

Band 6

ANWENDUNG NEURONALER NETZE IN DER FINANZWIRTSCHAFT

von

Bülent Acig

Kaiserslautern 2001

ISSN 1435-8484

Inhaltsverzeichnis

Inhaltsverzeichnis	II
Abbildungsverzeichnis	IV
Abkürzungsverzeichnis	V
Einleitung	1
A. Grundlagen künstlicher neuronaler Netze	4
I. Aufbau und Funktionsweise	4
1. Das biologische Vorbild.....	4
2. Funktion und Aufbau eines künstlichen Neurons	6
3. Funktion und Aufbau künstlicher neuronaler Netze	8
4. Das Lernen neuronaler Netze	10
II. Ausgewählte neuronale Netzwerke.....	12
1. Differenzierungsmerkmale neuronaler Netze	12
2. Das Perzeptron-Netz	14
3. Das Backpropagation-Netz	16
4. Das Counterpropagation-Netz.....	18
III. Neuronale Netze in der Praxis	20
1. Simulation neuronaler Netze durch Hard- und Software	20
2. Einsatzfelder neuronaler Netze	24
3. Neuronale Netze in der Betriebswirtschaft	26
B. Finanzwirtschaftliche Prognoseanwendungen	29
I. Grundlagen neuronaler Finanzprognosen	29
1. Zur prinzipiellen Prognostizierbarkeit von Finanzmarktdaten.....	29
2. Klassische Ansätze zur Finanzprognose	31
3. Vorgehensweise bei der Modellierung neuronaler Netze zur Finanzprognose.....	34
II. Wechselkursprognosen	39
1. Problemstellung.....	39

2. Modell von Poddig/Wallem.	42
3. Sonstige Studien zur Wechselkursprognose	45
III. Aktienkurs- und Indexprognose.....	47
1. Problemstellung.....	47
2. Modell von Gehring et al.	48
3. Sonstige Studien.....	52
IV. Abschließende Betrachtung zur Finanzprognose.....	54
1. Sonstige Einsatzmöglichkeiten für Finanzprognosen	54
2. Kritische Würdigung	57
C. Klassifikationsanwendungen neuronaler Netze im Finanzbereich.....	60
I. Grundlagen der neuronalen Klassifikation	60
1. Problemstellung und Ziele der neuronalen Klassifikation	60
2. Konventionelle Ansätze	62
II. Kreditwürdigkeitsprüfung von Konsumentenkrediten	64
1. Problemstellung.....	64
2. Modell von Enache.....	66
3. Weitere Studien zur Kreditwürdigkeitsprüfung von Konsumentenkrediten.....	70
III. Sonstige Klassifikationsanwendungen neuronaler Netze	73
1. Kundenklassifizierung.....	73
2. Bonitätsanalyse von Unternehmen.....	75
3. Klassifizierung von Bonds	76
IV. Kritische Würdigung.....	79
Zusammenfassung und Ausblick.....	81
Literaturverzeichnis.....	84

Abbildungsverzeichnis

Abbildung 1:	Interdisziplinarität neuronaler Netze	1
Abbildung 2:	Aufbau natürlicher Neuronen	4
Abbildung 3:	Aufbau einer Synapse	5
Abbildung 4:	Modell eines Units	6
Abbildung 5:	Vereinfachte Darstellung der Informationsverarbeitung in einem künstlichen Neuron	7
Abbildung 6:	Häufig verwendete Aktivierungsfunktionen	8
Abbildung 7:	Ein neuronales Netz als integriertes und offenes System	9
Abbildung 8:	Übersicht verschiedener Typen neuronaler Netze	13
Abbildung 9:	Leistungsfähigkeit von neuronalen Netzen in Abhängigkeit der Schichtenzahl	15
Abbildung 10:	Fehlerfläche eines neuronalen Netzes als Funktion der Gewichte w_1 und w_2	17
Abbildung 11:	3-schichtiges Backpropagation-Netz	18
Abbildung 12:	Vorwärtsgerichtetes CP-Netzwerk	19
Abbildung 13:	Ausgewählte neuronale Softwaresimulatoren	22
Abbildung 14:	Eine Klassifikation von Neurosystemen	23
Abbildung 15:	Übersicht klassischer Ansätze zur Finanzprognose	31
Abbildung 16:	Systematik fundamentaler Einflussfaktoren	33
Abbildung 17:	Kostenverlauf in den Lebensphasen eines neuronalen Netzes	35
Abbildung 18:	Entwicklungsphasen neuronaler Netze	36
Abbildung 19:	Rubriken der Einflussvariablen	43
Abbildung 20:	Renditeberechnung unterschiedlicher Handelsstrategien	44
Abbildung 21:	Vergleich der Ergebnisse bei Refenes und Zaidi	46
Abbildung 22:	Zweistufiges neuronales Netz zur Aktienkursprognose	50
Abbildung 23:	Ergebnisse der neuronalen Anlagesysteme	51
Abbildung 24:	Studien zur Zinsprognose mit neuronalen Netzen	55
Abbildung 25:	Prinzip der neuronalen Klassifikation	61
Abbildung 26:	Lineare versus nicht-lineare Trennung	64
Abbildung 27:	Potenzielle Einflussfaktoren auf die Kreditwürdigkeit	65
Abbildung 28:	Häufigkeiten der guten und schlechten Kreditfälle	67
Abbildung 29:	Ausgewählte Klassifikationsvariablen und deren Skalierung	68
Abbildung 30:	Trefferquoten des Ausgangsmodells	69
Abbildung 31:	Trefferquoten des verbesserten Modells	70
Abbildung 32:	Trefferquoten unterschiedlicher Verfahren	72
Abbildung 33:	Studien zur Bonitätsprüfung von Unternehmen	76
Abbildung 34:	Vergleich der Trefferquoten von linearer Regression und neuronalem Netz...	79

Abkürzungsverzeichnis

CP-Netz	Counterpropagation-Netz
KI	Künstliche Intelligenz
KNN	Künstliche Neuronale Netze
MLP	Multilayer-Perceptron

Einleitung

Es ist sicherlich noch übertrieben, im Zusammenhang mit künstlichen neuronalen Netzen von der Intelligenz, mal ganz abgesehen von der menschlichen Intelligenz zu sprechen. Die natürliche Intelligenz wird aber noch eine sehr lange Zeit der Forschungstrieb der Informationstechnologie bleiben.

So können Menschen und Tiere innerhalb von Millisekunden äußerst komplizierte Signale, wie Bilder, Geräusche und Gerüche aufnehmen und interpretieren. Ein Mensch ist in der Lage, ohne Mühe einen anderen, selbst nach langer Zeit und trotz Veränderungen, wiederzuerkennen. Entgegen der hohen Verarbeitungsgeschwindigkeiten von Computern, die 1000 mal höher sind als die im menschlichen Gehirn, könnten sie solche Aufgaben nicht bewerkstelligen. Alleine bei einfachen arithmetischen Operationen sind Computer dem menschlichen Gehirn überlegen.¹

Künstliche Neuronale Netze (KNN) gehören dem Bereich der Künstlichen Intelligenz (KI) an.

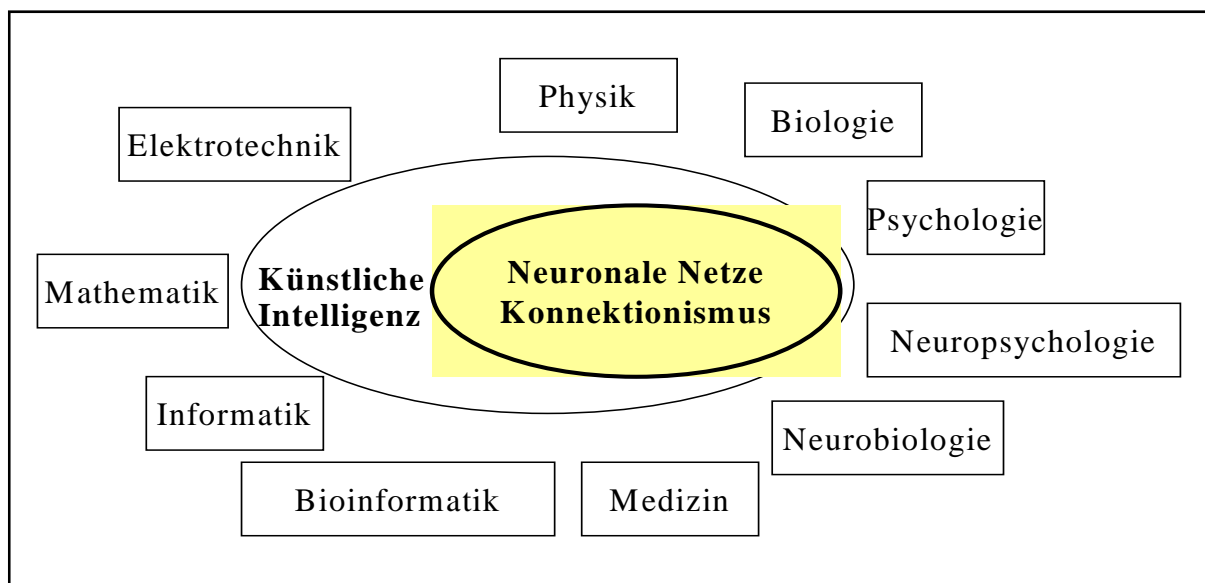


Abbildung 1: Interdisziplinarität neuronaler Netze²

Forschungsziel im Bereich KI ist die Nachahmung menschlicher Denkvorgänge. In diesem Bereich beschränkte sich die Wissenschaft lange Zeit darauf, Wissen explizit in Form von Regeln zu speichern und zu verarbeiten. Zu dieser sog. klassischen bzw. symbolischen KI zählt u.a. das Gebiet Expertensysteme.³ Künstliche neuronale Netze gehören dem Konnektionismus an. Dieser Terminus steht für einen Forschungszweig, der sich aus mehreren wissen-

¹ Vgl. Alex (Wiesbaden 1998), S. 77.

² Vgl. Zell (Bonn 1994), S. 24.

³ Vgl. Schmitter (Holzkirchen 1991), S. 1.

schaftlichen Disziplinen zusammensetzt (s. Abbildung 1). Ziel des konnektionistischen Ansatzes ist es, die Funktionsweise des Gehirns als ein Netzwerk von Neuronen zu erforschen und mit Hilfe von Computern zu imitieren. Im Gegensatz zur symbolischen KI wird hier das Wissen nicht explizit in Form von Regeln abgebildet, sondern implizit in Form von mathematischen Formeln und Gewichtungen gespeichert. Folglich wird dieser Forschungsbereich auch der subsymbolischen KI zugeordnet.⁴

Durch diesen Forschungsbereich erhofft man, in Zukunft künstliche neuronale Netzwerke zu entwickeln, die wie oben erwähnt, die höchst wünschenswerten Eigenschaften des menschlichen Gehirns beinhalten. Dazu gehören Beständigkeit und Fehlertoleranz. So sterben jeden Tag funktionsbeinhaltende Gehirnzellen ab oder erkranken, ohne dass die Tüchtigkeit des Gehirns erkennbar abnimmt. Die Lernfähigkeit und Flexibilität des Gehirns zählt seit jeher zu den Attributen, die man sich von einem Informationssystem erwünscht. Neuronale Netze sind, wie in den nachfolgenden Kapiteln zu sehen sein wird, in der Lage, sich durch Training auf veränderte Umweltbedingungen anzupassen, also zu lernen.

Die Informationsbearbeitung erfolgt in einem neuronalen Netz nicht sequentiell, sondern parallel. Dadurch kann, trotz einem, im Verhältnis zu Computern, langsamen Informationsfluss, eine Erhöhung der Verarbeitungsgeschwindigkeit erreicht werden. Weiterhin können neuronale Netze unscharfe, verrauschte und fehlerhafte Informationen erkennen und verarbeiten.⁵ Diese Ziele mit Rechnern zu erreichen, klingt noch etwas utopisch, ist aber schon teilweise realisiert worden.

Seit längerem existieren Anstrengungen, diese Eigenschaften speziell für die Finanzwirtschaft nützlich zu machen. Konventionelle Analyse- und Prognosesysteme scheinen hier veraltet und erneuerungsbedürftig zu sein. Als eine Alternative bieten sich hier künstliche neuronale Netze an, welche im Folgenden verkürzt als neuronale Netze bezeichnet werden sollen. Dieses moderne Tool konnte bereits zur Lösung von Aufgabenstellungen mit nicht-linearen und hochkomplexen Interdependenzen, welche für Finanzmarktdaten häufig vorliegen, bereits erfolgreich getestet werden.

In der nachfolgenden Arbeit soll auf die Thematik „Neuronale Netze in der Finanzwirtschaft“ näher eingegangen werden. Ziel ist es dabei, aufzuzeigen, wie und wofür neuronale Netze in

⁴ Vgl. Dorffner (Stuttgart 1991), S. 11-14.

⁵ Vgl. Hertz, u.a. (Redwood City 1991), S. 1.

der Finanzwirtschaft eingesetzt werden (konnten) und in wieweit Verbesserungen dadurch erreicht werden (konnten).

Zu Beginn der Arbeit wird das Basiswissen hierfür übermittelt. Hierzu wird unter der Überschrift „Grundlagen künstlicher neuronaler Netze“ eine allgemeine Einführung in Aufbau und Funktionsweise wiedergegeben. Weiterhin werden einige, in der Finanzwirtschaft häufig verwendete Netzwerktypen beschrieben.

