für jedes Unternehmen k durch die bestimmbaren individuellen Residuen wie folgt berechnen:²

$$RESID_k = y_k - p(y_k = 1) \quad (1.15)$$

Ob und welche Residuen statistische Ausreißer sind, kann nicht durch allgemeine statistische Maße gestützt werden.³ Wenn jedoch Residuen betragsmäßig Werte deutlich größer als 0,5 annehmen, ist das ein Hinweis auf Klassifikationsfehler.⁴ Um genau diese besser zu erkennen, werden die Residuen gewichtet. Eine mögliche Gewichtung stellen die standardisierten Residuen dar. Diese lassen sich wie folgt berechnen:⁵

$$ZRESID_k = \frac{y_k - p(y_k = 1)}{\sqrt{p(y_k = 1) * (1 - p(y_k = 1))}} \quad (1.16)$$

Für solche Ausreißer sind grundsätzlich zwei Ursachen denkbar:⁶

1. Die Daten sind atypisch. In diesem Falle sollten die Daten aus der Analyse ausgeschlossen werden.
2. Das Modell ist schlecht spezifiziert. Dies bedeutet, dass wichtige Einflussfaktoren bei der Modellbildung unberücksichtigt blieben und dass das Modell erweitert oder modifiziert werden muss.

¹ Vgl. BACKHAUS et al. (2011a), S. 277.

² Vgl. BACKHAUS et al. (2011a), S. 277.

³ Vgl. BACKHAUS et al. (2011a), S. 277.

⁴ Vgl. FROMM (2005), S. 27; BACKHAUS et al. (2011a), S. 277.

⁵ Vgl. SPSS Inc. (2007), S. 432.

⁶ Vgl. BACKHAUS et al. (2011a), S. 278.

2.1.1.2.5 Prüfung der Merkmalsvariablen

Die Trennfähigkeit der einzelnen Merkmalsvariablen können Erkenntnisse über ein Modell-Overfitting¹ liefern. Folgende zwei Tests können hierzu angewendet werden:²

1. Likelihood-Quotienten-Test
2. Wald-Statistik

Im Gegensatz zum Likelihood-Ratio-Test werden bei dem Likelihood-Quotienten-Test nicht alle Regressionskoeffizienten auf Null gesetzt. Bei dem Likelihood-Quotienten-Test werden verschiedene reduzierte Modelle gebildet, bei denen jeweils ein Regressionskoeffizient auf Null gesetzt wird und die Differenz der -2 LogLikelihoods (-2LL) zwischen vollständigem Modell (LL_V) und einem reduzierten Modell (LL_R) gebildet wird.³ Die zugrunde liegende Nullhypothese ist folgende:⁴

H_0 : Die Effekte des Regressionskoeffizienten b_h sind Null ($b_h = 0$).

Die alternative Hypothese ist folgende:⁵

H_1 : Die Effekte des Regressionskoeffizienten b_h sind ungleich Null ($b_h \neq 0$).

Die Signifikanz der resultierenden Größe ($LL_R - LL_V$) wird mittels der Chi-Quadrat-Verteilung geprüft.⁶ Der Freiheitsgrad ist abhängig von der Anzahl der getesteten Variablen.

Die WALD-Statistik ist an den t-Test der linearen Regression angelehnt.⁷ Wie bei dem Likelihood-Quotienten-Test wird die Nullhypothese, dass das konstante Glied b_0 oder ein

¹ Wenn die Modellbildungsdaten nicht systematische Fehler wie Messfehler enthalten, können Wirkungszusammenhänge zwischen abhängiger Variable und unabhängigen Variablen verallgemeinert werden und somit zu falschen Modellen führen. Vgl. o.A. (2015), o.S.

² Vgl. BACKHAUS et al. (2011a), S. 279.

³ Vgl. BACKHAUS et al. (2011a), S. 279.

⁴ Vgl. BACKHAUS et al. (2011a), S. 279.

⁵ Vgl. BACKHAUS et al. (2011a), S. 279.

⁶ Vgl. BACKHAUS et al. (2011a), S. 279.

⁷ Vgl. BACKHAUS et al. (2011a), S. 280.

bestimmter Regressionskoeffizient b_h Null ist, getestet. Die Formel der WALD-Teststatistik lautet:¹

$$WT = \left(\frac{b_h}{s_{b_h}} \right)^2 \quad (1.17)$$

mit

b_h = Regressionskoeffizienten für die unabhängige Variable h ($b_h \in \mathbb{R} \forall h = 1, 2, \dots, H$)

s_{bh} = Standardfehler von b_h ($s_{bh} \in \mathbb{R} \forall b_h \in \mathbb{R} \forall h = 1, 2, \dots, H$)

H = Anzahl der unabhängigen Variablen

Die Prüfgröße ist die Chi-Quadrat-Verteilung bei einem Freiheitsgrad von Eins. Ist der ermittelte WALD-Wert größer als der kritische Chi-Quadrat-Wert, ist der Einfluss des konstanten Gliedes b_0 oder der Regressionskoeffizienten b_h bestätigt.

¹ Vgl. BACKHAUS et al. (2011a), S. 280; ECKSTEIN (2012), S. 220 f. Die hier dargestellte WALD-Statistik bezieht sich einfachheitshalber auf b_h .

2.1.2 Komplexe Methoden der multivariaten Analyse: künstliche neuronale Netze und Self Enforcing Networks

2.1.2.1 Einführung in künstliche neuronale Netze

2.1.2.1.1 Biologisches Neuron und künstliches Neuron

Die Grundlogik der künstlichen neuronalen Netze orientiert sich an den biologischen Prozessen des Gehirns.¹ Das Gehirn besteht aus einer Vielzahl von vernetzten Nervenzellen, den sog. Neuronen. Abbildung 5 zeigt ein vereinfachtes biologisches Neuron.

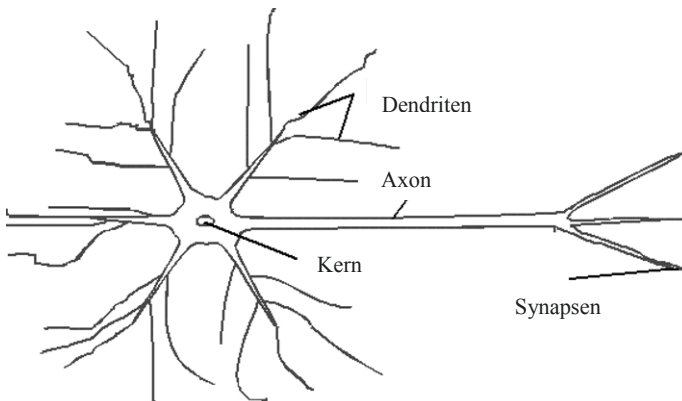


Abbildung 5: Biologisches Neuron²

Die Informationsverarbeitung geschieht wie folgt: Das Neuron empfängt über seine Dendriten erregende oder hemmende Signale von mehreren Neuronen. Diese werden im Empfängerneuron zu einem Gesamtsignal verdichtet. Überschreitet dieses einen gewissen Schwellenwert, wird ein elektrischer Impuls über das Axon, welches über seine Synapsen mit Dendriten anderer Neuronen verbunden ist, an die Folgeneuronen weitergeleitet.³ Je nach Typ der Synapse kann das Signal verstärkt (excitatorische Synapse) oder gehemmt (inhibitorische Synapse) weitergeleitet werden. Wird ein Signal durch eine excitatorische Synapse gesendet, steigt die Wahrscheinlichkeit, dass der Schwellenwert des Folgeneurons

¹ Vgl. KLÜVER et al. (2012), S. 115.

² Vgl. FANGMEYER (2003), o.S.

³ Vgl. BACKHAUS et al. (2011), S. 172 f.; KLÜVER et al. (2012), 115 f.; KÜTING et al. (2012), S. 380.

überschritten wird. Umgekehrt führt die Übertragung eines Signals über eine inhibitorische Synapse zur Reduzierung der Wahrscheinlichkeit einer Überschreitung des Schwellenwertes.¹ Durch Lernen können die excitatorischen oder inhibitorischen Wirkungen von Synapsen sich verändern. Zudem wachsen die Synapsen bei häufiger Nutzung und degenerieren bei seltener Nutzung, sodass die Folgeneuronen stärker bzw. schwächer von dem sendenden Neuron beeinflusst werden.²

Ähnlich wie biologische neuronale Netze sind auch künstliche neuronale Netze aufgebaut. Während bei den biologischen neuronalen Netzen biochemische Prozesse die Grundlage der Informationsverarbeitung bilden, sind bei künstlichen neuronalen Netzen geeignete mathematische Operationen für die Informationsverarbeitung zuständig.³ Abbildung 6 zeigt ein künstliches Neuron.

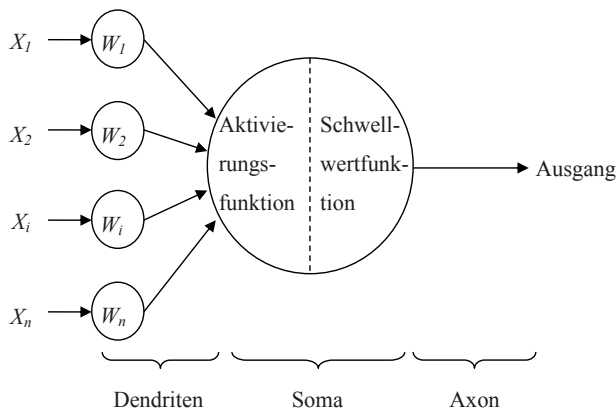


Abbildung 6: Künstliches Neuron⁴

Bei x_1 bis x_n handelt es sich um Signale sendender Neuronen und bei w_1 bis w_n um die Stärke der Verbindung zwischen einem sendenden Neuron und dem Empfängerneuron. Diese Signale können in Abhängigkeit von ihrem Vorzeichen entweder excitatorische oder inhibitorische Wirkungen haben, hierbei führt ein „+“ zur excitatorischen und ein „-“ zur

¹ Vgl. BACKHAUS et al. (2011), S. 173; KÜTING et al. (2012); S. 380.

² Vgl. BACKHAUS et al. (2011), S. 173.

³ Vgl. SCHMIDT et al. (2010), S. 141 f.; KÜTING et al. (2011), S. 381 f.

⁴ In Anlehnung an KLÜVER et al. (2012), S. 120.

inhibitorischen Wirkung. Die Informationsverarbeitung geschieht im Neuron durch unterschiedliche Funktionen. Zunächst wird durch die Inputfunktion der Zustand des Neurons zu einem bestimmten Zeitpunkt beschrieben.¹ Hierdurch wird der Nettoinput für jedes Neuron j , der durch die Summierung der Multiplikationen der Eingangssignale a_i mit den Gewichtswerten w_{ij} gebildet wird, berechnet.²

$$net_j = \sum_{i=1}^I w_{ij} * a_i \quad (2.1)$$

mit

net_j = Nettoinput des empfangenden Neurons j ($net \in \mathbb{R} \forall j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($a \in \mathbb{R} \forall i = 1, 2, \dots, I$)

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R} \forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

Dieser Nettoinput net_j wird von der Aktivierungsfunktion zur Berechnung des Aktivierungswertes a_j verwendet. In einigen Fällen wird auf die Unterscheidung zwischen Inputfunktion und Aktivierungsfunktion verzichtet.³ Dies führt dazu, das sendende Neuronen ihren Zustand als Signal ohne Modifikation übermitteln. Die am häufigsten verwendete Aktivierungsfunktion ist folgende:⁴

$$a_j = \sum_{i=1}^I w_{ij} * a_i \quad (2.2)$$

mit

¹ Vgl. KLÜVER et al. (2012), S. 118.

² Vgl. KLÜVER et al. (2012), S. 118.

³ Vgl. KLÜVER et al. (2012), S. 119 f.

⁴ Vgl. KLÜVER et al. (2012), S. 119; KÜTING et al. (2012), S. 382.

a_j = Aktivierungswert des empfangenden Neurons j ($a \in \mathbb{R} \forall j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($a \in \mathbb{R} \forall i = 1, 2, \dots, I$)

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R} \forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

Zur Erkennung von nicht-linearen Strukturen in den Eingangsdaten benötigt ein künstliches neuronales Netz eine nicht-lineare Aktivierungsfunktion.¹ Auch bei den nicht-linearen Aktivierungsfunktionen sind in der Praxis verschiedene Funktionen gebräuchlich.²

Die Übermittlung des Aktivierungswertes a_j an Empfängerneuronen kann von Schwellenwerten (auch Schwellwertfunktion genannt) abhängen. Schwellenwerte können bei der Ausgabefunktion eine Rolle spielen und zwar dann, wenn ein binärer Output benötigt wird.³ In der Praxis wird jedoch die Ausgabefunktion mit der Aktivierungsfunktion gleichgesetzt.⁴

¹ Vgl. REHKUGLER (1996), S. 572.