Den Kern dieser Arbeit stellen die Kapitel 2 und 3 dar. Hier werden die zwei wichtigsten Einsatzarten neuronaler Netze, Prognose und Klassifikation, separat behandelt. Für jede Einsatzart werden u.a. Zielsetzung, konventionelle Instrumente, Anwendungsbereiche, empirische Studien und eine kritische Würdigung aufgezeigt.

Abschließend wird die Arbeit zusammengefasst und ein Ausblick wiedergegeben.

A. Grundlagen künstlicher neuronaler Netze

I. Aufbau und Funktionsweise

1. Das biologische Vorbild

Schon der Terminus „künstliche neuronale Netze“ sagt aus, dass deren Grundstruktur an das natürliche Nervensystem angelehnt ist. Für das Verständnis künstlicher neuronaler Netze empfiehlt es sich daher, Struktur und Funktion des biologischen Vorbildes, nämlich das neuronale Netz lebender Organismen, zu betrachten. Daher soll an dieser Stelle eine vereinfachte Darstellung⁶ der Struktur und Funktionsweise des Nervensystems wiedergegeben werden.

Die Grundbausteine eines Nervensystems sind die Nervenzellen, auch genannt Neuronen (s. Abbildung 2). Ein Neuron setzt sich im wesentlichen aus 3 Bestandteilen zusammen: dem Zellkörper, mehreren Dendriten und einem Axon. Im Zellkörper liegt der Zellkern (Soma), welches das genetische Material der Zelle enthält. Darüber hinaus sorgt der Zellkörper für den Stoffwechsel und die Funktionsfähigkeit der Gesamtzelle.⁷

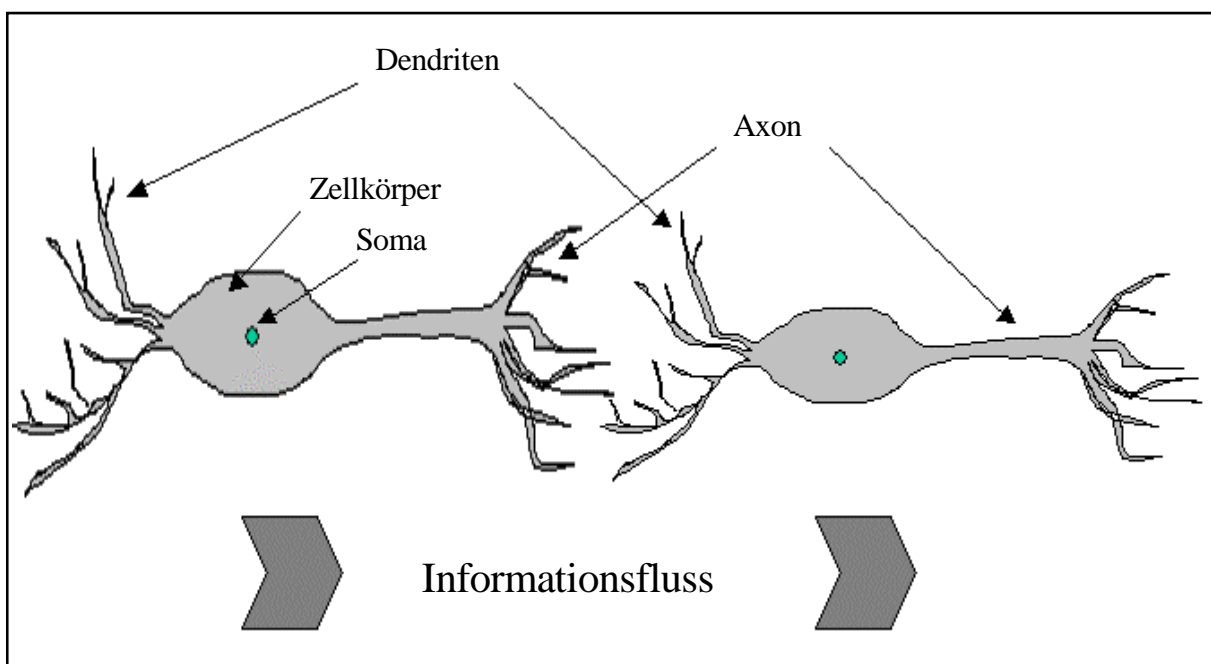


Abbildung 2: Aufbau natürlicher Neuronen

⁶ Mit Sicherheit kann man hier von einer Untertreibung reden.

⁷ Vgl. Wittkemper (Köln 1994), S. 9-10.

Die Dendriten sind feinverzweigte Ausläufer des Zellkörpers und dienen der Signalaufnahme, während das Axon, eine Leitungsbahn, die sich wiederum fein verzweigt, für die Signalweitergabe zuständig ist. Die Stelle, an der das Axon mit den Dendriten kontaktiert, wird als Synapse bezeichnet (s. Abbildung 3).

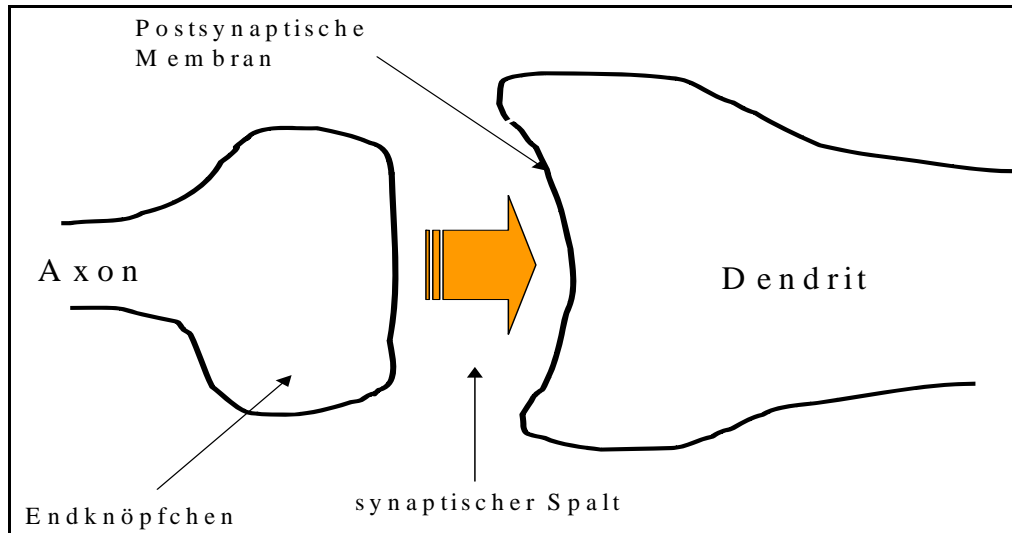


Abbildung 3: Aufbau einer Synapse

Diese Übergangsstellen tragen eine wesensbestimmende Funktion für das neuronale Netz. Am Endknöpfchen werden die Signale, welche in Form elektrischer Ladung ankommen, in eine chemische Substanz umgewandelt.⁸ Diese chemische Substanz, auch genannt Transmitter, wird durch ein elektrisches Potenzial, das an der Synapse gemessen werden kann, an die postsynaptische Membran transportiert. Somit werden die ankommenden Signale summiert und in Form von elektrischer Ladung an den Zellkern weitergeleitet. Ist ein bestimmter Schwellenwert im Zellkern erreicht, wird das empfangende Neuron selbst aktiv und gibt wiederum Signale über das Axon weiter.

Das elektrische Potenzial der Synapse bewirkt also eine Gewichtung der Eingangsinformation. Synapsen sind somit in der Lage, Informationen zu bewerten.⁹ Sie können erregend (excitatorisch) oder hemmend (inhibitorisch) wirken.¹⁰ Damit sind Synapsen wichtige Informationsträger, die ihre Bewertungsfunktionen zeitlich verändern können und infolgedessen das „Lernen“ repräsentieren.¹¹

⁸ Vgl. Fischbach (Heidelberg 1992), S. 34–37.

⁹ Füser (Wiesbaden 1995), S. 13.

¹⁰ Alex (Wiesbaden 1998), S. 80.

¹¹ Kinnebrok (Wien 1992), S. 14.

Das menschliche Gehirn besitzt ca. 10 Milliarden Neuronen. Ein Neuron ist wiederum mit 1000 bis 10000 anderen Neuronen verbunden. Wittkemper begründet die Performance neuronaler Netze folgendermaßen:

„Die Verarbeitungsgeschwindigkeit eines einzelnen Neurons ist verglichen mit heutigen Mikroprozessoren sehr gering, da die Neuronen jedoch gleichzeitig arbeiten, entsteht so ein hochgradig parallel arbeitendes und extrem effizientes Signalverarbeitungssystem, welches jeder heute bekannten Computerarchitektur unvergleichbar überlegen ist.“¹²

2. Funktion und Aufbau eines künstlichen Neurons

Analog zum biologischen Vorbild bestehen künstliche neuronale Netze aus mehreren Neuronen, die auch als Units bezeichnet werden (s. Abbildung 4).

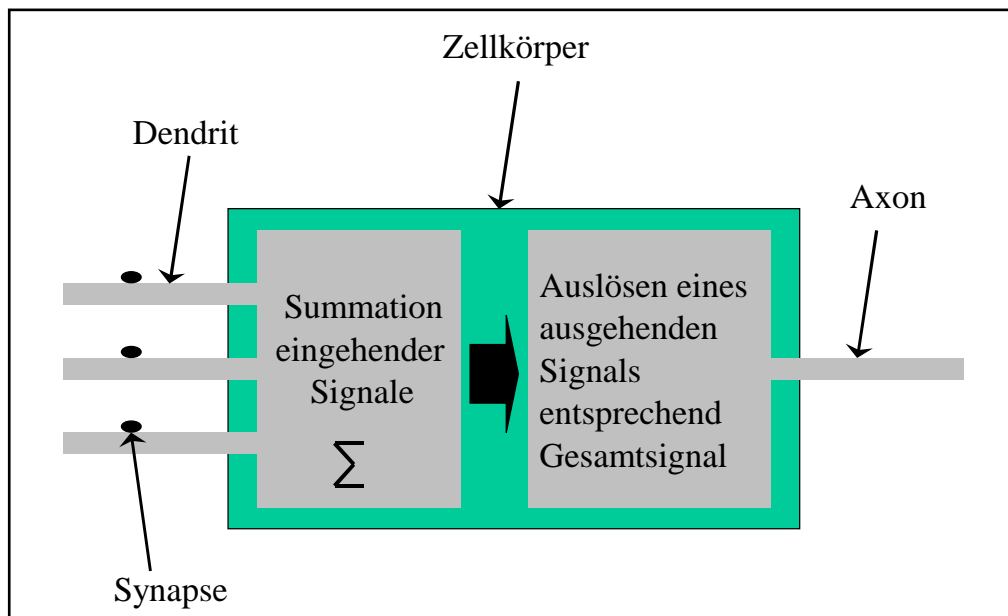


Abbildung 4: Modell eines Units¹³

Ein Signal läuft immer in eine Richtung. Jedem der n Eingänge eines Neurons wird ein Gewicht w zugewiesen, welches den Signalgewichtungen an den Synapsen der Nervenzellen entspricht. Die Gewichte w können sowohl positive als auch negative Werte annehmen. Dies hängt davon ab, ob das eintreffende Signal eine hemmende oder eine stimulierende Wirkung haben soll. Die Signalverarbeitung innerhalb des Neurons läuft wie folgt ab:¹⁴

¹² Wittkemper (Köln 1994), S. 11.

¹³ Vgl. Düsing (Hamburg 1997), S. 42.

¹⁴ Vgl. Hinton (Heidelberg 1992), S. 136. Um die Signalverarbeitung verständlicher darzustellen, wird in Anlehnung an Füsler (Wiesbaden 1995), S. 26 eine vereinfachte Darstellung mit fiktiven Gewichten herangezogen.

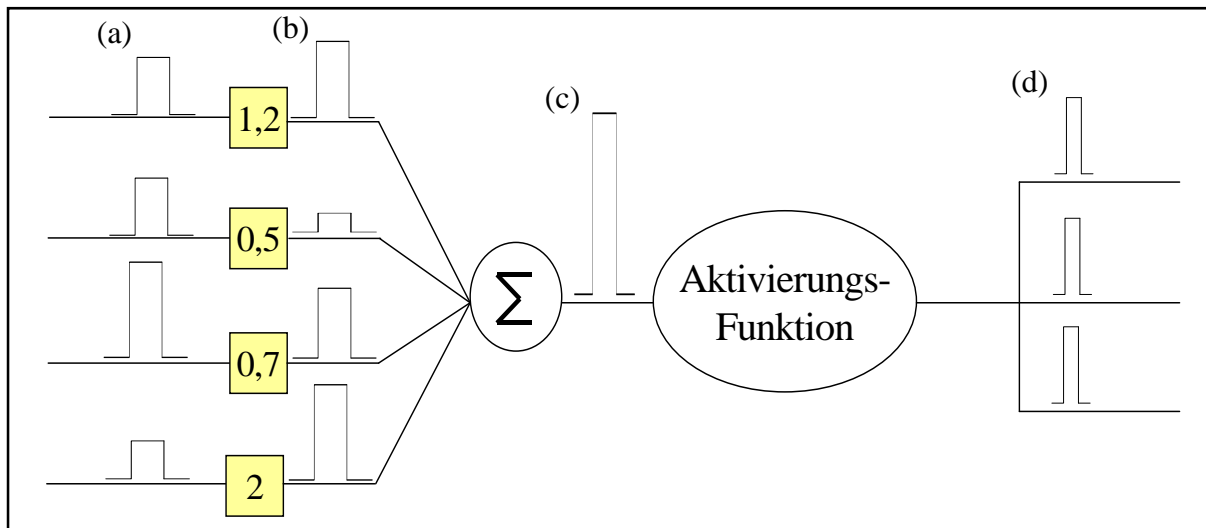


Abbildung 5: Vereinfachte Darstellung der Informationsverarbeitung in einem künstlichen Neuron

Die eintreffenden Signale (a) werden mit den jeweiligen Eingangsgewichten multipliziert. Dadurch entstehen die gewichteten Eingaben (b), welche dann überlagert werden. Gewöhnlich summiert man hierfür die gewichteten Eingaben. Als Ergebnis erhält man die gewichtete Gesamteingabe, was auch als Aktivierungspegel (c) bezeichnet wird. In allgemeiner Form lässt sich dieser folgendermaßen darstellen:¹⁵

$$a_i = w_{ij}o_j$$

mit a_i = Aktivierungspegel der Unit i

n = Anzahl der Eingänge

o_j = Ausgabesignal des vorgelagerten Unit j

w_{ij} = Verbindungsgewicht zw. den Units i und j

Jedes Neuron besitzt darüber hinaus eine Aktivierungsfunktion, auch genannt Transferfunktion. Über diese Transferfunktion, die nachfolgend näher erläutert werden soll, wird das Ausgangssignal (out) aus dem Aktivierungspegel bestimmt (s. auch Abbildung 6). Das Ausgangssignal wird über eine Leitung,¹⁶ welche sich wiederum verzweigen kann, gesendet. Daher sind alle momentanen Ausgangssignale eines Neurons gleich (d).

Die Art der Transferfunktion spielt eine wesentliche Rolle bei der erwünschten Funktion eines Neurons. Es existieren mehrere Transferfunktionen, wobei die nachfolgend dargestellten die am häufigsten verwendeten sind.¹⁷

¹⁵ Vgl. Wittkemper (Köln 1994), S. 12.

¹⁶ Entspricht dem Axon vom biologischen Vorbild.

¹⁷ Vgl. Zell (Bonn 1994), S. 77 und Wittkemper (Köln 1994), S. 13-14.

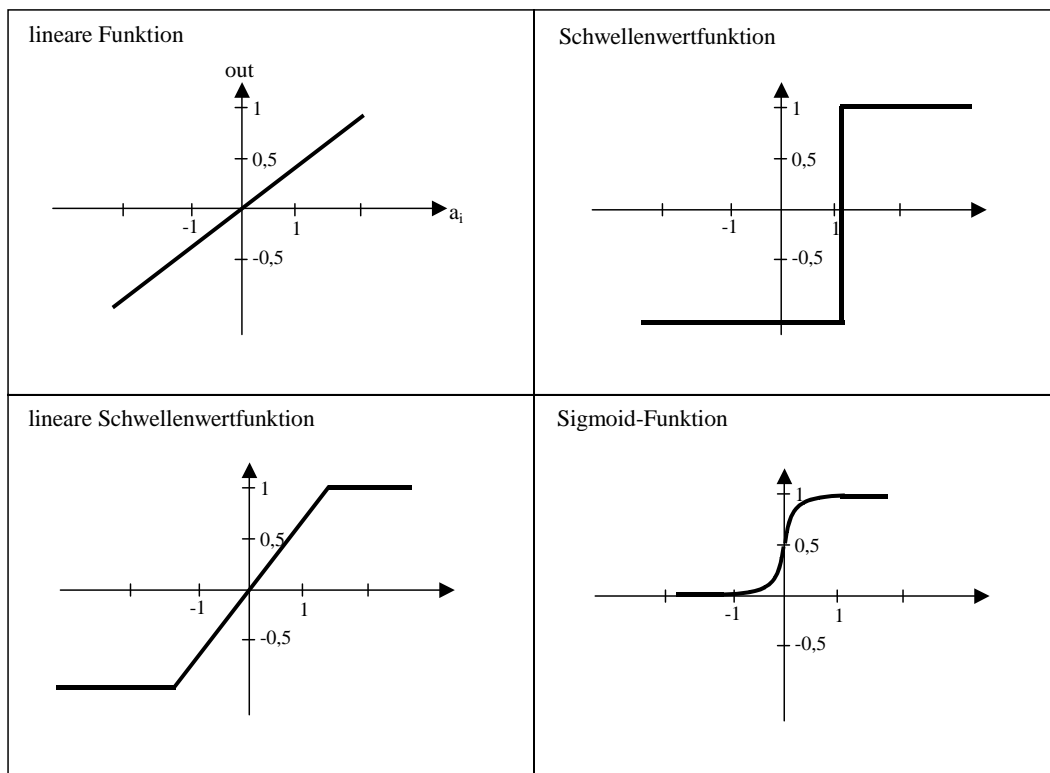


Abbildung 6: Häufig verwendete Aktivierungsfunktionen

Von diesen Transferfunktionen wird die Sigmoid-Funktion besonders bevorzugt. Der Grund dafür ist, dass sie über alle Werte differenzierbar (ableitbar) ist. Ferner weist diese Funktion am Nullpunkt höchste Sensibilität auf und ist unempfindlich gegenüber extremeren Werten (z.B. Ausreißer).¹⁸

Die Funktion eines künstlichen Neurons kann zusammengefasst als eine Entscheidung auf der Mikroebene betrachtet werden. Durch die Verknüpfung mehrerer Neuronen entsteht ein leistungsstarkes Netzwerksystem.

3. Funktion und Aufbau künstlicher neuronaler Netze

Nachdem Aufbau und Funktionsweise eines einzelnen Neurons dargestellt worden ist, soll nun gezeigt werden, wie sich das künstliche neuronale Netz aus diesen zusammensetzt und arbeitet. Isoliert betrachtet erscheinen die einzelnen Arbeitseinheiten primitiv. Aber durch die Vernetzung entsteht ein extrem leistungsfähiges System, mit dem jede beliebige Funktion dargestellt werden kann.

¹⁸ Vgl. Füsler (Wiesbaden 1995), S. 27-29.

In der Systemtheorie werden neuronale Netze als offene und dynamische Systeme beschrieben. Da die Verbindungsgewichte und somit die Topologien (s.u.) nicht starr sind, sondern sich verändern, spricht man von einem dynamischen System.¹⁹ Offen deshalb, weil jedes neuronale Netz in seine Umwelt integriert ist. Das Netz bekommt Eingangsinformationen aus der Umwelt und gibt Ausgangsinformationen an diese weiter (s. Abbildung 7).

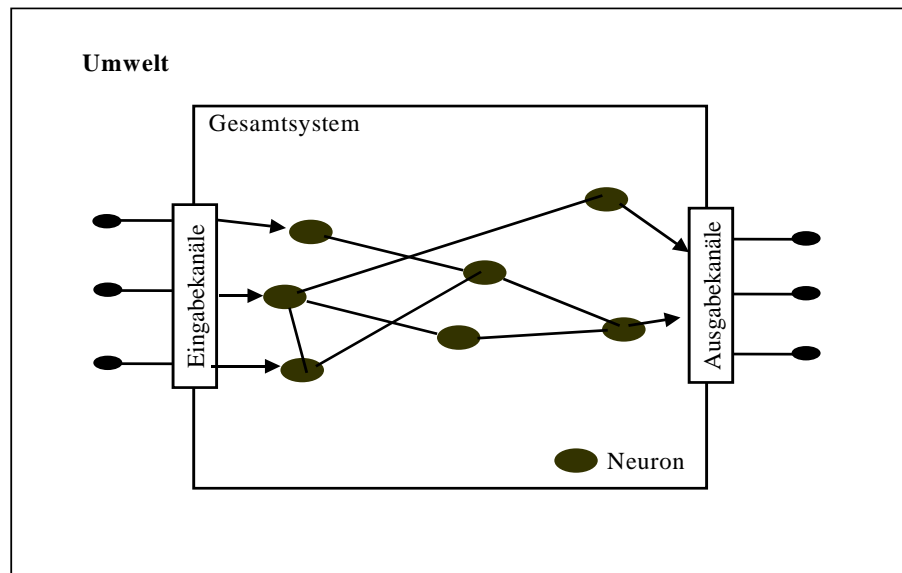


Abbildung 7: Ein neuronales Netz als integriertes und offenes System

Die Neuronen eines Netzes können bezüglich ihrer Aufgabenstellung unterschieden werden. Neuronen, deren Funktion die Signalaufnahme aus der Umwelt ist, werden als Input-Units bezeichnet, während Neuronen mit der Aufgabe, Information an die Außenwelt abzugeben, als Output-Units bezeichnet werden. Neuronen, welche keinen Kontakt zur Außenwelt haben, werden als Hidden Units bezeichnet.²⁰ Oftmals können Neuronen funktionell zu Schichten (layer) zusammengefasst werden. Entsprechend spricht man in diesem Fall von Input-, Output- und Hidden-Layer. In den Hidden-Layer können nicht-lineare, implizite Abhängigkeiten zwischen Eingangsinformation und dem dazugehörigen Ausgangssignal festgelegt werden. Das Problemlösungswissen ist in einem neuronalen Netz nicht in Form von Regeln gespeichert (s. auch Kap. 2.2.1), sondern auf einer niedrigeren Abstraktionsebene, was eine Rückverfolgung der Informationsverarbeitung beinahe unmöglich macht. Aus diesem Grund spricht man im Zusammenhang mit künstlichen neuronalen Netzen auch von einer „informationsverarbeitenden Black-Box“.²¹

¹⁹ Vgl. Lackes (München 2000), S. 21-22.

²⁰ Vgl. Wittkemper (Köln 1994), S. 14.

²¹ Füsler (Wiesbaden 1995), S. 22-23.

Für die Informationsverarbeitung ist neben der Lernregel, welche im nächsten Kapitel erörtert werden soll, die Netztopologie, also die Art der inneren Vernetzung, entscheidend. Generell wird zwischen 2 Arten von Netztopologien unterschieden.²²

(1) Vorwärtsgerichtete, rückkopplungsfreie Netze (feedforward-Netze)

- Hierarchische feedforward-Netze: Diese Netze sind in Ebenen (entspricht Schichten bzw. Layer) aufgeteilt. Es gibt nur Verbindungen von der einen zur hierarchisch nächstliegenden Schicht.
- Allgemeine feedforward-Netze: Der einzige Unterschied zu den hierarchischen feedforward-Netzen ist das Vorhandensein von Verbindungen, welche Ebenen überspringen (shortcut connection).

(2) Netze mit Rückkopplung (feedback-Netze)²³

- Netze mit direkter Rückkopplung (direct feedback-Netze): Dadurch dass Neuronen direkt mit sich verbunden sind, können sie sich selbst aktivieren bzw. hemmen.
- Netze mit indirekter Rückkopplung (indirect feedback-Netze): In diesem Fall kann es Rückkopplungen zwischen Ebenen geben.
- Netze mit möglichen Rückkopplungen innerhalb einer Ebene (lateral feedback-Netze): Diese Netze werden für Aufgaben eingesetzt, bei denen nur der stärkste Neuron aktiv werden soll. Daher werden diese auch als winner-takes-all-Netze bezeichnet.

Vollständig verbundene Netze (Hopfield-Netze): In diesem Fall liegen Verbindungen zwischen allen Neuronen vor.

4. Das Lernen neuronaler Netze

Nach der Modellierung eines Netzes für einen spezifischen Einsatzbereich (s. Kap. 2.1.3) folgt die Vorbereitung für den praktischen Einsatz. Hier wird das Netz in Betrieb genommen. Der Betrieb neuronaler Netze kann in drei Phasen eingeteilt werden: Lern-, Test- und Verarbeitungsphase.

In der Verarbeitungsphase wird das Netz zur Lösung des spezifischen Problems (z.B. Beurteilung der Kreditwürdigkeit einer Person) eingesetzt. Zuvor muss aber das Problemlösungswissen vom fertiggestellten Netz erlernt und anschließend getestet werden. In der Lernphase

²² Vgl. Baumgartner (Bamberg 1998), S. 71-72 und Zell (Wiesbaden 1998), S. 10.

²³ In der Literatur werden diese auch bezeichnet als rekursive bzw. rekurrente Netze.

werden dem Netz die Musterdaten (historische Daten) eingegeben. Durch diese Musterdaten generiert das Netz selbstständig Informationen, erkennt die Zusammenhänge und speichert das Problemlösungsverständnis implizit, also in einer subsymbolischen Form. Hierzu benötigen neuronale Netze weder Verständnis über Struktur und Zusammenhänge noch statistische Modelle.²⁴ Wie bereits angedeutet, lernt ein neuronales Netz fast ausschließlich durch die Veränderung der Verbindungsgewichte. Theoretisch gibt es eine Vielzahl von Lernprinzipien. Zu diesen gehören:

- (1) Modifikation der Verbindungsstärken w_{ij} , (Extremfälle dieses Prinzips sind das Löschen vorhandener Verbindungen und das Entwickeln neuer Verbindungen.)
- (2) Modifikation der Transferfunktion,
- (3) Bildung neuer Neuronen,
- (4) Löschen von vorhandenen Neuronen.²⁵

Die Lernarten werden in der Literatur oftmals in drei Kategorien eingeteilt:²⁶

- (1) Überwachtes Lernen (supervised learning): Hier werden dem Netz neben den Eingabedaten die dazugehörigen Ausgabedaten (Sollwerte) durch einen externen Lehrer vermittelt. Das Netz hat die Aufgabe, durch Modifikation ihrer Verbindungsgewichte diese Sollwerte zu approximieren.
- (2) Bestärkendes Lernen (reinforcement learning): In diesem Fall gibt ein externer Lehrer an, ob ein Ausgabewert richtig oder falsch ist. Angewandt wird diese Methode für Aufgabenstellungen, bei denen der genaue bzw. beste Ausgabewert nicht bekannt ist oder nicht existiert.
- (3) Unüberwachtes Lernen (unsupervised learning): Beim unüberwachten Lernen gibt es weder einen externen Lehrer noch Informationen bzgl. der Soll-Ausgabewerte. Das Netz versucht, in den Eingabemustern selbstständig Strukturen zu erkennen und zu kategorisieren.