² Vgl. SCHLITTGEN (2009), S. 387; SCHMIDT et al. (2010), S. 142 f.; BACKHAUS et al. (2011), S. 178 f.; KLÜVER et al. (2012), S. 119; KÜTING et al. (2012), S. 382. Mit verschiedenen Funktionen sind z.B. die Sinus-Funktion, die logistische Funktion und die Tangens-Hyperbolicus-Funktion gemeint. Vgl. KÜTING et al. (2012), S. 382.

³ Vgl. KLÜVER et al. (2012), S. 120; KÜTING et al. (2012), S. 382.

⁴ Vgl. KLÜVER et al. (2012), S. 120.

2.1.2.1.2 Interne und externe Topologien künstlicher neuronaler Netze

Das Gehirn besteht aus einer Vielzahl von vernetzten Nervenzellen. Analog zum biologischen Vorbild entsteht ein künstliches neuronales Netz durch die Vernetzung einer Vielzahl von künstlichen Neuronen.¹ Die Anordnung der Neuronen und die bestehenden Verbindungen werden als externe Topologie bezeichnet.² In der Praxis sind mehrschichtige künstliche neuronale Netze gebräuchlich.³ Abbildung 7 zeigt ein dreischichtiges künstliches neuronales Netz.

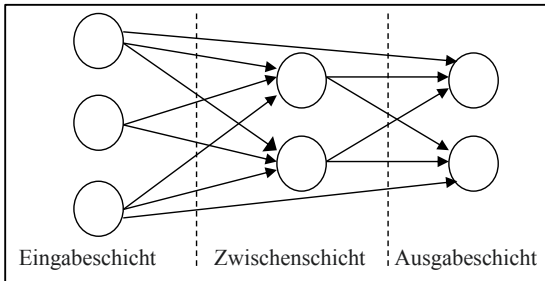


Abbildung 7: Dreischichtiges künstliches neuronales Netz⁴

Bei den Neuronen der Eingabeschicht (Input-Layer) handelt es sich um Inputneuronen, die lediglich Informationen aus der Umwelt aufnehmen und i. d. R. unverändert weitergeben.⁵ Die Neuronen der Zwischenschicht (Hidden-Layer) sind nicht mit der Außenwelt verbunden und dienen allein der Informationsverarbeitung.⁶ Diese Zwischenschicht ist bei geschichteten Feed-Forward-Netzen notwendig, da sonst gewisse Informationen nicht verarbeitet werden können.⁷ Die Neuronen der Ausgabeschicht (Output-Layer), Outputneuronen, stellen das Ergebnis des künstlichen neuronalen Netzes dar.⁸ Neben dieser Art von

¹ Vgl. BACKHAUS et al. (2011), S. 174.

² Hiermit ist lediglich die Zuordnung der Neuronen zu der Eingabeschicht, der Zwischenschicht und der Ausgabeschicht gemeint. Vgl. KLÜVER et al. (2012), S. 123.

³ Vgl. SCHLITGEN (2009), S. 387; BACKHAUS et al. (2011), S. 174 f.; KÜTING et al. (2012), S. 383 f.

⁴ In Anlehnung an KLÜVER et al. (2012), S. 122.

⁵ Vgl. PYTLIK (1995), S. 311.

⁶ Vgl. KÜTING et al. (2012), S. 383.

⁷ Hiermit ist eine Endweder-Oder-Funktion (XOR-Funktion) gemeint. Vgl. KLÜVER et al. (2012), S. 124.

⁸ Vgl. BACKHAUS et al. (2011), S. 174 f.; KLÜVER et al. (2012), S. 125 f.; KÜTING et al. (2012), S. 383.

geschichteten Netzen existieren auch nicht geschichtete Netze. Abbildung 8 zeigt ein nicht geschichtetes Netz.

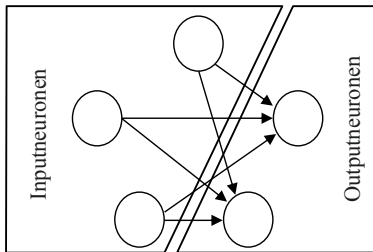


Abbildung 8: Nicht geschichtetes künstliches neuronales Netz¹

Bei nicht geschichteten Feed-Forward-Netzen werden Inputneuronen – ggf. alle vorhandenen Neuronen – und Outputneuronen – ggf. wieder alle vorhandenen Neuronen – festgelegt.² Die nicht geschichteten Feed-Forward-Netze sind die einfachsten Strukturen künstlicher neuronaler Netze. Die Ergebnisqualität eines nicht geschichteten Feed-Forward-Netzes kann grundsätzlich mit denen der geschichteten Netze ähnlich sein.

Die Neuronen können auf unterschiedliche Weise miteinander verbunden werden. Hier lassen sich, wie oben angeführt, zwei Arten von Netzen unterscheiden:³

1. Feed-Forward-Netze
2. Feed-Back-Netze bzw. rekurrente Netze

Bei den Feed-Forward-Netzen handelt es sich um Netze, die rückkopplungsfrei sind (vgl. Abbildung 7 und Abbildung 8).⁴ Mit Rückkopplung ist die Umkehrung der Aktivierungsausbreitung gemeint. Verbindungen bestehen nur zwischen den Neuronen der unterschiedlichen Schichten und Informationen fließen nur in eine Richtung über die Neuronen der Eingabeschicht hin zu den Neuronen der Ausgabeschicht.⁵

¹ In Anlehnung an KINNEBROCK (1994), S. 32.

² Vgl. KLÜVER (2015), S. 551.

³ Vgl. KLÜVER (2015), S. 551.

⁴ Vgl. KÜTING et al. (2012), S. 383; BACKHAUS et al. (2011), S. 174 f.; KLÜVER et al. (2012), S. 125 f.

⁵ Vgl. KÜTING et al. (2012), S. 383; BACKHAUS et al. (2011), S. 174 f.; KLÜVER et al. (2012), S. 125 f.

Bei den Feed-Back-Netzen können Rückkopplungen zwischen den Neuronen der unterschiedlichen Schichten bestehen.¹ Hier ist auch ein Informationsfluss und somit eine Aktivierungsausbreitung in beide Richtungen der Verbindung möglich.² Bei rekurrenten Netzen sind neben den Rückkopplungen zwischen Neuronen der unterschiedlichen Schichten auch Verbindungen zwischen Neuronen derselben Schicht möglich. Bei nicht geschichteten Netzen sind demnach Feed-Back-Netze per Definition rekurrente Netze. Abbildung 9 zeigt ein geschichtetes Feed-Back-Netz und ein nicht geschichtetes rekurrentes Netz.

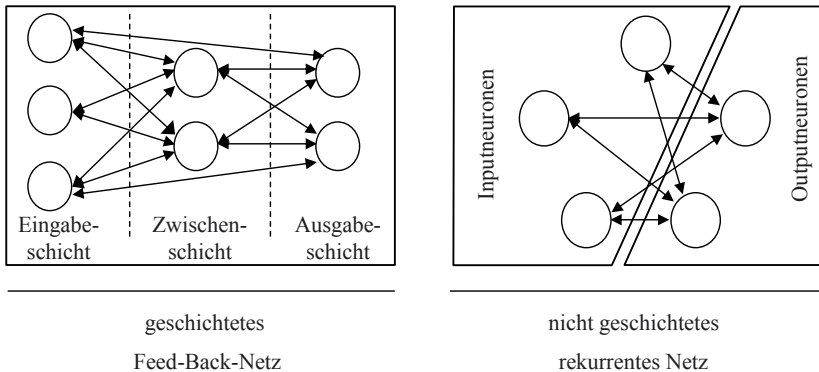


Abbildung 9: Geschichtetes Feed-Back-Netz und nicht geschichtetes rekurrentes Netz³

Es existiert keine Überlegenheit eines Netzmodells.⁴

Die interne Topologie wird durch eine Gewichtsmatrix repräsentiert.⁵ Diese beinhaltet bei den Feed-Forward-Netzen nur die Stärken der einzelnen Verbindungen zwischen den Neuronen unterschiedlicher Schichten und bei den Feed-Back-Netzen zusätzlich die Stärken der Rückkopplungen.⁶

¹ Vgl. KÜTING et al. (2012), S. 383; BACKHAUS et al. (2011), S. 174 f.; KLÜVER et al. (2012), S. 125 f.

² Vgl. KÜTING et al. (2012), S. 383; BACKHAUS et al. (2011), S. 174 f.; KLÜVER et al. (2012), S. 125 f.

³ In Anlehnung an KLÜVER et al. (2012), S. 126.

⁴ Vgl. KÜTING et al. (2012), S. 383.

⁵ Vgl. KLÜVER et al. (2012), S. 123.

⁶ Vgl. KLÜVER et al. (2012), S. 123.

2.1.2.1.3 Lernprozess der künstlichen neuronalen Netze

Ein Lernprozess ist die Voraussetzung zur Lösung eines Problems mithilfe der künstlichen neuronalen Netze.¹ Der Lernprozess der künstlichen neuronalen Netze orientiert sich hierbei an dem biologischen Vorbild. Die synaptischen Aktivitäten, die durch Lernen entweder eine Verbindung zwischen Neuronen stärken oder schwächen, werden im künstlichen neuronalen Netz durch eine Veränderung der Verbindungsgewichte erreicht.² Der Lernprozess im künstlichen neuronalen Netz kann in zwei Verfahren unterteilt werden:

1. überwachtes Lernen
2. unüberwachtes Lernen

Beim überwachten Lernen wird ein Ergebnis vorgegeben.³ Das Ziel des künstlichen neuronalen Netzwerks ist es, Verbindungsstrukturen zwischen Neuronen aufzubauen, die aus den Eingabedatensamples möglichst genau die Ausgabedatensamples ableiten.⁴ Das künstliche neuronale Netz vergleicht hierbei die ausgegebenen Ist-Werte mit den Soll-Werten und modifiziert ausgehend von den Ausgabeneuronen die Verbindungsgewichte, bis die Differenz zwischen den Ist- und Soll-Werten minimiert ist.⁵ Die Veränderung der Verbindungsgewichte erfolgt nach einer Lernregel.⁶

Beim unüberwachten Lernen wird kein Ausgabedatensample vorgegeben. Stattdessen werden systemimmanente Bewertungskriterien angewandt.⁷ Hierbei wird das Neuron mit dem höchsten Aktivierungswert, bei mehreren Eingabesignalen sind dies meist mehrere Neuronen, ausgewählt.⁸ Die nicht selektierten Neuronen werden in topologisch angeordnete

¹ Vgl. KÜTING et al. (2012), S. 385.

² Vgl. BACKHAUS et al. (2011), S. 175.

³ Vgl. KLÜVER et al. (2012), S. 127 f.

⁴ Vgl. BAETGE et al. (1994), S. 338.

⁵ Vgl. KÜTING et al. (2012), S. 385.

⁶ In der Praxis wird häufig die allgemeine Lernregel (WIDROW-HOFF-Regel; auch bekannt als Delta-Regel) oder eine modifizierte Version dieser angewendet. Vgl. KLÜVER et al. (2012), S. 128. Die Lernregel, mit der das SEN operiert, wird in Kapitel 2.1.2.2.5 erläutert.

⁷ Vgl. KLÜVER et al. (2012) S. 128.

⁸ Vgl. KLÜVER et al. (2012) S. 129.

Gruppen gruppiert.¹ Die ausgewählten Neuronen bilden Clusterzentren für die nicht ausgewählten Neuronen. Nach der Eingruppierung der nicht ausgewählten Neuronen zu Clusterzentren entstehen Cluster.² Das künstliche neuronale Netz ist solange im Lernprozess, bis die Entfernungen der Neuronen zu den Clusterzentren minimiert sind.³

2.1.2.1.4 Typen eines künstlichen neuronalen Netzes

Durch die Kombination der in Kapitel 2.1.2.1.2 eingeführten externen Topologien und der in Kapitel 2.1.2.1.3 dargestellten unterschiedlichen Lernverfahren ergeben sich vier Typen von künstlichen neuronalen Netzen. Abbildung 10 zeigt die unterschiedlichen Typen mit Beispielen für künstliche neuronale Netze.

	überwachtes Lernen	unüberwachtes Lernen
Feed-Forward-Netze	Multi-Layer-Perceptron (MLP)	1. Self-Organizing-Map (SOM) 2. Self-Enforcing-Network (SEN) ⁴
Feed-Back-Netze	Artmap	Self-Enforcing-Network (SEN) ⁵

Abbildung 10: Typen von künstlichen neuronalen Netzen mit Beispielen⁶

¹ Vgl. KLÜVER et al. (2012), S. 128 f.