Der bekannteste Repräsentant der Klasse des überwachten Lernens ist der Backpropagation-Algorithmus. Dieser wird für mehrschichtige Feedforward-Netze eingesetzt und ist besonders relevant für finanzwirtschaftliche Einsatzgebiete. Eine mathematische Beschreibung ist an

²⁴ Vgl. Miller (Mannheim 1994), S. 22-23.

²⁵ Vgl. Zell (Wiesbaden 1998), S. 9-11.

²⁶ Vgl. Zell (Bonn 1994), S. 93.

dieser Stelle weniger angebracht. Statt dessen soll hier der Backpropagation-Algorithmus, welcher in Kapitel 1.2.3 nochmals kurz aufgegriffen wird, dargestellt werden. Das Lernen erfolgt in 5 Schritten:

- (1) Übermittlung der Eingabedaten über die Input-Units.
- (2) Bearbeitung der Eingangsdaten durch das Netz (Vorwärtspropagierung) führen zu einem Output.
- (3) Tatsächlicher Output wird mit Soll-Output verglichen. Die Differenz wird bestimmt.
- (4) Diese Differenz wird von der Ausgabe- zur Eingabeschicht rückwärts propagiert.²⁷ Dadurch erhält man die für die Angleichung von Soll- und Ist-Output nötigen Veränderungen der Verbindungsgewichte Δw_{ij} .
- (5) Als nächstes werden dann die Modifikationen der Verbindungsgewichte vorgenommen und das Ganze (iterativ) mit neuen Eingangsmustern durchgeführt.

II. Ausgewählte neuronale Netzwerke

1. Differenzierungsmerkmale neuronaler Netze

Es existieren derzeit 20 bis 30 unterschiedliche Netztypen, von denen wiederum viele Varianten existieren. Diese können anhand mehrerer Merkmale differenziert werden. Dazu gehören u.a. die in den vorhergehenden Abschnitten dargestellten Aspekte neuronaler Netze. Folgende Kriterien können zur Unterscheidung dienen:²⁸

- Lernart: Man kann im groben differenzieren nach überwachtem, unüberwachtem und bestärkendem Lernen (Kap. 1.1.4).
- Architektur: Hier kann separiert werden nach einschichtigen und mehrschichtigen Netzwerken.
- Informationsfluss: Unterschieden wird in diesem Zusammenhang zwischen vorwärtsgerichteten und rekurrenten Netzen (Feedforward- und Feedbacknetze).
- Berechnungsart: Man unterscheidet zwischen dem deterministischen, welcher am häufigsten eingesetzt wird und dem stochastischen Berechnungsmodus. Beim stochastischen Berechnungsmodus sind die Ausgangssignale Zufälligkeiten ausgesetzt.

²⁷ Daher auch der Name Backpropagation.

²⁸ Vgl. Rigoll (Renningen-Malmsheim 1994), S. 82-84 und Füsler (Wiesbaden 1995), S. 38-40.

- **Verarbeitungsmodus:** Die synchrone Verarbeitung ist dadurch charakterisiert, dass alle Neuronen einer Schicht oder des gesamten Netzes gleichzeitig arbeiten. Beim asynchronen Modus setzen die Neuronen sukzessiv ein.

Dies sind die, in der Literatur am häufigsten genannten Differenzierungsmerkmale neuronaler Netze, wobei es sicherlich noch weitere Kriterien gibt, nach denen klassifiziert werden kann. So kann man z.B. nach der Aktivierungsfunktion (Sigmoid-Funktion, Schwellenwert-Funktion etc.), nach dem Wertebereich der Eingangs- und Ausgangssignale (binäre versus kontinuierliche Werte) oder nach der Verbindungsrichtung (unidirektional versus bidirektional) unterscheiden.

In der nächsten Tabelle ist eine grobe Übersicht bekannter Typen künstlicher neuronaler Netze, geordnet nach dem Arbeitsmodus (feedforward versus feedback), aufgelistet. Hierbei ist mit Bezug auf den Kern dieser Arbeit zu erwähnen, dass zu Zwecken der Prognose und Klassifikation im finanzwirtschaftlichen Bereich überwiegend die rückkopplungsfreien Netzwerkmodelle zum Einsatz gelangen.

Typen neuronaler Netze	
Feedforard-Netze	Feedback-Netze
Perzeptron	Boltzmann-Maschine
Multi-Layer-Perzeptron	Hopfield-Netze
Adaline	Cognition
Madaline	Reccurent Perceptron
Probabilistic Neural Network	Neocognition
Counterpropagation Netze	

Abbildung 8: Übersicht verschiedener Typen neuronaler Netze²⁹

Von diesen beispielhaft aufgezeigten Feedforward-Netzen werden zu finanzwirtschaftlichen Zwecken hauptsächlich Backpropagation-Netze³⁰ und (seltener) Counterpropagation-Netze eingesetzt. Das Perzeptron ist das älteste Netzwerk, auf deren Grundlage die meisten neueren Modelle entwickelt wurden. Aus diesem Grunde sollen als nächstes diese drei Netzwerkmodelle kurz vorgestellt werden.

²⁹ Vgl. Wilbert (Frankfurt a. M. 1996), S. 82.

³⁰ Das Backpropagation-Netz ist ein Spezialfall des Multi-Layer-Perzeptron (s. Kap. 2.3.3).

2. Das Perzeptron-Netz

Frank Rosenblatt entwickelte im Jahre 1958 das Perzeptron und schuf somit die Basis für die Entwicklung weiterer Netzwerkmodelle.³¹ Der Begriff Perzeptron steht dabei nicht für einen spezifischen Netztyp. Es ist ein Sammelbegriff für Feedforward-Netze, die einstufig aufgebaut sind und mit einem überwachten Lernalgorithmus trainiert werden. Mit dem Begriff „einschichtig“ ist hierbei das Nicht-Vorhandensein von verdeckten (aktiven) Schichten gemeint. Es können also mehrere Schichten existieren, von denen aber nur die Gewichte zwischen zwei Schichten modifiziert werden können.

Das klassische Perzeptron-Modell orientiert sich an der Signalverarbeitung des Auges. Hier besitzt das Netz drei Schichten. Die Eingabeschicht wird analog zum biologischen Vorbild (Netzhaut) *Retina* genannt. Die Neuronen dieser Schicht sind mit den Neuronen der nachfolgenden Schicht fest verbunden (keine Modifikationen möglich).³² Die Eingabewerte werden hier direkt an die nachfolgenden Neuronen weitergegeben. Die Neuronen der nachfolgenden Schicht funktionieren wie Musterdetektoren. Sie bündeln die Signale (Reize), die von der *Retina* ankommen. Diese Schicht wird auch als Assoziationsschicht bezeichnet. Die Neuronen dieser Schicht sind vollständig mit den Neuronen der Ausgabeschicht verbunden. Die Ausgabeneuronen arbeiten ausschließlich mit der Schwellenwertfunktion und können folglich nur binäre Ausgabewerte (z.B. (0,1) oder (-1,1)) erzeugen. Über einen überwachten Lernalgorithmus (Perzeptron-Lernregel)³³ müssen nun die Verbindungsgewichte zwischen Assoziations- und Ausgabeschicht so verändert werden, dass bestimmte Eingangsmuster der *Retina* mit der entsprechend erwünschten Reaktion (binäre Ausgabe) assoziiert werden.

³¹ Vgl. Rosenblatt, F.: *The Perceptron*, 1958, S. 386-408, nachgedruckt in: Anderson, J. und Rosenfeld, E.: *Foundations of Research*, MIT Press, Cambridge 1988.

³² Vgl. Rojas (Heidelberg 1996), S. 52-53.

³³ Eine Beschreibung des Perzeptron-Lernregels findet sich in Rieß (Aachen 1994), S. 26-28.

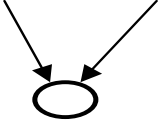
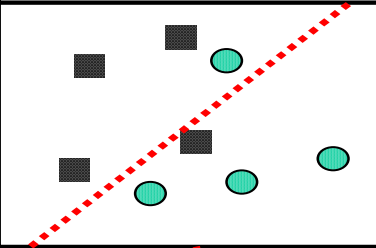
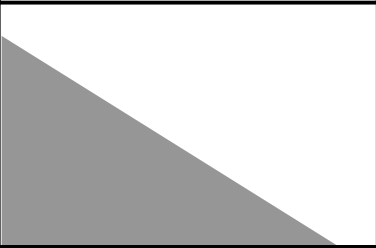
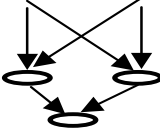
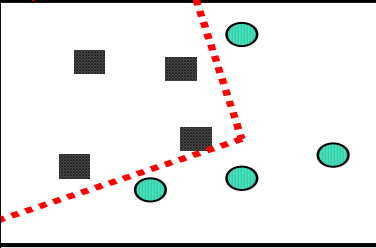
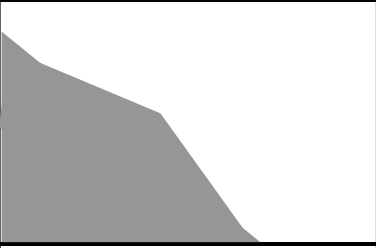
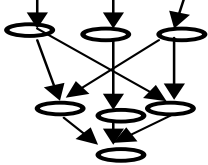
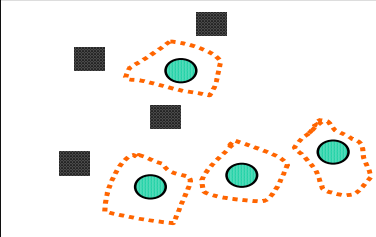
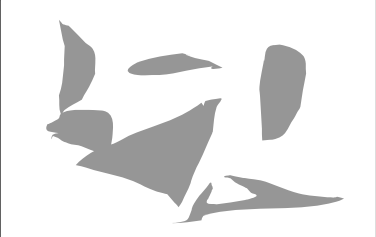
Struktur	Differenzierungsfähigkeit von Merkmalsausprägungen	Darstellbare Flächen
<p>1 Schicht</p> 		
<p>2 Schichten</p> 		
<p>3 Schichten</p> 		

Abbildung 9: Leistungsfähigkeit von neuronalen Netzen in Abhängigkeit der Schichtenzahl³⁴

Das Perzeptron wird dazu benutzt, bestimmte Muster zu erkennen und zu klassifizieren. Dabei kann ein derartiges Netz immer nur die Frage der Zugehörigkeit eines Musters zu einer vorher erlernten Klasse beantworten (ja oder nein). Die Einsatzgebiete klassischer Perzeptron-Netze sind stark beschränkt. Aufgrund der Tatsache, dass nur eine einstufige Informationsverarbeitung vorliegt, können nur Muster voneinander unterschieden werden, welche durch eine lineare Funktion voneinander getrennt werden können (s. Abbildung 9). Somit ist die Repräsentationsfähigkeit³⁵ solcher Netze extrem eingengt. Dieses Problem konnte durch die Kombination mehrerer Perzeptron-Netze und durch die folglich entstandenen zusätzlichen Verarbeitungsschichten behoben werden. Wie in Abbildung 9 zu sehen ist, kann dadurch die Klassifikationsfähigkeit praktisch unbegrenzt gesteigert werden.

Aus diesem Grund werden die klassischen Perzeptron-Netze heutzutage kaum eingesetzt. Sie bilden aber dennoch die Grundlage für die nachfolgenden, leistungsfähigeren Netzwerkmodelle und haben aus diesem Grund eine zumindest oberflächliche Darstellung verdient.

³⁴ (In Anlehnung an Lippman, R. P.: Nets, 1987) aus Schnurr (Wiesbaden 1997), S. 68. Eine detaillierte Darstellung dieser Problematik findet sich in Zell (Bonn 1994), S. 98-102.

³⁵ Damit ist die Fähigkeit des Netzes gemeint, eine Funktion realisieren zu können.

3. Das Backpropagation-Netz

Das Multilayer-Perceptron (MLP) ist eine Weiterentwicklung des Perzeptron-Netzes. Wie schon der Name sagt, bestehen diese Netzwerke im Gegensatz zu Perzeptron-Netzen aus mindestens einer weiteren, verdeckten Schicht (Hidden-Layer). *Minsky und Papert* kamen nach einer differenzierten Analyse der Fähigkeiten von Perzeptron-Netzen im Jahre 1988 zu dem Ergebnis, dass durch eine weitere verdeckte Schicht die Repräsentationsfähigkeit des Netzes erheblich zunimmt.³⁶ So ermöglichte schon die Zwischenschaltung (nur) einer verdeckten Schicht die Abbildung beliebiger Merkmalsräume (s. Abbildung 9). Dieser Sachverhalt wurde schon früher von einigen Wissenschaftlern erkannt. Der Grund aber, weshalb MLP-Netze nicht schon früher zum breiten Einsatz kamen, war, dass kein geeigneter Lernalgorithmus entwickelt werden konnte, welcher die zielgenaue Modifikation der Verbindungsgewichte in den verdeckten Schichten ermöglichte. Den Weg für den Einsatz von MLP-Netzen öffnete erst die präzise Beschreibung und Operationalisierung eines Lernverfahrens durch *Rumelhart, Hinton und Williams* im Jahre 1986.³⁷ Das Hauptproblem bei der Entwicklung eines Lernverfahrens für ein Netz mit einer oder mehreren verdeckten Schichten war das Auffinden der Fehlerquelle. Dadurch, dass die Neuronen in einer verdeckten Schicht arbeiten, konnte das fehlerverursachende Neuron nicht identifiziert werden. Gelöst wurde dieses Problem durch eine Lernregel, die den Fehler des Netzes als Differenz zwischen der tatsächlichen Ausgabe und der gewünschten Netzausgabe ermittelt. Als Fehlermaß wird hierbei die Summe der quadrierten Abweichungen zwischen Ist- und Soll-Ausgabewert der einzelnen Ausgabeneuronen hinzugezogen. Wenn man den Zielfunktionswert als eine Funktion der n verwendeten Verbindungsgewichte betrachtet, erhält man ein $n+1$ -dimensionales Fehlergebirge. Mit dem Gradientenabstiegsverfahren wird nun versucht, ein möglichst tiefes Tal zu erreichen (s. Abbildung 10).³⁸

³⁶ Vgl. Minsky/Papert (London 1988), S. 231.

³⁷ Vgl. Rumelhart/Hinton/Williams (Cambridge 1986), S. 319-328.

³⁸ Mit der Methode des steilsten Abstiegs wird versucht, möglichst schnell einen globalen Minimum der Fehlerfunktion zu erreichen.

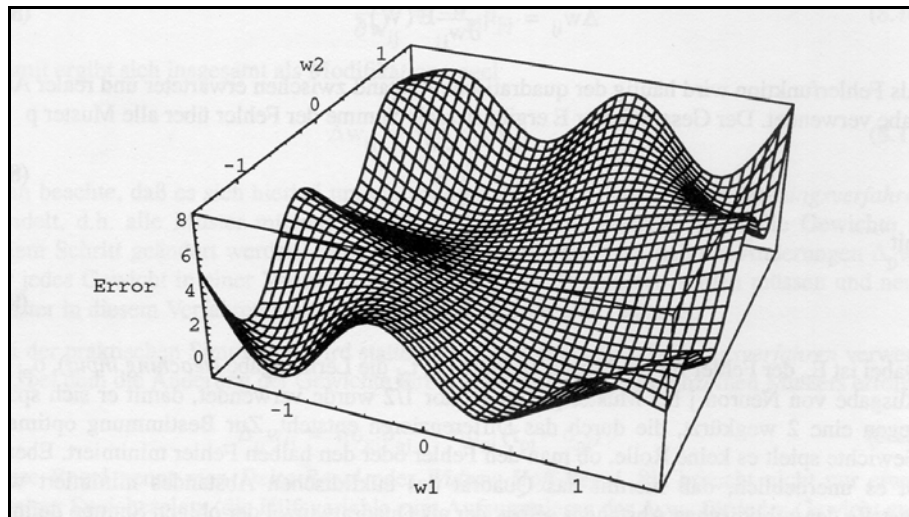


Abbildung 10: Fehlerfläche eines neuronalen Netzes als Funktion der Gewichte w_1 und w_2 ³⁹

Dazu werden die Fehler von jedem Ausgabeneuron relativ auf die vorgelagerten Neuronen heruntergebrochen und entsprechende Veränderungen der Verbindungsgewichte vorgenommen.⁴⁰ Da der Fehler bei dieser Methode rückwärts durch das Netz gegeben (propagiert) wird, bezeichnet man sie als Backpropagation-Algorithmus. Dieses Verfahren wird am häufigsten für MLP-Netze eingesetzt. Daher werden diese in der Literatur auch als Backpropagation-Netze bezeichnet. Somit sind Backpropagation-Netze streng genommen (sehr häufig vorkommende) Sonderfälle der MLP-Netze. Die Netztopologie des Backpropagation-Netzes ist der des Perzeptron-Netzes sehr ähnlich. Auch diese ist rückkopplungsfrei. Sie besitzt mindestens eine verdeckte Schicht, wobei im Grundmodell die Neuronen aufeinanderfolgender Schichten vollständig miteinander vernetzt sind (s. Abbildung 11).

³⁹ Vgl. Zell (Bonn 1994), S. 105.

⁴⁰ Eine detaillierte Beschreibung dieses Verfahrens findet sich bei Rojas (Heidelberg 1996), S. 149-172 und Zell (Wiesbaden 1998), S. 105-114.

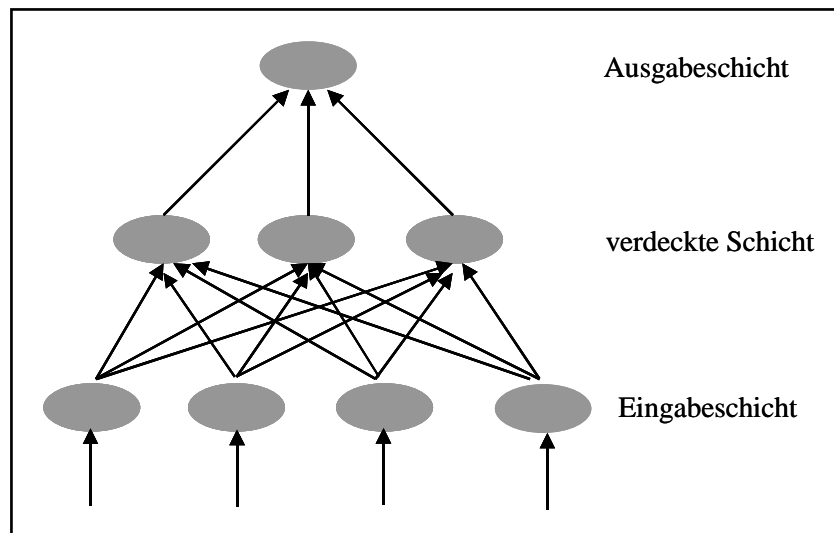


Abbildung 11: 3-schichtiges Backpropagation-Netz

Aufgrund der hohen Leistungsfähigkeit von Backpropagation-Netzen, was ihre Repräsentationsfähigkeit angeht, erfreut sich der Einsatz dieser Netztypen hoher Beliebtheit, insbesondere wenn es um den Einsatz in finanzwirtschaftlichen Bereichen geht. So ist es mit Backpropagation-Netzen möglich, jede beliebige auch nicht-lineare Funktion zu approximieren. Diese Eigenschaft wird, wie in den nachfolgenden Kapiteln zu sehen ist, speziell für Aufgaben der Prognose und der Klassifikation im Finanzbereich angewandt.

4. Das Counterpropagation-Netz

Das Counterpropagation-Netz (CP-Netz) ist ein verhältnismäßig moderner Netztyp. Es wurde im Jahre 1987 von *Hecht-Nielsen* entwickelt und wird im wesentlichen den Feedforward-Netzen zugeordnet. Unter den vielfältigen Modifikationen dieser Netztypen finden sich aber auch teilweise rückgekoppelte Varianten.⁴¹ Das wesentliche Potenzial von CP-Netzen ist die Fähigkeit, gespeicherte Muster über eingegebene Teilmuster abrufen zu können. Dies wird dadurch erreicht, dass die Informationen durch zwei miteinander gekoppelte und mit unterschiedlichen Lernalgorithmen arbeitende Netzarten geführt werden. In der sog. Kohonen-Schicht wird unüberwacht (selbstorganisierend) gelernt, während in der sog. Grossberg-Schicht eine überwachte Lernmethode zum Einsatz kommt. Die Kombination dieser zwei Netztypen bzw. Lernalgorithmen liefert Eigenschaften, welche keines der beiden voneinander isoliert besitzen könnte.⁴² So können bei CP-Netzen die Trainingszeiten im Verhältnis zu den im vorherigen Kapitel erläuterten Backpropagation-Netzen um den Faktor 100 reduziert wer-

⁴¹ Vgl. Wilbert (Frankfurt a. M. 1996), S. 82 und S. 93-94.

⁴² Vgl. Zell (Bonn 1994), S. 189.

den. Dafür sind CP-Netze aber nicht so flexibel einsetzbar wie Backpropagation-Netze.⁴³ Die Informationsverarbeitung in einem CP-Netz kann am besten schichtenweise beschrieben werden (s. Abbildung 12).

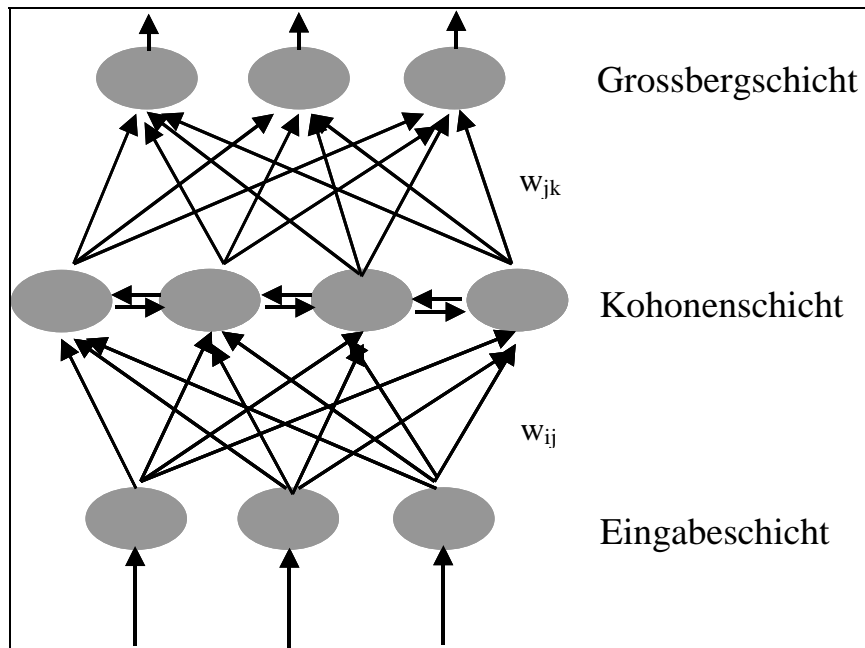


Abbildung 12: Vorwärtsgerichtetes CP-Netzwerk⁴⁴

Die Eingabeschicht dient meistens nur der Verteilung der Eingangsinformation an die Kohonenschicht. Als Aktivierungsfunktion wird die Identitätsfunktion (lineare Funktion) genommen. Generell ist jedes Eingabeneuron mit den Neuronen der Kohonenschicht über die Gewichte w_{ij} vollständig verbunden. Genauso sind alle Neuronen der Kohonenschicht mit den Neuronen der Grossbergschicht verbunden. Innerhalb der Kohonenschicht sind die Neuronen über hemmende Verbindungen miteinander verknüpft. Dadurch kann auf ein gegebenes Eingabevektor⁴⁵ immer nur ein Neuron, nämlich das Neuron mit der höchsten Nettoeingabe, aktiviert werden (winner-takes-all). Durch dieses Prinzip werden ähnliche Eingabevektoren auf dem gleichen Neuron der Kohonenschicht abgebildet. Folglich führt das unüberwachte Lernen der Kohonenschicht zu einer Modifikation der Verbindungsgewichte derart, dass für ähnliche Eingabemuster Cluster erzeugt werden. Während in der Kohonenschicht eine unbestimmte Klassenzuordnung stattfindet, hat die Grossbergschicht die Aufgabe, die gewünschte Zuordnung als Endergebnis auszugeben. Hierzu wird ein überwachtes Lernprinzip hinzugezogen. Dabei wird zuerst über mehrere Eingabemuster bestimmt, welche Neuronen der Koho-

⁴³ Vgl. Füsler (Wiesbaden 1995), S. 80.

⁴⁴ Vgl. Alex (Wiesbaden 1998), S. 80.

⁴⁵ Ein Eingabevektor ist ein codierter Eingabemuster.

nenschicht entsprechend aktiv werden. Danach werden genau die Verbindungsgewichte korrigiert, welche im Bezug zum sendenden Neuron stehen.⁴⁶ So erhält man nach erfolgreichem Training zu jedem aktiven Neuron der Kohonenschicht (dies entspricht der internen Klassenzuordnung eines Eingabemusters durch das Netz) die erwünschte Zuordnung in Form eines Ausgabevektors. Folglich ist die maximale Anzahl der differenzierbaren Klassen durch die Anzahl der Neuronen in der Kohonenschicht gegeben.

Aufgrund dieser dargestellten Eigenschaften werden CP-Netze zur Klassifikation, Mustererkennung und zur Mustervervollständigung eingesetzt. Speziell die Fähigkeit, verrauschte und fehlerhafte Daten zuzuordnen und vervollständigen zu können, machen es zu einem zukunfts-trächtigen, finanzwirtschaftlichen Instrumentarium.

III. Neuronale Netze in der Praxis

1. Simulation neuronaler Netze durch Hard- und Software

Das Potenzial neuronaler Netze wurde erst Mitte der achtziger Jahre von der Wissenschaft voll erkannt. Einer großen Anzahl von Wissenschaftlern und Anwendern wurde der Weg zu neuronalen Netzen erst durch die Arbeit von *Rumelhart et. al.*⁴⁷ eröffnet. Doch selbst dann wurde die breite Anwendung neuronaler Netze dadurch gehemmt, dass es lange dauerte, bis kommerzielle Programme für interessierte Anwender zur Verfügung standen. Des weiteren erfordert die Nutzung lernfähiger⁴⁸ neuronaler Netze die Haltung und Verarbeitung riesiger Datenmengen. Folglich sind die Nutzungsmöglichkeiten dieses Instrumentes eng gekoppelt an die Entwicklung der Computertechnologie, welches in den letzten Jahrzehnten einen enormen Auftrieb hatte.⁴⁹ *„Dabei ermöglichten neue Prozessoren, von denen heute schon einige hundert in sogenannten Parallelcomputern kooperieren, eine Umsetzung von Ideen, die teilweise schon seit Jahren in den Schubladen der Wissenschaftler lagen, aber nicht umgesetzt werden konnten, da die entsprechende Performance fehlte.“*⁵⁰ Die Grundlagenforschung auf dem Gebiet neuronaler Netze ist längst nicht abgeschlossen. Die rasante Entwicklung in der Informationstechnologie hat aber dazu beigetragen, dass der Abstand zwischen Theorie und technischer Realisierung neuronaler Netze stark zurückgegangen ist.