² Vgl. KLÜVER et al. (2012), S. 129.

³ Dies erfolgt durch die Anpassung der Gewichtsverbindungen zwischen den entsprechenden Neuronen. Vgl. KLÜVER et al. (2012), S. 129.

⁴ Das SEN bietet Funktionalitäten zur Realisierung von Feed-Forward- und Feed-Back-Netzen.

⁵ Das SEN bietet Funktionalitäten zur Realisierung von Feed-Forward- und Feed-Back-Netzen.

⁶ In Anlehnung an BACKHAUS et al. (2011), S. 176 und Klüver et al. (2012), S. 140 ff.

2.1.2.1.5 Trainieren eines künstlichen neuronalen Netzes

Das Trainieren eines künstlichen neuronalen Netzes ist der eigentliche Lernprozess. Die Informationsverarbeitung inkl. des Lernprozesses wird in Abbildung 11 dargestellt.

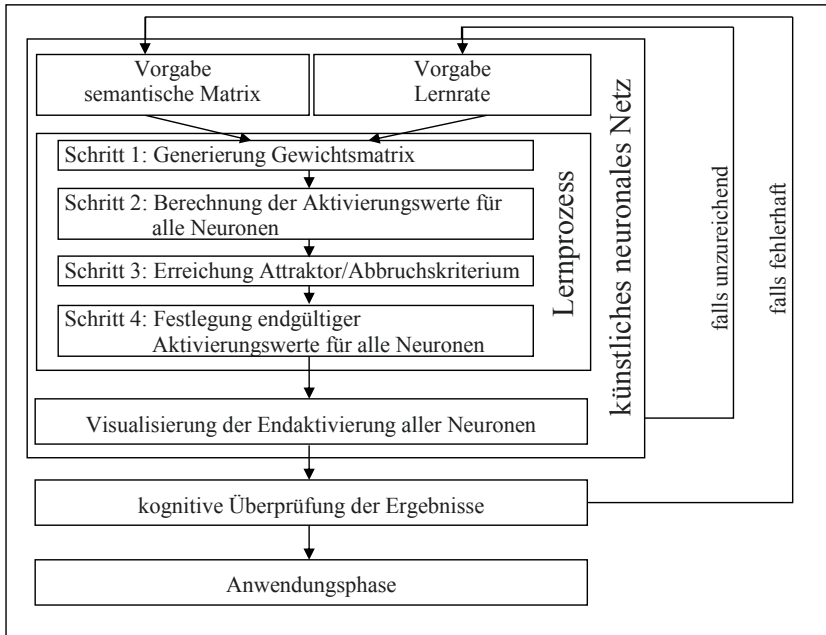


Abbildung 11: Ablaufschritte der Informationsverarbeitung¹

Falls die Werte der Gewichtsmatrix keinen hinreichenden Lernerfolg² ergeben, wird der Lernprozess wiederholt.³ Führt das wiederholte durchführen des Lernprozesses zu keinem hinreichenden Lernerfolg, sollten die Ergebnisse kognitiv überprüft werden, um dann An-

¹ In Anlehnung an KLÜVER et al. (2012), S. 138 ff.

² Der hinreichende Lernerfolg wird in der vorliegenden Arbeit anhand der richtig klassifizierten Unternehmen gemessen. Ein hinreichender Lernerfolg ist erreicht, wenn der prozentuale Anteil der richtig klassifizierten Unternehmen größer ist als die proportionale Zufallswahrscheinlichkeit gemäß der Formel $q^2 + (1 - q)^2$, mit q als der Anteil einer der Gruppen an der Gesamtzahl der Beobachtung, zu berechnen ist. Vgl. MORRISON (1969), S. 158.

³ Vgl. KLÜVER et al. (2012), S. 141.

passungen in der semantischen Matrix vornehmen zu können. Dabei hat der Anwender folgende Möglichkeiten, das Lernen zu beeinflussen:¹

1. Entwicklung neuer Verbindungen
2. Löschen existierender Verbindungen
3. Modifikation der Aktivierungsfunktion
4. Hinzufügen neuer Neuronen
5. Löschen von Neuronen
6. Modifikation der Gewichtswerte w_{ij} von Verbindungen

Die größte Bedeutung liegt in der Modifikation der Gewichtswerte w_{ij} von Verbindungen und der Hinzufügung oder Löschung von Neuronen.² Wurden Anpassungen im künstlichen neuronalen Netz vorgenommen, wird der Lernprozess mit einer geringen Lernrate wiederholt. Dieses Vorgehen wird so lange wiederholt, bis die Ergebnisse hinreichend sind.

Wird als Abbruchkriterium ein Maß verwendet, welches sich ausschließlich auf die Daten des Trainingsprozesses bezieht, was insbesondere bei dem überwachten Lernen der Fall ist, kann es zu einem Overfitting kommen.³ Beim Overfitting lernt das künstliche neuronale Netz nicht-systematische Fehler wie Messfehler und verallgemeinert diese. Deshalb sollte der Lernprozess [mit (d) als Lernschritt] beendet werden, wenn ein hinreichender Lernerfolg erreicht ist, da ansonsten die Fehler (E) bei der Klassifikation der Neuronen zunehmen.⁴ Um genau diesen hinreichenden Lernerfolg (D_{opt}) zu identifizieren, ist ein Datensatz notwendig, der nicht im Trainingsprozess verwendet wurde. In der Praxis wird bei dem überwachten Lernen daher die Unterteilung des erhobenen Datensatzes zufällig in Trainingsdaten (80%), Validierungsdaten (10%) und Testdaten (10%) empfohlen.⁵ Abbildung 12 veranschaulicht den Lernprozess:

¹ Vgl. ZELL (1994), S. 84.

² Vgl. BACKHAUS et al. (2011), S. 192.

³ Vgl. BACKHAUS et al. (2011), S. 196.

⁴ Vgl. BACKHAUS et al. (2011), S. 196.

⁵ Vgl. BACKHAUS et al. (2011), S. 197.

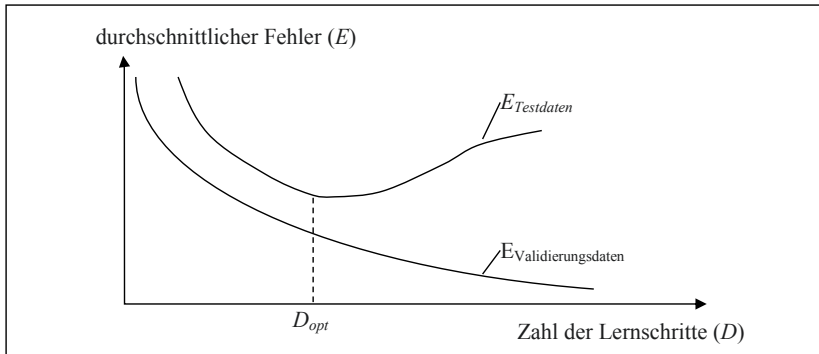


Abbildung 12: Lernprozesses¹

Bei dem unüberwachten Lernen ist diese Aufteilung nicht notwendig, da das Testen im Sinne des überwachten Lernens nicht durchgeführt wird. Der Datensatz bei dem unüberwachten Lernen wird analog zum Vorgehen bei der logistischen Regression in Daten zur Modellbildung (80%) und Modellvalidierung (20%) aufgeteilt.

Der Lernerfolg kann durch die Klassifikationsergebnisse überprüft werden. Hierzu sollten nicht ausschließlich die Modellbildungsdaten verwendet werden. Denn auch bei dem künstlichen neuronalen Netz führt die Ermittlung der Trefferquote auf Basis der Modellbildungsdaten zu tendenziell besseren Ergebnissen als die Ermittlung der Trefferquote auf Basis der Modellvalidierungsdaten. Das Klassifikationsergebnis der Modellvalidierungsdaten stellt eine um diesen Fehler korrigierte Trefferquote dar. Bei zwei Gruppen mit gleicher Größe ist eine zufällige Trefferquote von 50% zu erwarten.² Das künstliche neuronale Netz erbringt einen hinreichenden Lernerfolg, wenn die errechnete Trefferquote über der zufälligen Trefferquote liegt.

Um die Güte des künstlichen neuronalen Netzes zu bestimmen, werden die Klassifikationsergebnisse, also die Anzahl der richtig klassifizierten Unternehmen³, auf der Basis der Modellbildungs- und Modellvalidierungsdaten herangezogen.

¹ Vgl. BACKHAUS et al. (2011), S. 197.

² Bei ungleicher Gruppengröße ist die proportionale Zufallswahrscheinlichkeit gemäß der Formel $q^2 + (1-q)^2$, mit q als dem Anteil einer der Gruppen an der Gesamtzahl der Beobachtung, zu berechnen. Vgl. MORRISON (1969), S. 158.

2.1.2.1.6 Anwendungen des trainierten Netzes

Die Anwendungsmöglichkeiten des entwickelten künstlichen neuronalen Netzes beschränken sich auf Anwendungsdaten, bei denen zwischen Inputneuronen und Outputneuronen gleiche Kausalzusammenhänge unterstellt werden, die auch den Modellbildungsdaten zugrunde lagen. Bei dem Anwendungsdatensatz existieren im Vergleich zu den Modellbildungs- und Modellvalidierungsdaten keine Soll-Ausgabewerte, die zur Überprüfung der Ergebnisse zur Verfügung stehen, daher ist eine Fehleranalyse erst dann möglich, wenn eine Akquise eingetreten ist.¹

Zur Verbesserung des Verständnisses der Kausalzusammenhänge können Sensitivitätsanalysen eingesetzt werden.² Hierfür werden einzelne Eingangssignale variiert, um den Einfluss der Variation auf die Outputneuronen zu ermitteln.

¹ Vgl. BACKHAUS et al. (2011), S. 199.

² Vgl. BACKHAUS et al. (2011), S. 199.

2.1.2.2 Vorgehensmodell für den Einsatz von künstlichen neuronalen Netzen

2.1.2.2.1 Problemstrukturierung

Bei dem Einsatz von künstlichen neuronalen Netzen ist es notwendig, wie auch schon bei der logistischen Regression, dass Einflussgrößen (d.h. unabhängige Variablen) und Zielvariablen (d.h. abhängige Variablen) im Vorfeld definiert werden müssen.¹ Hierbei sollte die Bestimmung der Einflussgrößen auf der Grundlage von theoretisch begründeten Hypothesen zwischen Einflussgrößen und Zielvariablen vorgenommen werden.²

Abhängig vom Problemtyp ist ein Netztyp auszuwählen. Abbildung 13 zeigt eine mögliche Zuordnung verschiedener Netztypen zu Problemtypen.³

Problemtyp	Netztyp
Prognose	MLP
Zuordnung zu gegebener Klasse	MLP; SOM; SEN
Klassifikation	SOM; SEN

Abbildung 13: Zuordnung verschiedener Netztypen zu Problemtypen⁴

Das Ziel in der vorliegenden Arbeit, welches bei der Anwendung der künstlichen neuronalen Netze verfolgt wird, ist zunächst eine Klassifikation von Unternehmen anhand von Finanzkennzahlen, anhand derer die Unternehmen in die Klasse der Akquiseziele und in die Klasse der Nicht-Akquiseziele eingeteilt werden. Im Anschluss hieran sollen Unternehmen der deutschen Stahlindustrie als Akquiseziele oder Nicht-Akquiseziele identifiziert werden. Die Problemtypen können als „Klassifikation“ und „Zuordnung zu gegebener Klasse“ definiert werden.

¹ Vgl. BACKHAUS et al. (2011), S. 184.

² Vgl. BACKHAUS et al. (2011), S. 184.

³ In der Praxis existieren viele unterschiedliche Netztypen. Vgl. BACKHAUS et al. (2011), S. 184 f.; KLÜVER et al. (2012), S. 140. Die vorliegende Arbeit beschränkt sich auf einige wenige Netztypen, die für die Problemstellung relevant sind.

⁴ In Anlehnung an BACKHAUS et al. (2011), S. 185; KLÜVER et al. (2012), S. 140. Die Zuordnung von SEN erfolgte durch den Autor der vorliegenden Arbeit.

2.1.2.2.2 Netztypauswahl

Bei dem überwachten Lernen besteht das Risiko, dass das künstliche neuronale Netz nicht nur allgemeingültige Strukturen aus dem Ausgabedatensample lernt, sondern auch spezifische Strukturen aus dem Ausgabedatensample gelernt werden.¹ Damit ist konkret ein Overfitting gemeint, wodurch eventuell nicht allgemein für die Stahlindustrie gültige Hypothesen, die durch die in Kapitel 2.2 definierten Finanzkennzahlen beschrieben werden, im künstlichen neuronalen Netz berücksichtigt werden.