⁴⁶ Vgl. Füsler (Wiesbaden 1995), S. 84-85.

⁴⁷ Vgl. Rumelhart et. al. (Cambridge 1986).

⁴⁸ Das Anspruchsvollste bei der technischen Umsetzung neuronaler Netze ist das Training.

⁴⁹ Es wird angenommen, dass dieser Auftrieb zumindest in den nächsten 2 Jahrzehnten anhalten wird.

⁵⁰ Füsler (Wiesbaden 1995), S. 4.

Die technische Realisierung oder Implementierung neuronaler Netze kann prinzipiell auf zwei unterschiedliche Weisen vorgenommen werden. Ein neuronales Netz kann zum einen auf konventionellen Rechnern rein softwaremäßig simuliert werden. Des weiteren ergibt sich die Möglichkeit, spezielle Hardware zu entwickeln.⁵¹ Die rein softwaremäßige Implementierung ist natürlich sehr kostengünstig, da sie auf vorhandene und zu üblichen Zwecken eingesetzte Rechner installiert werden kann. Softwaresimulatoren für neuronale Netze bestehen im wesentlichen aus drei Komponenten:

- Simulatorkern
- graphische Benutzerschnittstelle und
- Datenvisualisierung

Der Simulatorkern bildet, wie der Name schon sagt, das Herzstück. Dieser enthält die Netztopologie, Lernstrategien, Initialisierungs- und Aktivierungsfunktionen. Hier werden alle Operationen für die Simulation in der Trainings- und Einsatzphase durchgeführt und die Schnittstellen zu den anderen Komponenten unterstützt. Die graphische Benutzerschnittstelle gewährleistet den Dialog zwischen Benutzer und System. Dabei kann der Benutzer (meist über ein Fenster) das Netzwerk und die Arbeitsweise variieren. Die Datenvisualisierung kann als eine graphische Erklärungskomponente gesehen werden, welche die graphische Analyse von Netzwerken ermöglicht.⁵²

Entsprechend diesem Aufbauprinzip existiert heute eine kaum überschaubare Anzahl von Softwaresimulatoren von öffentlichen und privaten Organisationen. In

Abbildung 13 sind einige davon aufgelistet.

⁵¹ Vgl. Rojas (Heidelberg 1996), S. 399.

⁵² Vgl. Baumgartner (Bamberg 1998), S. 81-83.

Produkt	Beschreibung	Homepage http://www...
Stuttgarter Neuronale Netze Simulator (SNNS)	Nicht-kommerziell, frei zugänglich, plattformunabhängig, Universalwerkzeug, hergestellt am Lehrstuhl IPVR der Uni Stuttgart	Informatik.uni-stuttgart.de/ipvr/bv/projekte/neuroinformatik/abschlussbericht/node19.html
Neural Works Professional II/Plus	Kommerziell, weit verbreitetes und bekanntes Tool, nahezu plattformunabhängig, Universalwerkzeug	Neuralware.com
SENN	Kommerziell, lauffähig auf UNIX-Systemen, ursprünglich als Universalwerkzeug konzipiert, wurde dann aber auf finanzwirtschaftlichen Einsatz zugeschnitten, hergestellt von Siemens-Nixdorf	sni-usa.com/snat/senn/tutorial/
IBM Neural Network Utilità (NNU)	Kommerziell, lauffähig auf IBM eigenen Plattformen (RS/600, OS/400, OS/2), seit neuem auch lauffähig auf einigen Windowsplattformen	Research.ibm.com/people/almasi/nnu&brochure.html
NEURONET INVESTOX	Kommerziell, lauffähig auf allen gängigen Windowsversionen (ab Win 95), spezielle „Börsensoftware“ (Aktienkursprognose und Portfoliomanagement), hergestellt von KnöpfelSoftwareEntwicklung	Investox.com

Abbildung 13: Ausgewählte neuronale Softwaresimulatoren

Viele von diesen stellen Universalwerkzeuge mit einer flexiblen Aufbaustruktur dar.⁵³ Der Nachfragesog durch das erhöhte Interesse an der finanzwirtschaftlichen Nutzung neuronaler Netze bewegte die Hersteller, oftmals auch für bestimmte Finanzbereiche spezialisierte Softwaretools zu entwickeln. Aber auch kleinere Unternehmen versuchen, die Gunst der Stunde⁵⁴ auszunutzen, indem sie auf bestimmte Anwendungen (z.B. der Aktienkursprognose) zugeschnittene Softwarepakete liefern.

Obwohl Softwaresimulatoren recht einfach handhabbar und im Vergleich zu Hardwarelösungen sehr kostengünstig sind, weisen diese einige Defizite auf,⁵⁵ welche in bestimmten Fällen eine Hardwareimplementierung derzeit unumgänglich erscheinen lassen. So muss bei einer Softwaresimulation auf die massiv parallele Arbeitsweise verzichtet werden. Des weiteren stellt der Lernvorgang das Kernproblem dar. Die rein softwaremäßige Implementierung ist

⁵³ Hierbei handelt es sich um eher größere Organisationen, wie z.B. IBM, Siemens-Nixdorf oder einige Universitäten.

⁵⁴ Ein Beispiel ist das wachsende Interesse an Börsengeschäften durch Privatanleger seit der Neuemission der Telekom-Aktie.

⁵⁵ Vgl. Kratzer (München 1990), S. 138.

sehr rechen- und folglich zeitintensiv. Mit zunehmender Größe und Anzahl der Verbindungen steigt die Rechenzeit für den Lernvorgang exponentiell.⁵⁶ Würde man von einem „durchtrainierten“ Netz ausgehen, welches keine Veränderungen der Verbindungsgewichte mehr benötigt, wäre eine reine Softwaresimulation wahrscheinlich uneingeschränkt und unabhängig von der Netzgröße einsetzbar. Andererseits müsste bei gegebenen zeitlichen Ressourcen im Falle größerer Netze auf die Hardwareimplementierung zurückgegriffen werden. Besonders bei kurzfristigen Finanzprognosen, welche eine hohe Anpassungsfähigkeit bei relativ kurzen Trainingszeiten erfordern, ist die Umsetzung auf neuronale Hardware oftmals unumgänglich.

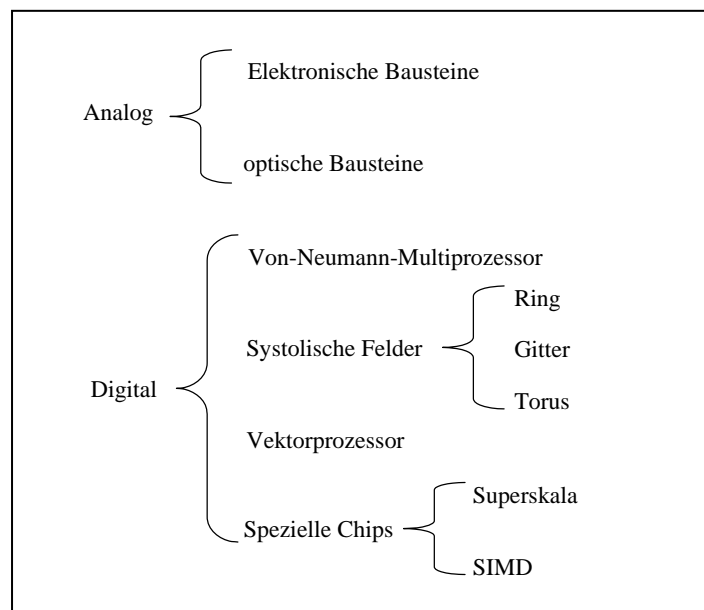


Abbildung 14: Eine Klassifikation von Neurosystemen⁵⁷

Es gibt hierfür eine Reihe von innovativen Hardwarelösungen, deren Darstellung in diesem Kontext nicht relevant ist. Stattdessen soll eine Systematisierung von neuronaler Hardware wiedergegeben werden (s. Abbildung 14). Für interessierte Leser sei auf die relevanten Literaturstellen hingewiesen.⁵⁸

Bemerkenswert ist, dass aufgrund der großen Fortschritte in der Mikroelektronik und der Optoelektronik neuronale Netze mit hoher Rechengeschwindigkeit und hoher Neuronendichte innerhalb kleinster Dimensionen realisiert werden können. So können neuronale Netze, implementiert in Mikrochips, für verschiedenste Echtzeitanwendungen in unmittelbarer Maschinennähe (z.B. Staubsauger, Bordcomputer von Autos, Motoren etc.) angebracht werden, ohne

⁵⁶ Vgl. Rojas (Heidelberg 1996), S. 399-401.

⁵⁷ Vgl. Rojas (Heidelberg 1996), S. 401.

⁵⁸ Zell (Bonn 1994), S. 431-491, Rojas (Heidelberg 1996), S. 399-425.

aufwendige Regelungstechnik einzusetzen.⁵⁹ Eine Übersicht über die Einsatz-gebiete wird in den nachfolgenden Kapiteln gegeben.

2. Einsatzfelder neuronaler Netze

Wie in vorherigen Kapiteln schon erwähnt, zeigen neuronale Netze ihre überzeugende Performance auf Einsatzgebieten, wo unscharfe und verrauschte Informationen vorliegen, die im mathematischen Kontext mehrdimensionale, nicht-lineare Abbildungen darstellen. Es ist sicherlich einleuchtend, dass es kaum Bereiche gibt, bei denen Instrumente mit solchen Eigenschaften nicht nachgefragt werden. Folglich ist das Anwendungsspektrum neuronaler Netze auch enorm groß.⁶⁰ Im folgenden sollen exemplarisch einige Gebiete vorgestellt werden, wo neuronale Netze zur Lösung bestimmter Aufgabenbereiche eingesetzt werden. Von einer Systematik kann dabei nicht die Rede sein. Vielmehr soll hier dem Leser die gegenwärtig unüberschaubare Breite der Anwendungsmöglichkeiten neuronaler Netze demonstriert werden.

- **Biomedizin:** Zur Entwicklung und Herleitung neuronaler Techniken wurde das Gehirn als Vorbild genommen. Folglich werden diese in der Medizin speziell dazu verwendet, ein besseres Verständnis für die Abläufe im Gehirn zu erhalten. In Zukunft könnten neuronale Techniken also dazu dienen, die noch weitestgehend unerforschten Abläufe im menschlichen Gehirn aufzudecken. Neuronale Techniken werden heutzutage speziell zur Diagnose und Überwachung von Gehirn und Körper herangezogen.⁶¹
- **Chemie:** In der Chemie werden neuronale Netze zur Klassifikation und zum Vergleich von Molekülen, Gemischen u.ä. benutzt. Ein konkretes Beispiel ist die Ähnlichkeitsanalyse biologisch aktiver Moleküle.⁶²
- **Steuer- und Regelungstechnik:** Die Regelungstechnik ist ein besonders sensibler Bereich im Bezug auf den Einsatz neuronaler Netze. Einige Fachleute prognostizieren, dass schon in naher Zukunft die konventionelle Regelungstechnik durch neuronale Techniken substituiert wird.⁶³ Es gibt bereits heute schon neuronale Anwendungen, welche mit konventioneller Regelungstechnik kaum mehr zu bewältigen wären. Neuronale Netze sind insbesondere für die Steuerung von autonomen Einheiten unverzichtbar. In der Robotik führen die konventionellen, sensorbasierten Regelungstechniken aufgrund der hohen Datenmengen, welche in Echtzeit zu bewältigen sind, oftmals

⁵⁹ Vgl. Rigoll (Renningen-Malmsheim 1994), S. 10-11.

⁶⁰ Vgl. Kinnebrock (München 1994), S. 103.

⁶¹ Vgl. Seraphin (München 1994), S. 203-214.

⁶² Vgl. Zell (Bonn 1994), S. 501-509.

⁶³ Vgl. Kinnebrock (München 1994), S. 110-113.

zu schwerfälligen und in der Komplexität begrenzten Bewegungsabläufen. Neuronale Netze bieten sich in diesem Zusammenhang als eine innovative Lösung an.⁶⁴ So werden neuronale Netze gegenwärtig zur Steuerung von Robotern, Raketen und unbemannten Fahrzeugen eingesetzt.

- Spracherkennung: Werkzeuge zur Spracherkennung gewinnen in letzter Zeit stark an Bedeutung. Es gibt schon in einigen alltäglich genutzten Objekten Spracherkennungssoftware. So sind viele Handys schon mit dieser Technologie ausgestattet. Selbst einige große Autohersteller haben die Zukunftstracht der Spracherkennung erkannt und gedenken, das Autofahren dadurch zu erleichtern. Jedoch mangelte es bis jetzt an der Funktionalität und der Effizienz dieser Systeme. Neuronale Netze könnten in diesem Zusammenhang zu Erfolgssprüngen beitragen.
- Meteorologie: In der Meteorologie werden neuronale Netze verwendet, um anhand von Daten, die von Satelliten empfangen werden, aktuelle klimatische Zustände zu analysieren und auf deren Grundlage Wettervorhersagen zu treffen.
- Optische Bild- und Zeichenerkennung: Da neuronale Netze mit verrauschten und inkonsistenten Informationen gut zurecht kommen, eignen sich diese besonders zur Muster- und Zeichenerkennung. Folglich setzen Versicherungen, Banken und Behörden neuronale Netze ein, um die massenhafte Bearbeitung von Formularen ökonomisch zu bewältigen.⁶⁵ Weiterhin werden neuronale Netze zur Zuordnung von Handschriften und zur Erkennung von Unterschriften herangezogen. Ursprünglich kamen neuronale Netze im militärischen Bereich zum Einsatz. Zum Zwecke der Ortung übernahmen neuronale Netze die Aufgabe, Ziele zu klassifizieren, Aufklärungsbilder auszuwerten und Sonardaten zu interpretieren.⁶⁶

Darüber hinaus gibt es noch eine Reihe von Anwendungsfeldern, auf die hier nicht näher eingegangen werden soll. Dazu gehören u.a. Datenkompression und Datenaufbereitung in der Informatik, Herleiten von Spielstrategien und musikalisches Komponieren.