Der Autor der vorliegenden Arbeit wird sich deshalb bei der Auswahl eines künstlichen neuronalen Netzes auf die Netze beschränken, welche unüberwacht lernen. Wie in Kapitel 2.1.2.1.2 aufgeführt, existiert keine generelle Überlegenheit eines Netztypen.² Die Netztypen SEN und SOM sind als Netztypen, die unüberwacht lernen, auswählbar. Das SEN bietet im Vergleich zur SOM die Möglichkeit, Feed-Forward- und Feed-Back-Netze zu konstruieren.³ Zudem sind SOMs nicht deterministisch und liefern bei der wiederholten Klassifikation der gleichen Objekte (Akquiseziele und Nicht-Akquiseziele) unterschiedliche Ergebnisse. Die Reproduzierbarkeit der Klassifikationsergebnisse ist nicht vorhanden. Das SEN hingegen ist deterministisch und ermöglicht eine Reproduzierbarkeit der Klassifikationsergebnisse.⁴ Das SEN ist für die vorliegende Problemstellung besser geeignet.

Das SEN ist ein Tool der COBASC-Research Group der Fakultät für Wirtschaftswissenschaften der Universität Duisburg-Essen.

2.1.2.2.3 Konzeptualisierungsprämissen

Für die Verwendung des SEN (und auch anderer Netztypen) ist es notwendig einige Konzeptualisierungsprämissen zu definieren. Insbesondere ist die Klärung der elementaren Zusammenhänge zwischen KNN-Stahl-Prognosemodellen und dem SEN erforderlich. In diesem Kapitel erfolgt diese Klärung. In den Kapitel 2.1.2.2.4 und 2.1.2.2.5 werden dann die technischen Beziehungen und Funktionsweisen der einzelnen Elemente des SEN be-

¹ Vgl. KÜTING et al. (2012), S. 385.

² Vgl. KÜTING et al. (2012), S. 383.

³ Vgl. COBASC (2014), o.S.

⁴ Vgl. COBASC (2014), o.S.

schrieben. Die Bearbeitung der Elemente des SEN erfolgt idealtypisch in einer bestimmten Reihenfolge im SEN.¹ Dieser Reihenfolge folgen die Kapitel 2.1.2.2.4 und 2.1.2.2.5. An dieser Reihenfolge orientiert sich auch die Klärung der elementaren Zusammenhänge zwischen KNN-Stahl-Prognosemodellen und dem SEN.

Im SEN werden Attribute angelegt, die Objekte und Eingabe-Vektoren beschreiben. Diese Attribute sind in der vorliegenden Arbeit die Finanzkennzahlen, die die Hypothesen zu Übernahmen und Fusionen operationalisieren und im Rahmen der logistischen Regression als unabhängige Variablen beschrieben wurden. Objekte und Eingabe-Vektoren sind demnach Unternehmen. Die Differenzierung von Objekten und Eingabe-Vektoren erfolgt in den folgenden Absätzen. Die Attribute können im SEN unterschiedliche Ausprägungen aufweisen, zum einen bedingt durch Rechenoperationen und zum anderen bedingt durch Berücksichtigung von Objekten und Eingabe-Vektoren.

Die Attributsausprägungen können in die tatsächliche Attributsausprägung eines Attributs, in die minimale Attributsausprägung, in die maximale Attributsausprägung, in die normalisierte minimale Attributsausprägung und in die normalisierte maximale Attributsausprägung unterschieden werden. Die tatsächliche Attributsausprägung eines Attributs gibt den beobachteten Wert einer Finanzkennzahl an. Die minimale Attributsausprägung gibt unter Berücksichtigung aller Objekte und Eingabe-Vektoren den kleinsten beobachteten Wert einer Finanzkennzahl wieder. Die maximale Attributsausprägung gibt unter Berücksichtigung aller Objekte und Eingabe-Vektoren den größten beobachteten Wert einer Finanzkennzahl wieder. Abhängig von der gewählten Kodierung werden diese maximalen und minimalen Attributsausprägungen normalisiert und stellen die normalisierten maximalen bzw. minimalen Attributsausprägungen dar.

Die Kodierung kann unipolar zwischen 0 und 1, bipolar zwischen -1 und 1 und frei gesetzt werden.

Objekte und Eingabe-Vektoren sind Unternehmen also Akquiseziele und Nicht-Akquiseziele. In dem Fall von Objekten werden Akquiseziele und Nicht-Akquiseziele als Referenztypen dargestellt.² Das heißt, die verfügbaren Daten werden zunächst analog zur logistischen Regression zufällig in Daten zur Modellbildung und Daten zur Modellvalidierung aufgeteilt. Die Modellbildungsdaten werden dann zur Bildung von Referenztypen verwen-

¹ Vgl. KLÜVER et al. (2015), S. 139 ff.

² Vgl. KLÜVER et al. (2012), S. 142.

det. In der vorliegenden Arbeit wurden die Modellbildungsdaten in Akquiseziel und Nicht-Akquiseziel aufgeteilt. Im Anschluss wurden Mittelwerte für alle Finanzkennzahlen zum einen für die Akquiseziele und zum anderen für die Nicht-Akquiseziele gebildet. Die Mittelwerte für Akquiseziele und Nicht-Akquiseziele stellen die Referenztypen und somit die zwei Objekte der vorliegenden KNN-Stahl-Prognosemodelle dar. Diese Attributsausprägungen werden auf die gewählte Kodierung unter Berücksichtigung der Attributsausprägungsgrenzen normalisiert¹ und ggf. mit dem cue validity factor (cfv) multipliziert.

Eingabe-Vektoren sind ebenfalls Unternehmen, sind aber im Vergleich zu Objekten keine Referenztypen. Die Eingabe-Vektoren sind alle Unternehmen der Modellbildungs- und Modellvalidierungsdaten. Sie werden ebenfalls durch die festgelegten Attribute beschrieben. Diese Attributsausprägungen werden auf die gewählte Kodierung unter Berücksichtigung der Attributsausprägungsgrenzen normalisiert und ggf. mit dem cfv multipliziert.

Der cfv ist ein Multiplikator. Durch den cfv ist eine Gewichtung der Finanzkennzahlen möglich. Eine unterschiedliche Gewichtung von einzelnen Finanzkennzahlen, also Attributen, sollte aber begründet geschehen.

Unter Berücksichtigung der ausgewählten Lernregel, der Lernrate und der Lernschritte² wird aus der normalisierten und gewichteten semantischen Matrix, die die Objekte als Referenztypen für Akquiseziel und Nicht-Akquiseziel mit den jeweiligen Mittelwerten der Finanzkennzahlen beinhaltet, die Gewichtsmatrix mit Gewichtswerten errechnet. Diese Gewichtswerte dienen in Verbindung mit der Aktivierungsfunktion und den Ausprägungen der Finanzkennzahlen der Eingabe-Vektoren, also den Finanzkennzahlen der einzelnen Unternehmen, der Klassifikation der Unternehmen in Akquiseziele und Nicht-Akquiseziele.

¹ Die Normalisierung wird in Kapitel 2.1.2.2.5 der vorliegenden Arbeit beschrieben.

² Das Lernen im SEN wird in Kapitel 2.1.2.2.5 ausführlich beschrieben.

2.1.2.2.4 Festlegung der Netztopologie

Das SEN besteht aus drei gekoppelten Teilen. Der erste Teil besteht aus dem *Attribute-Editor*, der *semantischen Matrix* und dem *Eingabe-Vektor*.¹ Der zweite Teil des SEN ist das eigentliche Netzwerk.² Hierbei handelt es sich um ein einschichtiges rekurrentes Netzwerk, in dem alle Neuronen miteinander verbunden sein können (externe Topologie).³ Die Anzahl der Neuronen entspricht der Anzahl der Objekte und Attribute. Für die vorliegende Arbeit wird die Standardeinstellung des SEN, die Feed-Forward-Topologie, verwendet.⁴ Die Festlegung der Beziehungen zwischen Attributen und Objekten in der semantischen Matrix bestimmt die interne Topologie.⁵ Der dritte Teil des SEN besteht aus den Ergebnisansichten.⁶ Diese Ergebnisansichten stellen zusammen den Visualisierungsteil dar.⁷

2.1.2.2.5 Informationsverarbeitung in den Neuronen

Der *Attribute-Editor* legt den Rahmen der Normalisierung fest und dient zum Anlegen, Entfernen und Bearbeiten von Attributen. In der *kompakten Ansicht* ist es möglich, den Namen eines Attributs, die minimale Attributsausprägung (r_{min}), die maximale Attributsausprägung (r_{max}) und einen Standard-Wert (r_{def}) der Attribute anzulegen.⁸ Zudem kann in der *kompakten Ansicht* im *Attribute-Editor* die Kodierung für die normalisierte Attributsausprägung (n_{min} und n_{max}) festgelegt werden.⁹ Die Kodierung gibt das Intervall an, auf

¹ Vgl. KLÜVER et al. (2012), S. 140.

² Vgl. KLÜVER et al. (2012), S. 140.

³ Vgl. KLÜVER et al. (2012), S. 140.

⁴ Erläuterungen zu der Festlegung der externen Topologie folgen in Kapitel 3.3.2.

⁵ Vgl. KLÜVER et al. (2012), S. 140. Die interne Topologie kann in der Gewichtsmatrix angepasst werden. Es ist möglich, Verbindungen zwischen Objekten und Attributen, Objekten und Objekten sowie Attributen und Attributen zu definieren.

⁶ Vgl. COBASC (2014), o.S.

⁷ Vgl. KLÜVER et al. (2012), S. 142.

⁸ Vgl. COBASC (2014), o.S.

⁹ Vgl. COBASC (2014), o.S.

das die Attributsausprägungen normalisiert werden.¹ Es besteht die Möglichkeit der uni- und bipolaren Kodierung.² Bei der unipolaren Kodierung wird die minimale Attributsausprägung (r_{min}) auf 0 (n_{min}) und die maximale Attributsausprägung (r_{max}) auf 1 (n_{max}) gesetzt.³ Bei der bipolaren Kodierung wird die minimale Attributsausprägung (r_{min}) auf -1 (n_{min}) und die maximale Attributsausprägung (r_{max}) auf 1 (n_{max}) gesetzt.⁴ Abbildung 14 zeigt die *kompakte Ansicht* des *Attribute-Editors*.



Name	Standard	Minimum	Maximum	Kodierung
Return on Equity	0,00	-0,91	1,11	[-1; 1]
Free Cash Flow /...	0,00	-0,27	0,34	[-1; 1]
Operating Profit ...	0,00	-2,39	0,57	[-1; 1]
Asset Turnover	0,31	0,00	3,41	[0; 1]
Earnings Growth	-7,14	-8,08	8,79	[-1; 1]
Earnings Before ...	-808,00	-1.587,07	29.609,10	[-1; 1]
Return on Capita...	-0,05	-0,11	0,90	[-1; 1]
Earnings Before ...	-0,54	-1,59	0,97	[-1; 1]
Earnings Before ...	-8,09	-8,09	14,53	[-1; 1]
Sales Growth	-0,63	-0,84	1,94	[-1; 1]
Net Liquid Asset...	0,02	0,00	0,58	[-1; 1]
Long-Term Debt	0,00	6,81	5,17	[-1; 1]

Abbildung 14: Kompakte Ansicht des Attribute-Editors des SEN (Ausschnitt)⁵

In der *Experten-Ansicht* werden die Eingabemöglichkeiten erweitert. Zu den beschriebenen Funktionalitäten kommen weitere Eingabemöglichkeiten hinzu. Es können frei wählbare Kodierungsgrenzen gesetzt werden.⁶ Zusätzlich ist die Eingabe eines cue validity factors möglich. Der cue validity factor (cvf) ist ein Multiplikator und bestimmt den Einfluss, den jedes einzelne Attribut auf das Endergebnis hat.⁷ Je höher der cue validity factor, desto

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

³ Vgl. COBASC (2014), o.S.

⁴ Vgl. COBASC (2014), o.S.

⁵ Vgl. COBASC (2014), o.S.

⁶ Vgl. COBASC (2014), o.S.

⁷ Vgl. COBASC (2014), o.S.

wichtiger das entsprechende Attribut. Abbildung 15 zeigt die *Experten-Ansicht* des *Attribute-Editors*.

Name	r _{def}	r _{min}	r _{max}	n _{min}	n _{max}	cvf	v _{min}	v _{max}
Return o...	0,00	-0,91	1,11	-1,00	1,00	0,60	-0,60	0,60
Free Ca...	0,00	-0,27	0,34	-1,00	1,00	1,00	-1,00	1,00
Operatin...	0,00	-2,39	0,57	-1,00	1,00	1,00	-1,00	1,00
Asset T...	0,31	0,00	3,41	0,00	1,00	0,60	0,00	0,60
Earning...	-7,14	-8,08	8,79	-1,00	1,00	1,00	-1,00	1,00
Earning...	-808,00	-1.587,07	29.609,10	-1,00	1,00	1,00	-1,00	1,00
Return o...	-0,05	-0,11	0,90	-1,00	1,00	0,60	-0,60	0,60
Earning...	-0,54	-1,59	0,97	-1,00	1,00	1,00	-1,00	1,00
Earning...	-8,09	-8,09	14,53	-1,00	1,00	1,00	-1,00	1,00
Sales G...	-0,63	-0,84	1,94	-1,00	1,00	1,00	-1,00	1,00
Net Liqu...	0,02	0,00	0,58	-1,00	1,00	1,50	-1,50	1,50

Abbildung 15: Experten Ansicht des Attribute-Editors des SEN (Ausschnitt)¹

Nach erfolgreicher Abbildung der Attribute erfolgt die Generierung von Objekten. Ein Standard-Objekt kann mittels der im *Attribute-Editor* festgelegten Standard-Werte in der *semantischen Matrix* durch Hinzufügen eines neuen Objektes generiert werden.

Die Festlegung der logisch-semantischen Beziehungen zwischen Attributen und Objekten in der *semantischen Matrix* bestimmt die interne Topologie.² Die Beziehungen können binär codiert sein oder reell, wobei letzteres die Möglichkeit zur Bewertung der Stärke der Beziehung bietet.³ Die *semantische Matrix* bildet die Datenbasis des SEN und besteht aus drei Ansichten.

Nach dem Anlegen von Objekten in der Ansicht *Rohdaten* können die Attributsausprägungen angepasst werden, um Referenztypen zu modellieren.⁴ Referenztypen (im vorliegenden Fall zwei, Akquiseziel und Nicht-Akquiseziel) dienen der Klassifikation der Modell-

¹ Vgl. COBASC (2014), o.S.

² Vgl. KLÜVER et al. (2012), S. 140.

³ Vgl. KLÜVER et al. (2012), S. 140.

⁴ Vgl. COBASC (2014), o.S.

bildungsdaten in Akquiseziel und Nicht-Akquiseziel.¹ Abbildung 16 zeigt die Ansicht *Rohdaten* der *semantischen Matrix*.

Objekt...	Retu...	Free...	Oper...	Asse...	Earni...	Earni...	Retu...	Earni...	Earni...	Sale...	Net L...	Long...	Sho...
Akqu...	0,19	0,02	0,07	1,03	-0,26	235...	0,09	0,12	0,89	0,11	0,23	0,57	0,4
Nicht...	0,22	0,03	0,10	1,07	0,14	1,31...	0,10	0,18	1,09	0,20	0,21	0,49	0,3

Abbildung 16: Rohdaten-Ansicht der semantischen Matrix (Ausschnitt)²

In der Ansicht *Normalisiert* werden die Daten, die in ihrem natürlichen Datenformat in der Ansicht *Rohdaten* eingegeben wurden, in dem zuvor im *Attribute-Editor* festgelegtem Kodierungsintervall ($n_{\min}$ und $n_{\max}$) normalisiert dargestellt.³ Die Normalisierung der Attributsausprägungen erfolgt nach der folgenden Berechnungsvorschrift:⁴

$$v_{\text{norm}} = \left[\frac{v_{\text{raw}} - r_{\min}}{r_{\max} - r_{\min}} * (n_{\max} - n_{\min}) \right] - n_{\min} \quad (2.3)$$

mit

v_{norm} = normalisierte Attributsausprägung ($v_{\text{norm}} \in \mathbb{R}$)

v_{raw} = tatsächliche Attributsausprägung ($v_{\text{raw}} \in \mathbb{R}$)

$r_{\min}$ = minimale Attributsausprägung ($r_{\min} \in \mathbb{R}$)

$r_{\max}$ = maximale Attributsausprägung ($r_{\max} \in \mathbb{R}$)

¹ Vgl. KLÜVER et al. (2012), S. 142.

² Vgl. COBASC (2014), o.S.

³ Vgl. COBASC (2014), o.S.

⁴ Vgl. COBASC (2014), o.S.

n_{min} = normalisierte minimale Attributsausprägung ($n_{min} \in \mathbb{R}$)

n_{max} = normalisierte maximale Attributsausprägung ($n_{min} \in \mathbb{R}$)

Abbildung 17 zeigt die Ansicht *Normalisiert* der *semantischen Matrix*.

Objekte...	Retu...	Free...	Oper...	Asse...	Earni...	Earni...	Retu...	Earni...	Earni...	Sale...	Net L...	Long...	Sho...
Akqu...	0,09	-0,05	0,66	0,30	-0,07	-0,88	-0,60	0,34	-0,21	-0,32	-0,21	0,23	-0,88
Nicht...	0,12	-0,02	0,68	0,31	-0,03	-0,81	-0,58	0,38	-0,19	-0,25	-0,28	0,22	-0,88

Abbildung 17: Normalisierte Ansicht der semantischen Matrix (Ausschnitt)¹

In der Ansicht *Gewichtet* werden die normalisierten Attributsausprägungen mit dem cue validity factor (*cvf*) multipliziert dargestellt.² Die Daten, die in der Ansicht *Gewichtet* dargestellt werden, sind die eigentlichen Daten der *semantischen Matrix*. Die Berechnungsvorschrift lautet wie folgt:³

$$v_{sm} = v_{norm} * cvf \quad (2.4)$$

mit

v_{sm} = Wert aus der semantischen Matrix ($v_{sm} \in \mathbb{R}$)

v_{norm} = normalisierter Attributswert ($v_{norm} \in \mathbb{R}$)

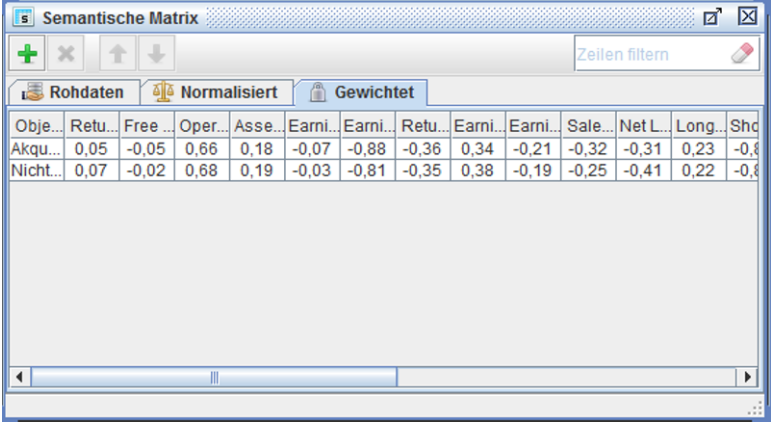
cvf = cue validity factor ($cvf \in \mathbb{R}$)

Abbildung 18 zeigt die Ansicht *Gewichtet* der *semantischen Matrix*.

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

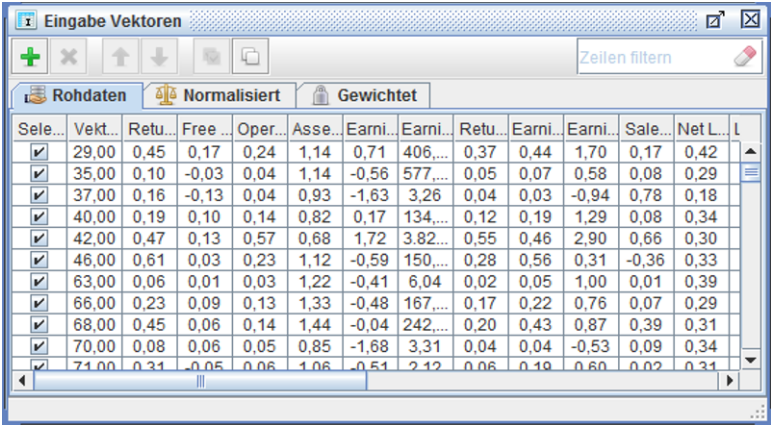
³ Vgl. COBASC (2014), o.S.



Objekte	Retu...	Free...	Oper...	Asse...	Earni...	Earni...	Retu...	Earni...	Earni...	Sale...	Net L...	Long...	Sho...
Akqu...	0,05	-0,05	0,66	0,18	-0,07	-0,88	-0,36	0,34	-0,21	-0,32	-0,31	0,23	-0,8
Nicht...	0,07	-0,02	0,68	0,19	-0,03	-0,81	-0,35	0,38	-0,19	-0,25	-0,41	0,22	-0,8

Abbildung 18: Gewichtete Ansicht der semantischen Matrix (Ausschnitt)¹

Das Fenster *Eingabe-Vektoren* dient zur Erfassung der Eingabe-Vektoren, die im Rahmen des Lernprozesses den Objekten zugeordnet werden. Hier werden die Modellbildungs- und Modellvalidierungsdaten abgebildet. Die Erfassung der *Eingabe-Vektoren* erfolgt analog zu der Erfassung der Objekte in der *semantischen Matrix*. Abbildung 19 zeigt das Fenster *Eingabe-Vektor*.



Sele...	Vekt...	Retu...	Free...	Oper...	Asse...	Earni...	Earni...	Retu...	Earni...	Earni...	Sale...	Net L...	L...
☒	29,00	0,45	0,17	0,24	1,14	0,71	406...	0,37	0,44	1,70	0,17	0,42	
☒	35,00	0,10	-0,03	0,04	1,14	-0,56	577...	0,05	0,07	0,58	0,08	0,29	
☒	37,00	0,16	-0,13	0,04	0,93	-1,63	3,26	0,04	0,03	-0,94	0,78	0,18	
☒	40,00	0,19	0,10	0,14	0,82	0,17	134...	0,12	0,19	1,29	0,08	0,34	
☒	42,00	0,47	0,13	0,57	0,68	1,72	3,82...	0,55	0,46	2,90	0,66	0,30	
☒	46,00	0,61	0,03	0,23	1,12	-0,59	150...	0,28	0,56	0,31	-0,36	0,33	
☒	63,00	0,06	0,01	0,03	1,22	-0,41	6,04	0,02	0,05	1,00	0,01	0,39	
☒	66,00	0,23	0,09	0,13	1,33	-0,48	167...	0,17	0,22	0,76	0,07	0,29	
☒	68,00	0,45	0,06	0,14	1,44	-0,04	242...	0,20	0,43	0,87	0,39	0,31	
☒	70,00	0,08	0,06	0,05	0,85	-1,68	3,31	0,04	0,04	-0,53	0,09	0,34	
☒	71,00	0,31	-0,05	0,06	1,06	-0,51	2,12	0,06	0,19	0,60	0,02	0,31	

Abbildung 19: Eingabe-Vektoren des SEN (Ausschnitt)²

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

Der zweite Teil des SEN, die Gewichtsmatrix, ist das eigentliche Netzwerk und stellt die interne Topologie dar.¹ Hierbei handelt es sich um ein nicht geschichtetes rekurrentes Netzwerk, wo alle Neuronen miteinander verbunden sein können (externe Topologie).² Die Anzahl der Neuronen entspricht der Anzahl der Objekte und Attribute. Die *Gewichtsmatrix* bietet eine *kompakte Ansicht* und eine *Experten-Ansicht*.

In der *kompakten Ansicht* werden nur die Gewichte angezeigt, die durch die Anwendung der Lernregel auf die *semantische Matrix* generiert werden. Dies sind die Gewichtswerte $[w_{ij}(t)]$ von Attributen des Netzes zu den jeweiligen Objekten. Die *Gewichtsmatrix* enthält zu Beginn nur Gewichtswerte w_{ij} zwischen einem sendenden Neuron i und einem empfangenden Neuron j von 0 $[w_{ij}(t) = 0]$.³ Die Gewichtswerte werden durch folgende Lernregel (Self-Enforcing Rule) generiert:⁴

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \text{ und} \quad (2.5)$$

$$\Delta w_{ij} = c * v_{sm} \quad (2.6)$$

mit

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

t = Zeitpunkt ($t = 0, 1, 2, \dots, T$)

c = Lernrate ($c \in \mathbb{R}^{>0}$)

v_{sm} = Wert aus der semantischen Matrix ($v_{sm} \in \mathbb{R}$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

Zusätzlich ist die Eingabe der Lernrate und der Anzahl der Lernschritte möglich. Die Lernrate gibt das Maß an, in dem die Gewichte der *Gewichtsmatrix* in einem Lernschritt beeinflusst werden. Die Lernschritte geben an, wie viele Wiederholungen des Lernprozesses durchgeführt werden sollen. Durch die vorgegebene Lernrate und die *semantische Matrix*

¹ Vgl. KLÜVER et al. (2012), S. 140.