Zu berücksichtigen ist hierbei, dass viele Anwendungsmöglichkeiten noch nicht einmal bekannt sind, da viele (zivile und militärische) Organisationen aus Konkurrenzgründen ihre Aktivitäten im Bereich der Neuroinformatik geheim halten.⁶⁷ Die konkreten betriebs-

⁶⁴ Vgl. Zell (Bonn 1994), S. 541-554.

⁶⁵ Vgl. Haun (Renningen-Malmsheim 1998), S. 128.

⁶⁶ Vgl. Kratzer (München 1990), S. 172.

⁶⁷ Vgl. Kinnebrock (München 1994), S. 104.

wirtschaftlichen Anwendungsgebiete wurden hier fast vollständig ausgelassen. Eine grobe Darstellung dieser Einsatzmöglichkeiten soll im nachfolgenden Kapitel wiedergegeben werden.

3. Neuronale Netze in der Betriebswirtschaft

Nachdem im vorherigen Kapitel die gesamten, eher technischen Einsatzmöglichkeiten neuronaler Netze genannt wurden, sollen hier die Anwendungen in betriebswirtschaftlichen Bereichen dargestellt werden. Eine einfache Klassifizierung der Einsatzfelder ist auch hier nicht möglich. So wird z.B. die im vorherigen Kapitel genannte Muster- und Signalerkennung in der Betriebswirtschaft zur Qualitätskontrolle in Qualitätsmanagementsystemen herangezogen. Oftmals werden die betrieblichen Funktionsbereiche zur Klassifikation betriebswirtschaftlicher Anwendungen neuronaler Netze eingesetzt.⁶⁸ Die betrieblichen Funktionsbereiche können aufgeteilt werden in Marketingbereich, Produktionsbereich und Finanzbereich. Die Anwendungen neuronaler Netze im Finanzbereich, als untergeordnetes Gebiet der Betriebswirtschaft, bilden den Schwerpunkt dieser Arbeit und werden noch einer intensiven Aufarbeitung unterworfen. Hier sollen nun einige Anwendungen im Marketing- und im Produktionsbereich exemplarisch aufgezeigt werden.

- Materialbedarfsprognose: Die Aufgabe der betrieblichen Materialwirtschaft ist die zeit- und mengenmäßige Bereitstellung von Materialien, welche zur Leistungserstellung und Leistungsverwertung benötigt werden. In modernen PPS-Systemen (Produktionsplanung- und -steuerung) werden neuronale Netze zur Prognose von Zeitreihen herangezogen. Hierbei gibt es eine Reihe von Konzepten mit variierender Größe. So werden teilweise mehrere neuronale Netze miteinander gekoppelt, wobei z.B. ein Netz den Gesamtabsatz prognostiziert und ein weiteres den Materialbedarf.⁶⁹ Weiterhin können die Aspekte der Produktion auf Lager und die optimale Lagerhaltung von Roh-, Hilfs-, Betriebsstoffen und Fertigwaren mitberücksichtigt werden.
- Ablaufplanung: In PPS-Systemen, werden neuronale Netze auch eingesetzt, um Fertigungsaufträge betrieblichen Ressourcen zuzuordnen. Als Beispiel sei hier das System NEPAL genannt, welches aus mehreren kooperierenden Netzen besteht und die Ab-

⁶⁸ Vgl. Adam (Münster 1995), S. 591-592 und Corsten/May (Wiesbaden 1996), S. 1.

⁶⁹ Vgl. Lackes (München 2000), S. 210-230.

laufplanung in drei Schritten (Variantenauswahl, Maschinenauswahl und Zeitplanung) löst.⁷⁰ Die Ablaufplanung ist ein typisches Optimierungsproblem.

- Travelling-Salesman-Problem (TSP): Das TSP steht für eine ganze Reihe von kombinatorischen Optimierungsproblemen. Dazu gehört neben der Ablaufplanung die Ressourcenallokation, welche u.a. die Maschinenbelegungsplanung und die Zeit- und Arbeitsplanung beinhaltet.⁷¹
- Personalbeurteilung: Neben dem altbewährten Vorstellungsgespräch werden in modernen Personalmanagement-Systemen neuronale Netze verwendet, um aus verschiedenen Kennziffern die Eignung eines Bewerbers für das Unternehmen festzustellen.⁷² Sicherlich werden neuronale Netze in diesem Bereich die konventionellen Auswahlverfahren nicht ersetzen. Sie können aber als Entscheidungs-unterstützungssystem sehr nützlich sein.
- Absatz- und Umsatzprognose: Die enorme Performance neuronaler Netze bei Prognosen machen es im Vertrieb zu einem nachgefragten Instrument. *Düsing* stellt ein Modell auf Basis neuronaler Netze vor, welches für einen Einzelhandelsunternehmen eine Prognose von Absatz und Umsatz mit einem Prognosehorizont von einer Woche vornimmt.⁷³
- Marktsegmentierung: Die Marktsegmentierung ist Aufgabenbereich der Marktforschung. Ziel der Marktsegmentierung ist es, einen Gesamtmarkt in möglichst homogene Klassen einzuteilen, um die Effizienz der Marketingaktivitäten zu steigern. Neuronale Netze kommen in diesen Bereichen als Klassifikatoren zum Einsatz. So werden z.B. in der Kraftfahrzeug-Versicherungsbranche speziell seit der Aufhebung der gesetzlichen Tarifvorschriften im Jahre 1994 neuronale Netze dazu benutzt, Kunden in Risikogruppen einzuordnen und individuell maßgeschneiderte Tarife zu liefern.⁷⁴

Des Weiteren werden neuronale Netze zur Beurteilung von Marktrisiken und zur Prognose von Marktreaktionen herangezogen.

Bei ausreichend guter Informationsversorgung sind die Einsatzpotenziale neuronaler Werkzeuge praktisch unbegrenzt. Da auch in Zukunft von einer Effizienzsteigerung der Ressource

⁷⁰ Vgl. Märtens/Sauer (Wiesbaden 1998), S. 205-214.

⁷¹ Vgl. Wilbert (Frankfurt a. M. 1996), S. 144-147.

⁷² Vgl. Lackes (München 2000), S. 192-210.

⁷³ Vgl. Düsing (Hamburg 1997), S. 166-187.

⁷⁴ Vgl. Lackes (München 2000), S. 230-253.

„Information“ auszugehen ist, wird sich der Einsatz neuronaler Netze in der Betriebswirtschaft wahrscheinlich intensivieren.

B. Finanzwirtschaftliche Prognoseanwendungen

I. Grundlagen neuronaler Finanzprognosen

1. Zur prinzipiellen Prognostizierbarkeit von Finanzmarktdaten

Die Finanzprognose ist ein Problembereich, wo die Entscheidungsträger nur wenige aussagekräftige Informationen über zukünftige Entwicklungen von Wertpapierkursen zur Verfügung haben. Entscheidungen werden oftmals auf der Basis von unvollständigen und subjektiven Informationen getroffen. Weiterhin werden die Entscheidungsträger häufig mit Daten überflutet, die aus Gründen zeitlicher Restriktionen und hoher Komplexität nicht zu wertvollen Informationen verdichtet werden können. Jede exklusive Information könnte aber sofort in Geld transformiert werden. Folglich ist die Suche von Finanzinstituten nach neueren Ansätzen zur Gewinnung von neuen und exklusiven Informationen aus Datenpools groß. Neuronale Netze stellen in diesem Zusammenhang ein vielversprechendes Instrument dar. Doch vorher stellt sich die grundsätzliche Frage, ob Prognosen bzgl. Finanzmarktdaten überhaupt möglich sind. Nach der „efficient market theory“ machen Finanzprognosen keinen Sinn. Die „Theorie effizienter Märkte“ besitzt auch heute noch viele Anhänger. Sie ist aus dem Modell des vollkommenen Wettbewerbs und seinen Prämissen für Gütermärkte entwickelt worden. Zu diesen Prämissen gehören u.a. vollständige Transparenz, unendlich schnelle Informationswahrnehmung und sofortige Reaktion aller Marktteilnehmer.⁷⁵ Dieterle beschreibt die Hypothese folgendermaßen:

„Die Hypothese effizienter Märkte geht davon aus, dass alle Informationen, welche die Preisbildung von Wertpapieren beeinflussen, auch in den Wert der Papiere eingehen, weswegen an effizienten Märkten keine kursbestimmenden Informationen verloren gehen bzw. bei der Preisfindung nicht beachtet werden und es unmöglich ist, mittels Handelsentscheidungen aufgrund bekannter Informationen Zusatzgewinne zu realisieren.“⁷⁶

Jede vorhandene Information ist demnach schon in den Preisen der Wertpapiere integriert. Folglich ist die zukünftige Entwicklung rein zufällig. Jedoch ist diese Theorie in der Praxis sicherlich kaum haltbar. Zum einen existieren solche effizienten Märkte überhaupt nicht. Neue Informationen erreichen nicht jeden Anleger, mal ganz abgesehen vom simultanen Antreffen der Informationen. Des Weiteren werden Informationen unterschiedlich wahrgenommen, was von der Intelligenz eines jeden abhängt. Von einer unendlich schnellen Reaktion

⁷⁵ Vgl. Widdel (München 1996), S. 21-24.

⁷⁶ Dieterle (Berlin 1999), S. 45.

kann genauso wenig die Rede sein, da Kauf- und Verkaufsentscheidungen situationsbedingt und entsprechend den menschlichen Denkprozessen Verzögerungen unterliegen.⁷⁷

Dennoch leuchtet die Theorie effizienter Märkte zumindest auf den ersten Blick, beim scheinbar regellosen Auf und Ab der Aktien- und Devisenkurse ein. Diese sehen wie zufällig aus und wären demnach nicht prognostizierbar. Aber was diesen Anschein hat, muss nicht zwangsläufig zufällig sein. Ein gutes Beispiel hierfür ist der Zufallszahlengenerator eines Rechners. Kein Mensch kann hier eine Gesetzmäßigkeit erkennen. Dennoch werden diese Zahlen nach einer Formel berechnet. Es handelt sich hier also nur um scheinbare Zufallszahlen. Daher werden sie auch als Pseudo-Zufallszahlen bezeichnet.⁷⁸ Es stellt sich daher die Frage, ob die Kursverläufe nur pseudo-zufällig sind. Aus mathematischer Sicht würde die Frage lauten: Ist die Wahrscheinlichkeitsverteilung der Kursverläufe normalverteilt oder liegt eine andere, möglicherweise artverwandte Häufigkeitsverteilung vor? Dieser Frage gingen schon einige Wissenschaftler nach. So will *E. Peters* herausgefunden haben, dass bei den wöchentlichen Erträgen des Standard & Poors Index der New York Stock Exchange von 1928 bis 1989 keine Normalverteilung, sondern die ähnelnde leptokurtosische Häufigkeitsverteilung vorliegt. Dies zeigt, dass Preisänderungen entgegen der Annahme der Effizienzhypothese nur bedingt zufällig sind. Vielmehr scheinen bestimmte Tendenzen und Trends vorzuliegen.⁷⁹ Bei der Betrachtung der Entwicklung von Wertpapierkursen anhand von Charts, kann man feststellen, dass eine gewisse Selbstähnlichkeit unabhängig vom Betrachtungshorizont ein Tag, eine Woche oder ein Jahr vorliegt. Es sind gewisse, von Zeit zu Zeit wiederkehrende Muster, welche zwar nie ganz identisch sind, sich aber doch sehr ähneln. Peters spricht in diesem Zusammenhang von Fraktalen.⁸⁰ Die Fraktaltheorie entstammt aus den nicht-linearen, dynamischen Systemen und wird in der volkstümlichen Sprache auch als Chaostheorie bezeichnet. Die Chaostheorie untersucht dynamische Systeme, deren Bewegungen zufällig aussehen, denen aber ein bestimmtes nicht-lineares Wirkungsgesetz zugrunde liegt. Man spricht hier auch von einem deterministischen Chaos.⁸¹ Sind Wirkungsgesetz und Ausgangszustand bekannt, kann demnach das Verhalten eines Systems vorhergesagt werden. Folglich wäre es möglich, mit entsprechenden Rechnerkapazitäten Kursbewegungen auf den Finanzmärkten zu prognostizieren. Trotzdem ist die Prognostizierbarkeit von Finanzmarktdaten

⁷⁷ Vgl. Thoma (München 1996), S. 20-22.

⁷⁸ Vgl. Uhlig (München 1995), S. 1.

⁷⁹ Vgl. Peters (New York 1994), S. 22-24.

⁸⁰ Vgl. Uhlig (München 1995), S. 118.