² Vgl. KLÜVER et al. (2012), S. 140.

³ Vgl. KLÜVER et al. (2012), S. 141.

⁴ Vgl. KLÜVER et al. (2012), S. 141. Die Basis dieser Lernregel ist das allgemeine Lernschema (General Enforcing Rule Schema; $\Delta w_{ij} = \pm c * |1 - w_{ij}(t)|$)

wird die *Gewichtsmatrix* generiert. Standardmäßig ist im SEN die Feed-Forward-Topologie eingestellt.¹ Dies führt dazu, dass keine Gewichte zwischen Objekten und Objekten, Objekten und Attributen sowie Attributen und Attributen definiert sind. Diese zuletzt genannten Gewichte sind in der *kompakten Ansicht* der *Gewichtsmatrix* nicht zu sehen. Abbildung 20 zeigt die *kompakte Ansicht* der *Gewichtsmatrix*.

	Retu...	Free ...	Oper...	Asse...	Earni...	Earni...	Retu...	Earni...	Earni...	Sale...	Net L...	Long...	Sho...
Akqu...	0,03	-0,02	0,33	0,09	-0,04	-0,44	-0,18	0,17	-0,10	-0,16	-0,16	0,12	-0,4
Nicht...	0,04	-0,01	0,34	0,09	-0,01	-0,41	-0,18	0,19	-0,09	-0,13	-0,21	0,11	-0,4

Abbildung 20: Kompakte Ansicht der Gewichtsmatrix (Ausschnitt)²

Die *Experten-Ansicht* gibt die Möglichkeit, die vollständige *Gewichtsmatrix* zu bearbeiten. Durch die Eingabe von Gewichten zwischen Attributen und Attributen, Objekten und Attributen sowie Objekten und Objekten ist es so möglich, ein Feed-Backward-Netz oder ein rekurrentes Netz zu entwickeln. Abbildung 21 zeigt die *Experten-Ansicht* der *Gewichtsmatrix*.

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

	Akqu...	Nicht...	Retu...	Free ...	Oper...	Asse...	Earni...	Earni...	Retu...	Earni...	Earni...	Sale...	
Akqu...	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Nicht...	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Retu...	0,03	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Free ...	-0,02	-0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Oper...	0,33	0,34	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Asse...	0,09	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Earn...	-0,04	-0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Earn...	-0,44	-0,41	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Retu...	-0,18	-0,18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Earn...	0,17	0,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
Earn...	-0,10	-0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	

Abbildung 21: Experten-Ansicht der Gewichtsmatrix (Ausschnitt)¹

Nach der Berechnung der Gewichtswerte werden diese in der Aktivierungsfunktion² verwendet. Als Aktivierungsfunktion stehen im SEN die bereits in Kapitel 2.1.2.1.1 eingeführte lineare Aktivierungsfunktion (2.1) sowie fünf weitere Funktionen zur Verfügung.

1. Lineare Aktivierungsfunktion:³

$$a_j = \sum_{i=1}^I w_{ij} * a_i \quad (2.2)$$

mit

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($i = 1, 2, \dots, I$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

¹ Vgl. COBASC (2014), o.S.

² Das SEN verzichtet auf eine Schwellenwertfunktion. Vgl. COBASC (2014), o.S.

³ Vgl. COBASC (2014), o.S.

2. Tangens Hyperbolicus:¹

$$a_j = \tanh(\text{net}_j) \quad (2.7)$$

mit

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

$$\text{net}_j = \sum_{i=1}^I w_{ij} * a_i$$

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($i = 1, 2, \dots, I$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

3. Logistische Funktion:²

$$a_j = \frac{1}{1 + e^{-\text{net}_j}} \quad (2.8)$$

mit

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

e = 2,71828183 (Eulersche Zahl)

$$\text{net}_j = \sum_{i=1}^I w_{ij} * a_i$$

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($i = 1, 2, \dots, I$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

Die lineare Aktivierungsfunktion erleichtert das Verständnis des Netzwerkprozesses wesentlich, hat aber einen Nachteil.³ Sind in der *semantischen Matrix* viele 1-er enthalten,

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

³ Vgl. KLÜVER et al. (2012), S. 140.

eine hohe Lernrate und eine hohe Anzahl an Lernschritten gewählt wird, der Aktivierungswert sehr schnell groß. Um dieses Problem zu lösen, sind folgende Aktivierungsfunktionen im SEN vorhanden:¹

4. Lineare Mittelwertfunktion:

$$a_j = \frac{\sum_{i=1}^I w_{ij} * a_i}{I} \quad (2.9)$$

mit

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($i = 1, 2, \dots, I$)

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

Bei vorgegebenen Eingangssignalen a_i zwischen 0 und 1 bleibt a_j unter 1 und nur in Extremfällen zwischen 1 und 2.²

5. Logarithmisch lineare Funktion:

$$a_j = \sum_{i=1}^I \begin{cases} \log_3(1 + a_i) * w_{ij} & \text{wenn } a_i \geq 0 \\ \log_3(|1 + a_i|) * -w_{ij} & \text{sonst} \end{cases} \quad (2.10)$$

mit

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($i = 1, 2, \dots, I$)

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

¹ Vgl. COBASC (2014), o.S.

² Vgl. KLÜVER et al. (2012), S. 141.

Bei einem Eingangssignal a_i von 0 bis 1 werden die logarithmischen Werte negativ, aus diesem Grund ist der Summand +1 hinzugefügt. Dies ergibt Werte zwischen 1 und 2. Um diese Werte zwischen 0 und 1 einzugrenzen, wird als logarithmische Basis die 3 verwendet.¹

6. Enforcing Activation Function:

$$a_j = \frac{\sum_{i=1}^I w_{ij} * a_i}{1 + \left| \sum_{i=1}^I w_{ij} * a_i \right|} \quad (2.11)$$

mit

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

a_i = Eingangssignal des sendenden Neurons i ($i = 1, 2, \dots, I$)

w_{ij} = Gewichtswert zwischen sendendem Neuron i und empfangendem Neuron j ($w_{ij} \in \mathbb{R}$
 $\forall i = 1, 2, \dots, I \wedge j = 1, 2, \dots, J$)

n_{min} = normalisierte minimale Attributsausprägung ($n_{min} \in \mathbb{R}$)

n_{max} = normalisierte maximale Attributsausprägung ($n_{max} \in \mathbb{R}$)

I = Anzahl sendender Neuronen

J = Anzahl empfangender Neuronen

Unter Berücksichtigung der elementaren Zusammenhänge zwischen KNN-Stahl-Prognosemodellen und dem SEN (siehe Kapitel 2.1.2.2.3) sowie weiterer Konkretisierungen der Elemente der o.g. Funktionen soll im Folgenden das Verständnis für die Modellbildung, die in Kapitel 3.3.2 erfolgt, geschaffen werden.²

Um Aktivierungswerte (a_j) für die Akquiseziele und Nicht-Akquiseziele (hier: Eingabevektoren) zu erhalten, müssen zunächst die Gewichtsmatrix und die Aktivierungswerte (a_j) für die Referenztypen der Akquiseziele und Nicht-Akquiseziele generiert werden. Hierfür sind die Finanzkennzahlausprägungen der Referenztypen (v_{raw}) notwendig. Diese Finanzkennzahlausprägungen werden nach der gewählten Vorschrift normalisiert. Hierfür wird die Funktion 2.3 verwendet:

¹ Vgl. KLÜVER et al. (2012), S. 141.

² Hierbei wird aus Übersichtsgründen die lineare Funktion als Aktivierungsfunktion ausgewählt. Die integrierte Funktion sieht bei der Auswahl einer anderen Aktivierungsfunktion anders aus. Die konkrete Auswahl der Aktivierungsfunktion für das KNN-Stahlprognosemodell erfolgt in Kapitel 3.3.2.

$$v_{norm} = \left[\frac{v_{raw} - r_{min}}{r_{max} - r_{min}} * (n_{max} - n_{min}) \right] - n_{min} \quad (2.3)$$

Die normalisierten Finanzkennzahlausprägungen (v_{norm}) werden mit den *cvf* gewichtet. Das Ergebnis dieser Gewichtung sind die Werte für die *semantische Matrix* (v_{sm}).

$$v_{sm} = v_{norm} * cvf \quad (2.4)$$

Unter Berücksichtigung der Lernrate (c) erfolgt die Berechnung der Änderung der Gewichtswerte (Δw_{ij}).

$$\Delta w_{ij} = c * v_{sm} \quad (2.6)$$

Für die Berechnung der *Gewichtsmatrix*, die zum Anfangszeitpunkt ($t=0$) keine Gewichtswerte (w_{ij}) enthält wird die Formel 2.4 verwendet.

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \quad (2.5)$$

Ist die Gewichtsmatrix generiert, können unter Berücksichtigung der Aktivierungsfunktion, Aktivierungswerte (a_j) für die Referenztypen abgebildet werden.

$$a_j = \sum_{i=1}^I w_{ij} * a_i \quad (2.2)$$

Durch die Anwendung der generierten Gewichtsmatrix und der ausgewählten Aktivierungsfunktion werden sodann Aktivierungswerte (a_j) für die Akquiseziele und Nicht-Akquiseziele (hier: *Eingabe-Vektoren*) ermittelt. Die Eingangssignale (a_i) sind die Finanzkennzahlausprägungen der Akquiseziele und Nicht-Akquiseziele (hier: *Eingabe-Vektoren*). Die Klassifikation der *Eingabe-Vektoren* zu Referenztypen erfolgt anhand der ermittelten Aktivierungswerte (a_j).¹

Der dritte Teil des SEN sind die Ergebnisansichten.² Diese Ergebnisansichten stellen zusammen den Visualisierungsteil dar.³ Dieser ist für die Visualisierung der Ähnlichkeiten der einzelnen Objekte verantwortlich. Es stehen drei Visualisierungen zur Verfügung.¹

¹ Bei der Klassifikation eines Unternehmens als Akquiseziel oder Nicht-Akquiseziel ist eine Klassifikationsvorschrift festzulegen. Für die vorliegenden KNN-Stahl-Prognosemodelle gilt: Eingabe-Vektoren werden dem Referenztyp mit dem höchsten Aktivierungswert zugeordnet.

² Vgl. COBASC (2014), o.S.

³ Vgl. KLÜVER et al. (2012), S. 142.

1. Karten-Visualisierung

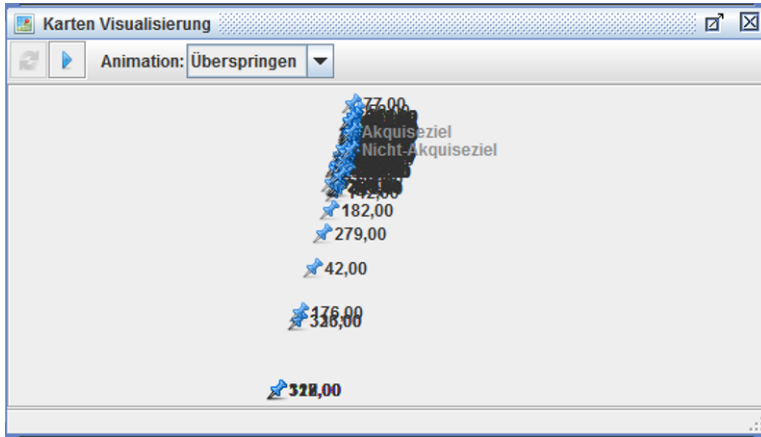


Abbildung 22: Karten-Visualisierung des SEN²

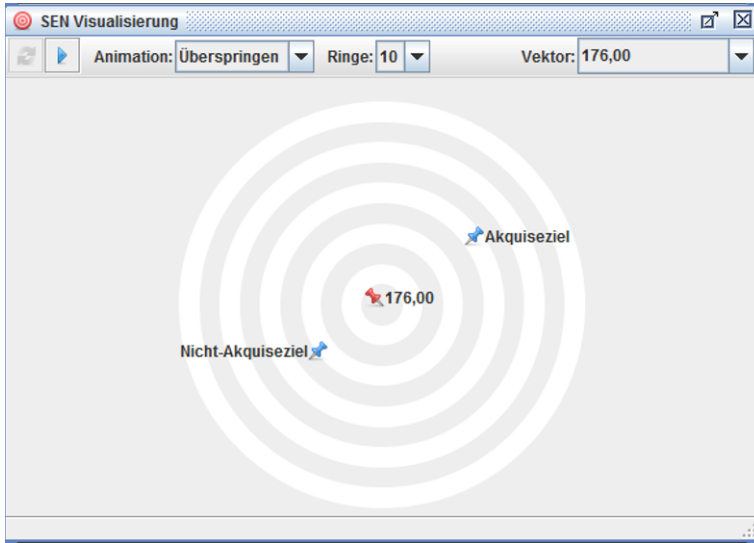
Die *Karten-Visualisierung* stellt alle Eingabe-Vektoren im Verhältnis zu den Objekten dar. Je ähnlicher die Eingabe-Vektoren und Objekte sind, desto näher zueinander werden sie platziert und umgekehrt.³

¹ In der aktuellen Version des SEN stehen vier Visualisierungen zur Verfügung.

² Vgl. COBASC (2014), o.S.

³ Vgl. COBASC (2014), o.S.

2. SEN-Visualisierung

Abbildung 23: SEN-Visualisierung¹

In der *SEN-Visualisierung* können einzelne Eingabe-Vektoren ausgewählt werden, die im Anschluss im Zentrum der konzentrischen Kreise abgebildet werden. Die Objekte werden in gleichen Winkelabständen um den Eingabe-Vektor herum angeordnet. Diese Darstellung dient zur Betrachtung einzelner Eingabe-Vektoren und um festzustellen, wie groß die Ähnlichkeit des betrachteten Eingabe-Vektors zu den einzelnen Objekten ist. Die Abstände des Eingabe-Vektors zu den Objekten werden mit der folgenden Formel ermittelt:²

$$r_j = \begin{cases} 1 - a_{rel}, & \text{wenn } a_{rel} \geq 0 \\ 1 - \frac{\tanh(a_{rel})}{2}, & \text{sonst} \end{cases} \quad (2.12)$$

mit

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

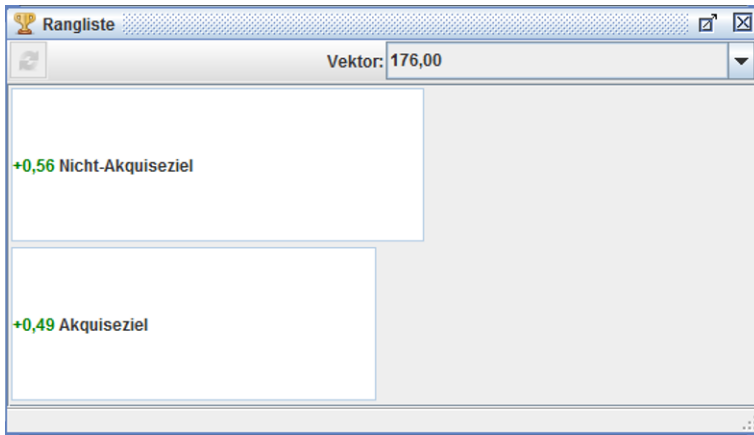
r_j = Winkelabstand ($n_{min} \in \mathbb{R}^{>0}$)

$$a_{rel} = \frac{a_j}{a_{max}}$$

a_j = Aktivierungswert des empfangenden Neurons j ($j = 1, 2, \dots, J$)

a_{max} = maximaler Aktivierungswert aller empfangenden Neuronen

J = Anzahl empfangender Neuronen



3. Rangliste

Abbildung 24: Rangliste des SEN¹

Die *Rangliste* stellt für den ausgewählten Eingabe-Vektor die Aktivierungswerte (a_j) der Objekte in absteigender Reihenfolge in Balkenform dar.² Je länger der Balken, desto ähnlicher ist der ausgewählte Eingabe-Vektor (in diesem Fall Eingabe-Vektor 319) mit dem angezeigten Objekt. Zusätzlich zu der Visualisierung werden die Ergebnisse tabellarisch dargestellt.

¹ Vgl. COBASC (2014), o.S.

² Vgl. COBASC (2014), o.S.

2.2 Mergers & Acquisitions-Hypothesen zur Identifizierung von Mergers & Acquisitions-Kandidaten

Zwischen 1971 und 1984 sind vermehrt empirische Untersuchungen mit öffentlich zugänglichen Finanzkennzahlen zur Vorhersage von Akquisezielen und Nicht-Akquisezielen erfolgt.¹ Diese Untersuchungen hatten als Ergebnis Vorhersage-Modelle auf Basis statistischer Modelle, die bei der Klassifikation 65% bis 92% und bei der Validierung 63% bis 70% der betrachteten Unternehmen korrekt zuordneten.² Erst durch die Untersuchungen von PALEPU (1986) wurde deutlich, dass diese vorangegangenen Untersuchungen fundamentale Fehler bei der Auswahl des Umfangs der Datensamples hatten, denn typischer Weise wurden in diesen Arbeiten gleich große Samples an Akquisezielen und Nicht-Akquisezielen gewählt.³ Erst durch die Systematisierung des Auswahlverfahrens für den Umfang der Datensamples wurde dieser Fehler behoben.⁴ Nicht nur wegen der Formalisierung des Auswahlverfahrens ist die Untersuchung von PALEPU (1986) von hoher Bedeutung, denn PALEPU (1986) ist der Erste, der basierend auf Hypothesen zu Übernahmen und Fusionen Attribute bestimmt, mit denen die Vorhersage von Akquisezielen erfolgt.⁵ Dies hat den Vorteil, dass die betrachteten Attribute nicht willkürlich ausgewählt werden.

In der vorliegenden Arbeit finden daher die von PALEPU (1986) betrachteten Hypothesen zu Übernahmen und Fusionen Berücksichtigung. Diese Hypothesen sind bis heute die zent-

¹ Vgl. SIMKOWITZ et al. (1971), S. 1 ff.; STEVENS (1973), S. 149 ff.; BELKAOUI (1978), S. 94; DIETRICH et al. (1984), S. 397.

² Vgl. BARNES (1998), S. 574.

³ STEVENS (1973), S. 150 hatte z. B. zur Entwicklung des Modells 40 Akquiseziele und 40 Nicht-Akquiseziele und zur Validierung 20 Akquiseziele und 20 Nicht-Akquiseziele ausgewählt.

⁴ Hierbei muss eine zufällige Auswahl an Akquisezielen und Nicht-Akquisezielen, die im anzahlmäßigen Verhältnis zum am gesamten Markt herrschenden Verhältnis an Akquisezielen und Nicht-Akquisezielen stehen, erfolgen. Vgl. PALEPU (1986), S. 10 ff.

⁵ In den Untersuchungen von SIMKOWITZ et al. (1971), S. 1 ff.; STEVENS (1973), S. 149 ff.; BELKAOUI (1978), S. 95; DIETRICH et al. (1984), S. 393 ff. betrachten die Autoren populäre Finanzkennzahlen, wobei die Auswahl willkürlich geschieht, und reduzieren diese anhand ihrer statistischen Signifikanz auf eine final zu betrachtende Anzahl an Finanzkennzahlen. Bspw. starteten SIMKOWITZ et al. (1971) mit 24 Finanzkennzahlen und reduzierten diese auf 7. Vgl. SIMKOWITZ et al. (1971), S. 8.

ralen Elemente bei dem Versuch, M&As mittels Finanzkennzahlen zu prognostizieren.¹ Um neuere Erkenntnisse, die seit der empirischen Untersuchung von PALEPU (1986) gewonnen wurden, zu berücksichtigen, werden weitere Finanzkennzahlen, die sich als relevant erwiesen haben, als unabhängige Variablen den Hypothesen zu Übernahmen und Fusionen zugeordnet.

2.2.1 Hypothese des ineffizienten Managements

Die „inefficient management hypothesis“ wird als Hypothese des ineffizienten Managements beschrieben. Ein ineffizientes Management wird definiert als ein Management, welches nicht in der Lage ist, den Shareholder Value zu steigern, oder seine eigenen Interessen auf Kosten der Aktionäre steigert.² Die Hypothese des ineffizienten Managements basiert auf den Annahmen,

1. dass die „schlechte Leistung“ eines Unternehmens auf ein Management, welches rücksichtslos, ineffizient und auf sich selbst fokussiert ist, zurückgeführt werden kann und,³
2. dass das Management nach einer Übernahme durch die neuen Eigentümer diszipliniert oder durch ein neues Management ersetzt werden kann.⁴

Somit können nicht ausgeschöpfte Opportunitäten zur Kostenreduzierung sowie Umsatz- und Gewinnsteigerung durch ein neues, effizienteres Management realisiert werden.⁵ Die „schlechte Leistung“ wird anhand verschiedener Finanzkennzahlen gemessen. Tabelle 4 zeigt eine Übersicht empirischer Untersuchungen zur Hypothese des ineffizienten Managements und den dabei betrachteten Finanzkennzahlen. Zusätzlich wird die Wirkungsbe-

¹ Vgl. HASBROUCK (1985), S. 361; PALEPU (1986), S. 30 f.; AMBROSE et al. (1992), S. 583 ff.; SONG et al. (1993), S. 456; MEADOR et al. (1996), S. 22 f.; BARNES (1998), S. 297 f.; BEŠTER (2000), S. 15; AGRAWAL et al. (2003), S. 742 ff.; DOUMPOS et al. (2004), S. 208 f.; TSAGKANOS et al. (2006), S. 189 f.; BASU et al. (2008), S. 216; UCER (2009), S. 50 ff.; ERDOGAN (2012), S. 75; BECCALLI et al. (2013), S. 268 ff.

² Vgl. MWALE (2007), S. 19.

³ Vgl. AGRAWAL et al. (1992), S. 1605 ff.; MWALE (2007), S. 10.

⁴ Vgl. WALSCH (1988), S. 180; MANNE (1965), S. 112; BECCALLI et al. (2013), S. 269.

⁵ Vgl. BREALEY et al. (2008), S. 887; WESTON et al. (1990), S. 250 f.

ziehung zwischen jeder Finanzkennzahl und der Übernahmewahrscheinlichkeit dargestellt. Ein „-“ bedeutet, dass bei einer größer werdenden Finanzkennzahl die Übernahmewahrscheinlichkeit sinkt, und umgekehrt bedeutet ein „+“ eine steigende Übernahmewahrscheinlichkeit, wenn die Finanzkennzahl größer wird. Ein „+/-“ bedeutet, dass keine Definierung der Wirkungsbeziehung erfolgt ist.

	Gross Value Added / Net Operating Profit																	✓
	Gross Value Added / Total Asset																	✓
	Market Capitalization / Equity					✓												
	Dividend Growth					✓												
	Dividend / Equity					✓												
	Share Price / Earnings					✓												
	Earnings Before Tax Growth					✓												
	Earnings Before Tax / Equity					✓												
	Return on Capital Employed				✓													
	Earnings Before Tax					✓												✓
	Earnings Growth										✓							
	Sales Growth										✓							
	Asset Turnover										✓							
	Operating Profit Margin										✓							
	Free Cash Flow / Total Assets										✓							
	Return on Equity	✓						✓			✓			✓				
	Average Excess Return	✓	✓						✓									
		PALEPU (1986)	AMBROSE et al. (1992)	POWELL (1997)	BARNES (1998)	CUDD et al. (2000)	MWALE (2007)	BRAR et al. (2009)	ERDOGAN (2012)									

Wahrscheinlichkeit der Übernahme ¹																		
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Tabelle 4: Empirische Untersuchungen zur Hypothese des ineffizienten Managements mit den betrachteten Finanzkennzahlen und Wirkungsbeziehungen¹

¹ Vgl. PALEPU (1986), S. 19; AMBROSE et al. (1992), S. 579; POWELL (1997), S. 1013; BARNES (1998), S. 579; CUDD et al. (2000), S. 107; MWALE (2007), S. 45; BRAR et al. (2009), S. 436; ERDOGAN (2012), S. 74.

2.2.2 Hypothese des Wachstums-Ressourcen-Ungleichgewichts

Die „growth-resource mismatch hypothesis“ ist in der vorliegenden Arbeit als Hypothese des Wachstums-Ressourcen-Ungleichgewichts formuliert.¹ Bei dieser Hypothese wird das Zusammenwirken der Wachstumsmöglichkeiten und der Finanzmittel eines Unternehmens betrachtet.² Impliziert wird eine Übernahmewahrscheinlichkeit bei zwei Typen von Unternehmen:

1. Unternehmen mit geringen Wachstumsmöglichkeiten, aber mit hohen Finanzmitteln³, sowie
2. Unternehmen mit hohen Wachstumsmöglichkeiten, aber mit geringen Finanzmitteln⁴,

Hierbei werden folgende Annahmen getroffen:

1. Die Finanzmittel des Akquiseziels können profitabler bei dem Akquiseakteur investiert werden und
2. die Wachstumsmöglichkeiten des Akquiseziels können günstiger durch den Akquiseakteur finanziert werden.⁵

Die „Wachstumsmöglichkeiten“ und die „Finanzmittel“ werden anhand verschiedener Finanzkennzahlen gemessen. Tabelle 5 zeigt eine Übersicht empirischer Untersuchungen zur Hypothese des Wachstums-Ressourcen-Ungleichgewichts und der dabei betrachteten Finanzkennzahlen. Zusätzlich wird die Wirkungsbeziehung zwischen jeder Finanzkennzahl und der Übernahmewahrscheinlichkeit dargestellt. Ein „-“ bedeutet, dass bei einer größer werdenden Finanzkennzahl die Übernahmewahrscheinlichkeit sinkt, und umgekehrt bedeutet ein „+“ eine steigende Übernahmewahrscheinlichkeit, wenn die Finanzkennzahl größer wird. Ein „+/-“ bedeutet, dass keine Definierung der Wirkungsbeziehung erfolgt ist.

¹ Die „financial leverage hypothesis“ und die „liquidity hypothesis“ werden unter dieser Hypothese zusammengefasst. Einige Autoren betrachten den Verschuldungsgrad und die Liquidität separat. Vgl. DOUMPOS et al. (2004), S. 201; ERDOGAN (2012), S. 73. Aus Sicht des Autors sind diese Attribute stark voneinander abhängig und werden aus diesem Grund zusammen betrachtet.

² Vgl. PALEPU (1986), S. 17; MEADOR et al. (1996), S. 12; ERDOGAN (2012), S. 73.

³ Vgl. PALEPU (1986), S. 17; MEADOR et al. (1996), S. 12, ERDOGAN (2012), S. 73.

⁴ Vgl. PALEPU (1986), S. 17; ERDOGAN (2012), S. 73.

⁵ Vgl. PALEPU (1986), S. 17; CUDD et al. (2000), S. 107; ERDOGAN (2012), S. 73.

2.2.3 Größenhypothese

Die als „size hypothesis“ formulierte Hypothese ist in der vorliegenden Arbeit als Größenhypothese beschrieben. Die „Größe“ wird definiert als die relative Größe eines Unternehmens in der eigenen Branche und kann grundsätzlich mit Hilfe des Anlagevermögens definiert werden.¹ Bei der Größenhypothese wird argumentiert, dass die Wahrscheinlichkeit einer Übernahme mit der Größe des Unternehmens abnimmt.² Die Größenhypothese basiert auf den Annahmen,

1. dass die Transaktionskosten einer Übernahme abhängig von der Größe des Akquisziels sind und
2. dass die Abwehrmaßnahmen des Managements mit steigender Größe des Akquisziels zunehmen.³

Resultierend nimmt die Anzahl der Akquiseakteur mit der Größe des Akquisziels ab.⁴ Das Anlagevermögen wird in einigen empirischen Untersuchungen durch zusätzliche Finanzkennzahlen erweitert. Tabelle 6 zeigt eine Übersicht empirischer Untersuchungen zur Größenhypothese und der dabei betrachteten Finanzkennzahlen. Zusätzlich wird die Wirkungsbeziehung zwischen jeder Finanzkennzahl und der Übernahmewahrscheinlichkeit dargestellt. Ein „-“ bedeutet, dass bei einer größer werdenden Finanzkennzahl die Übernahmewahrscheinlichkeit sinkt.

¹ Vgl. PALEPU (1986), S. 18; MEADOR et al. (1996), S. 12; POWELL (1997), S. 1014; SINGH (1975), S. 505 f.; ERDOGAN (2012), S. 75.

² Vgl. PALEPU (1986), S. 18; MEADOR et al. (1996), S. 12; POWELL (1997), S. 1014; SINGH (1975), S. 505 f.; ERDOGAN (2012), S. 72.

³ Vgl. PALEPU (1986), S. 18; MEADOR et al. (1996), S. 12; POWELL (1997), S. 1014; SINGH (1975), S. 505 f.; ERDOGAN (2012), S. 72.

⁴ Vgl. PALEPU (1986), S. 18; MEADOR et al. (1996), S. 12; POWELL (1997), S. 1014; SINGH (1975), S. 505 f.; ERDOGAN (2012), S. 72.

	Total Assets / Market	Market Capita- lization / Market	Number of Employees / Market	Sales Growth / Market
PALEPU (1986)	✓			
AMBROSE et al. (1992)	✓			
POWELL (1997)	✓			
BARNES (1998)		✓		
CUDD et al. (2000)	✓			
BRAR et al. (2009)		✓	✓	✓
ERDOGAN (2012)	✓			

Wahrscheinlichkeit der Übernahme ¹	-	-	-	-
--	---	---	---	---

Tabelle 6: Empirische Untersuchungen zur Größenhypothese mit den betrachteten Finanzkennzahlen²

2.2.4 Hypothese der unterbewerteten Vermögens

Die „asset undervalue hypothesis“ oder „market to book hypothesis“ wird in der vorliegenden Arbeit als Hypothese des unterbewerteten Vermögens bezeichnet. Diese Hypothese impliziert, dass Unternehmen, die einen im Vergleich zum Buchwert unterbewerteten Marktwert haben, günstige Akquizeziele sind.³ Hierbei wird die Annahme getroffen, dass Unternehmen bei Expansionsvorhaben zwei Alternativen, 1. alternative Neuinvestition und 2. alternative Übernahme eines bestehenden Unternehmens, haben und sich für die günsti-

¹ Die dargestellten Wirkungsbeziehungen zwischen Finanzkennzahl und Wahrscheinlichkeit der Übernahme gelten für alle betrachteten Untersuchungen.

² Vgl. PALEPU (1986), S. 18; AMBROSE et al. (1992), S. 579; POWELL (1997), S. 1014; BARNES (1998), S. 579; CUDD et al. (2000), S. 108; BRAR et al. (2009), S. 436; ERDOGAN (2012), S. 74.

³ Vgl. PALEPU (1986), S. 18; POWELL (1997), S. 1013. Alternativ ist die Betrachtung des Wiederbeschaffungswertes statt des Buchwertes. Bei dieser Betrachtung wird der Marktwert in Relation zum Wiederbeschaffungswert gesetzt, das Ergebnis ist das sogenannte TOBIN'S q. Vgl. HASBROUCK (1984), S. 351; WESTON et al. (1990), S. 199. Das Problem hierbei ist, dass der Wiederbeschaffungswert nicht kontinuierlich berichtet wird und somit die Datenverfügbarkeit für die Auswertung nicht gewährleistet ist. Vgl. POWELL (1997), S. 1013.

gere Alternative entscheiden.¹ Die Hypothese des unterbewerteten Vermögens wird anhand verschiedener Finanzkennzahlen gemessen. Tabelle 7 zeigt eine Übersicht empirischer Untersuchungen zur Hypothese des unterbewerteten Vermögens und der dabei betrachteten Finanzkennzahlen. Zusätzlich wird die Wirkungsbeziehung zwischen jeder Finanzkennzahl und der Übernahmewahrscheinlichkeit dargestellt. Ein „-“ bedeutet, dass bei einer größer werdenden Finanzkennzahl die Übernahmewahrscheinlichkeit sinkt. Ein „+/-“ bedeutet, dass keine Definierung der Wirkungsbeziehung erfolgt ist.

	Market Capitalization / Equity	Dividend Yield	Earnings Yield Past	Earnings Yield Future
PALEPU (1986)	✓			
BRAR et al. (2009)	✓	✓	✓	✓
CUDD et al. (2000)	✓			
HASBROUCK (1984)	✓			
POWELL (1997)	✓			
MEADOR et al. (1996)	✓			
AMBROSE et al. (1992)	✓			

Wahrscheinlichkeit der Übernahme ²	-	+/-	+/-	+/-
---	---	-----	-----	-----

Tabelle 7: Empirische Untersuchungen zur Hypothese des unterbewerteten Vermögens mit den betrachteten Finanzkennzahlen³

¹ Vgl. PALEPU (1986), S. 18; POWELL (1997), S. 1013; WESTON et al. (1990), S. 198 f.; ERDOGAN (2012), S. 73.

² Die dargestellten Wirkungsbeziehungen zwischen Finanzkennzahl und Wahrscheinlichkeit der Übernahme gelten für alle betrachteten Untersuchungen.

³ Vgl. HASBROUCK (1984), S. 351; PALEPU (1986), S. 18; AMBROSE et al. (1992), S. 579; MEADOR et al. (1996), S. 15; POWELL (1997), S. 1014; CUDD et al. (2000), S. 108; BRAR et al. (2009), S. 436.

2.2.5 Kurs-Gewinn-Hypothese

Die „price-earnings hypothesis“ wird in der vorliegenden Arbeit als Kurs-Gewinn-Hypothese bezeichnet. Bei dieser Hypothese werden der Aktienkurs und der Gewinn pro Aktie in Relation gesetzt.¹ Gemutmaß wird eine steigende Übernahmewahrscheinlichkeit mit fallendem Kurs-Gewinn-Verhältnis.² Hierbei wird angenommen,

1. dass die Erwartungen der Investoren³ bezüglich des Gewinnwachstums bei Unternehmen mit einem geringen Kurs-Gewinn-Verhältnis zu pessimistisch und bei einem hohen Kurs-Gewinn-Verhältnis zu optimistisch sind⁴ sowie
2. dass das geringere Kurs-Gewinn-Verhältnis des Akquiseziels durch die Übernahme aufgewertet wird.

Hierdurch wird ein impliziter Marktwertzuwachs bei dem Akquiseziel allein durch die Übernahme impliziert.⁵

Explizit bedeutet dies eine negative Wirkungsbeziehung zwischen Kurs-Gewinn-Verhältnis und Übernahmewahrscheinlichkeit.

2.2.6 Weitere Mergers & Acquisitions-Hypothesen zur Identifizierung von Mergers & Acquisitions-Kandidaten

Im folgenden Abschnitt werden weitere M&A-Hypothesen zur finanzkennzahlenbasierten Identifizierung von zukünftigen M&A-Kandidaten übersichtlich aufgeführt. Diese M&A-Hypothesen finden in der vorliegenden Arbeit keine Berücksichtigung, da diese im Vergleich zu den o. g. M&A-Hypothesen eine geringere wissenschaftliche Bedeutung bei der finanzkennzahlenbasierten Identifizierung von M&As haben.

¹ Vgl. CUDD et al. (2000), S. 109; MEADOR et al. (1996), S. 15; PALEPU (1986), S. 18.

² Vgl. CUDD et al. (2000), S. 109; MEADOR et al. (1996), S. 15; PALEPU (1986), S. 18.

³ Investoren sind in diesem Zusammenhang Aktionäre eines Unternehmens.

⁴ Vgl. BASU (1977), S. 663; DREMAN et al. (1995), S. 21.

⁵ Vgl. CUDD et al. (2000), S. 109; MEADOR et al. (1996), S. 15; PALEPU (1986), S. 18.

POWELL (1997) betrachtet die „Free Cash Flow hypothesis“, wonach Unternehmen, die schlechte Leistungen erbringen, sowie Unternehmen, die gute Leistungen erbringen und hohe Free-Cash-Flow-Reserven bilden, Akquiseziele werden.¹

MITCHELL et al. (1990) belegen in ihrer empirischen Untersuchung, dass Unternehmen, die durch eine Übernahme das Eigenkapital gesenkt haben, zu Akquisezielen werden. Umgekehrt werden Unternehmen, die Eigenkapital durch Übernahmen erhöht haben, nicht zu Akquisezielen.²

BARNES (1998) betrachtet zusätzlich zu der Hypothese des ineffizienten Managements, der Hypothese des Wachstums-Ressourcen-Ungleichgewichts und der Größenhypothese noch die „Anticipatory share price change hypothesis“, welche die Entwicklung des Aktienkurses kurz vor der Übernahme betrachtet.³

Zur Hypothese des ineffizienten Managements und zur Hypothese der unterbewerteten Aktivposten, die beide zu den Effizienz-Hypothesen gehören, kommen folgende Hypothesen hinzu:⁴

1. Differential managerial efficiency
2. Operating synergy
3. Pure diversification
4. Strategic realignment to changing environments,

die jedoch bei der finanzkennzahlbasierten Identifizierung von M&As kaum eine Rolle spielen.

¹ Vgl. POWELL (1997), S. 1013; JANSEN (2008), S. 261.

² Vgl. MITCHELL et al. (1990), S. 383 f.; JANSEN (2008), S. 261.

³ Vgl. BARNES (1998), S. 579; JANSEN (2008), S. 261.

⁴ Vgl. WESTON et al. (1990), S. 191 ff.; BREALEY et al. (2008), S. 885 ff.

Fusions- und Übernahmekandidaten in der deutschen
Stahlindustrie

Ein Vergleich zwischen binär logistischen Regressionen
und künstlichen neuronalen Netzen

Önder, F.

2016, XXX, 237 S. 32 Abb., 19 Abb. in Farbe., Hardcover

ISBN: 978-3-658-15373-1