⁸¹ Vgl. Uhlig (München 1995), S. 3.

nicht endgültig geklärt. Ein wesentliches, erkanntes Problem ist die Instabilität auf Märkten. Kapitalmärkte sind nämlich nicht stationär. Das bedeutet, dass die Regeln, nach denen Preisbewegungen ablaufen, sich verändern.⁸² Weiterhin ergeben sich kursbeeinflussende Ereignisse, wie z.B. Kriege oder Naturkatastrophen, die natürlich kaum prognostizierbar sind.

Will ein Finanzinstitut also erfolgreich sein, muss ein intelligentes und adaptives Werkzeug mit höchst flexiblen Eigenschaften herangezogen werden, welches auch den Anforderungen des stetigen Wandels gerecht wird. Dass neuronale Netze hierfür besonders gut geeignet sind, wird in den folgenden Kapiteln näher zu sehen sein.

2. Klassische Ansätze zur Finanzprognose

Die bedeutendsten Anwendungsgebiete der Finanzprognose und -analyse finden sich in der Zins-, Wertpapier- und der Wechselkursproblematik. Mit den verwendeten Techniken wird versucht, mögliche Trends zu identifizieren und zu analysieren. Auf deren Grundlage werden dann Modelle realisiert, welche Simulationen von Zeitreihen und somit Prognosen ermöglichen. Im folgenden Abschnitt sollen die in der Praxis am häufigsten angewandten Methoden zur Finanzprognose vorgestellt werden. Diese können in drei Klassen unterteilt werden: technische Analyse, fundamentale Analyse und die „intuitiv, intellektuellen“ Methoden (s. Abbildung 15).

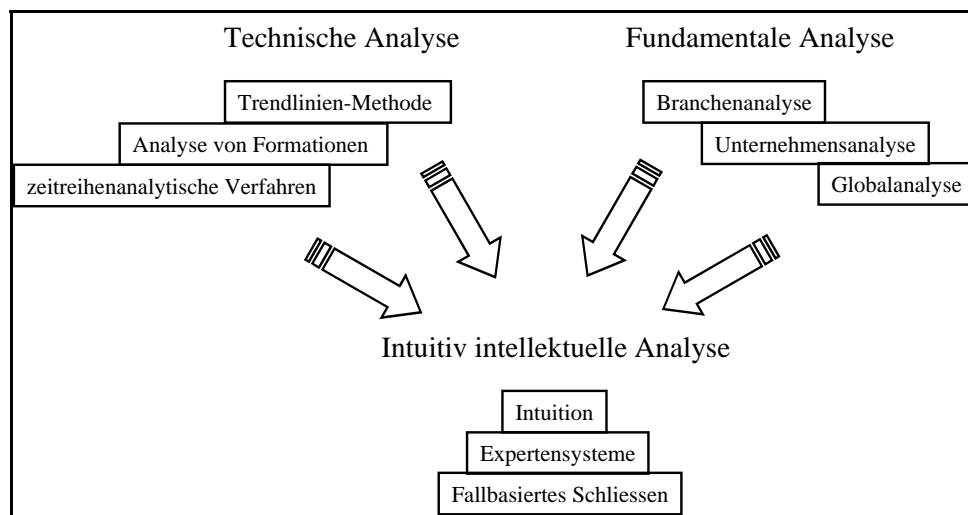


Abbildung 15: Übersicht klassischer Ansätze zur Finanzprognose

Die technische Analyse von Aktien und Währungen befasst sich ausschließlich mit den historischen Kursentwicklungen. Dabei wird davon ausgegangen, dass alle kursbeeinflussenden Faktoren sich in den Kursentwicklungen widerspiegeln. Folglich können diese Faktoren auch

⁸² Vgl. Baumgartner (Bamberg 1998), S. 45.

aus den Kursentwicklungen herausgelesen werden. Dazu werden die Kursentwicklungen durch graphische Hilfsmittel, sog. Charts, dargestellt.⁸³ Die Grundlage der technischen Analyse bildet die Dow-Theorie. Hierbei handelt es sich eher um eine Heuristik, welche auf der Beobachtung von Zeitreihen gründet, als um eine ökonomische Theorie. Danach verlaufen Märkte in Trends. Diese Trends aufzudecken und Trendumkehrungen rechtzeitig zu erkennen, ist Aufgabe der technischen Analyse.⁸⁴ Die technische Analyse wird in der Literatur weiterhin untergliedert in:

- Trendlinien-Methode
- Analyse von Formationen und
- zeitreihenanalytische Verfahren.

Bei der Trendlinienmethode werden Trendgeraden in Charts eingezeichnet. Damit hofft man, zusätzliche Informationen bzgl. Timing von Käufen und Verkäufen zu verschaffen. Mit der Analyse der Formationen wird versucht, typische, oft auftretende Muster im Kursverlauf zu identifizieren, die bestehende Trends bestätigen oder auf eine Trendumkehr hinweisen.⁸⁵ Die Analyse der Formationen ist deshalb interessant, weil sie in die modernen Ansätze der Finanzanalyse passen, welche von nicht-linearen und fraktalen Prozessen von Zeitreihen der Finanzmarktdaten ausgehen (s. Kap. 2.1.1).

Bei den zeitreihenanalytischen Verfahren werden ebenfalls historische Kursentwicklungen als Informationsbasis genommen. Sie gehen folglich auch nicht über den Ansatz der technischen Analyse hinaus. Diese Verfahren werden nicht durch grafische Darstellungen gestützt, sondern sind mathematisch und statistisch fundiert.⁸⁶ Daher spricht man auch von der statistisch-technischen Analyse. Wichtige Verfahren, die öfter als Benchmarks für neuronale Netze genommen werden, sind die ARIMA-Modelle und die Spektralanalyse. Das ARIMA-Modell wird dazu eingesetzt, historische Zeitreihen durch zwei überlagerte Prozesse (AR und MA⁸⁷) aus sich selbst heraus zu erklären.⁸⁸

Die Spektralanalyse entstammt den Naturwissenschaften. Analog zum Sonnenlicht, das sich aus einzelnen Komponenten zusammensetzt, werden Zeitreihen von Finanzmarktdaten als

⁸³ Vgl. Miller (Mannheim 1994), S. 16.

⁸⁴ Vgl. Steiner/Bruns (Stuttgart 1994), S. 280-281.

⁸⁵ Vgl. Poddig (Uhlenbruch 1999), S. 99-104.

⁸⁶ Vgl. Baumgartner (Bamberg 1998), S. 45.

⁸⁷ Autoregressive Prozesse und Moving Average Prozesse.

⁸⁸ Vgl. Loistl (München 1994), S. 157-166.

Überlagerungen zyklischer Bewegungen gesehen, welche die einzelnen Einflussfaktoren darstellen.⁸⁹

Im Gegensatz zur technischen Analyse orientiert sich die fundamentale Analyse nicht an historischen Kursentwicklungen, sondern an den fundamentalen Einflussfaktoren. Danach hängen die Wertpapierkurse von der Ertragskraft der jeweiligen Unternehmen ab. Ziel dieser Methode ist es folglich, die fundamentalen Einflussfaktoren zu analysieren und die Auswirkungen frühestmöglich abzuschätzen. Die Einflussfaktoren bei der fundamentalen Analyse werden in der Literatur oftmals in drei Analyseebenen aufgeteilt (s. Abbildung 16).⁹⁰

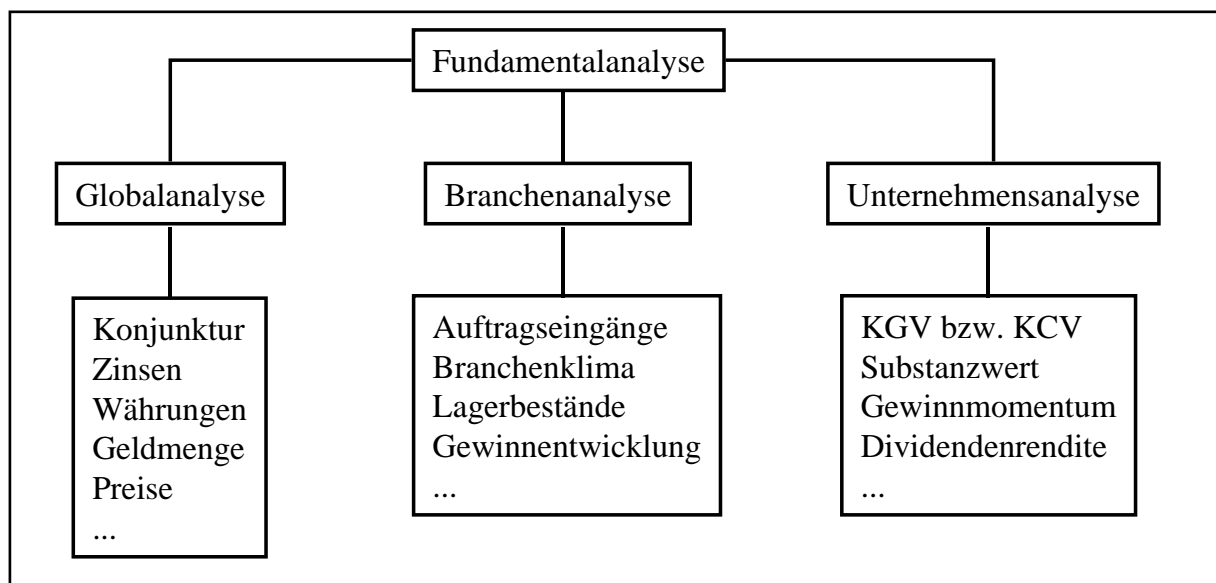


Abbildung 16: Systematik fundamentaler Einflussfaktoren

Die intuitiv intellektuellen Methoden basieren auf den biologischen neuronalen Netzen des menschlichen Gehirns. Intuition und Gefühle ersetzen in der Realität, zum Teil bedingt durch zeitliche Restriktionen, oftmals die Rationalität in der Entscheidungsfindung. So nannten Börsenmakler nach einer Studie der Webster Universität Wien an erster Stelle das Gefühl als Entscheidungsgrundlage.⁹¹ Dieses wird aber wiederum beeinflusst von der Erfahrung, die ein Mensch u.a. auch durch den Einsatz der oben dargestellten fundamentalen und technischen Analysen gemacht hat. In der Vergangenheit hat es schon viele Anstrengungen mit der Absicht gegeben, diese Erfahrung einer breiteren Masse von Anwendern zugänglich zu machen. Diesen Zweck verfolgen Expertensysteme, welche auch als wissensbasierte Systeme bezeichnet werden. Hier wird das Wissen von Experten soweit möglich in logische Regeln gefasst.

⁸⁹ Vgl. Baumgartner (Bamberg 1998), S. 51-52.

⁹⁰ Vgl. Obst (Stuttgart 2000), S. 861-864.

⁹¹ Vgl. Hocking/Oberlechner (Wien 1996), S. 10.

Somit wird Nichtexperten ermöglicht, das Wissen von Experten für einen Problembereich (zumindest teilweise) zu adaptieren.⁹²

Das Fallbasierte Schliessen (Case Based Reasoning) ist dem Expertensystem sehr ähnlich und wird in der Literatur auch teilweise als solches bezeichnet. Hier werden die gesammelten Erfahrungen als Fälle in Datenbanken gespeichert. Aktuelle Fälle werden mit denen aus der Datenbank verglichen. Beim Vorliegen eines ähnlichen Falles wird das darüber verfügbare Wissen auf den neuen Fall angewendet.⁹³

Ein wesentliches Problem intuitiv intellektueller Methoden ist die Übertragbarkeit. So können diese Methoden nur durch vorhandenes Wissen und durch Vergleiche mit vergangenen Fällen gelöst werden. Daher stellt die Marktdynamik ein wesentliches Problem dar. Ein flexibles und adaptives System, wie es die neuronalen Netze mit ihrer Lern- und Generalisierungsfähigkeit darstellen, könnten hierbei Abhilfe schaffen. Wie neuronale Netze zur Prognose von Zeitreihen modelliert und eingesetzt werden, wird im nächsten Kapitel dargestellt.

3. Vorgehensweise bei der Modellierung neuronaler Netze zur Finanzprognose

Für eine Vielzahl von Einsatzgebieten in der Finanzwirtschaft benötigt man heutzutage überhaupt kein detailliertes Wissen mehr über Aufbau und Funktionsweise von neuronalen Netzen. Denn es sind inzwischen zahlreiche Programme im Handel (s. Kap. 1.3.1), mit denen der reine Anwender ohne weitere Probleme neuronale Netze (z.B. für die Aktienkursprognose) einsetzen kann. Hier soll aber davon abgesehen und ein Entwicklungsschema skizziert werden, um das Verständnis über dieses Instrumentarium auszuweiten. *„Was es zu wissen gilt, ist die Methodik ... mit der eine konkrete Problemstellung in eine Neuronale Netzanwendung überführt werden kann.“*⁹⁴

In diesem Zusammenhang ist es wichtig, zu wissen, dass die Entwicklung von neuronalen Netzen mit relativ großen Risiken bezüglich Kosten und der letztlich zu erwartenden Leistungsfähigkeit behaftet ist. So wird die erreichte Leistungsfähigkeit, abgesehen von der prinzipiellen Eignung, hauptsächlich beeinflusst von der Qualität, dem Umfang der verfügbaren Daten und den meist zufällig gewählten, aktuellen Parametereinstellungen des Netzes.⁹⁵

⁹² Vgl. Schimpe (Aachen 1997), S. 8-10.

⁹³ Vgl. Biendarra (Frankfurt a. M. 1998), S. 84-86.

⁹⁴ Haun (Renningen-Malmsheim 1998), S. 192.

⁹⁵ Vgl. Füser (Wiesbaden 1995), S. 41.

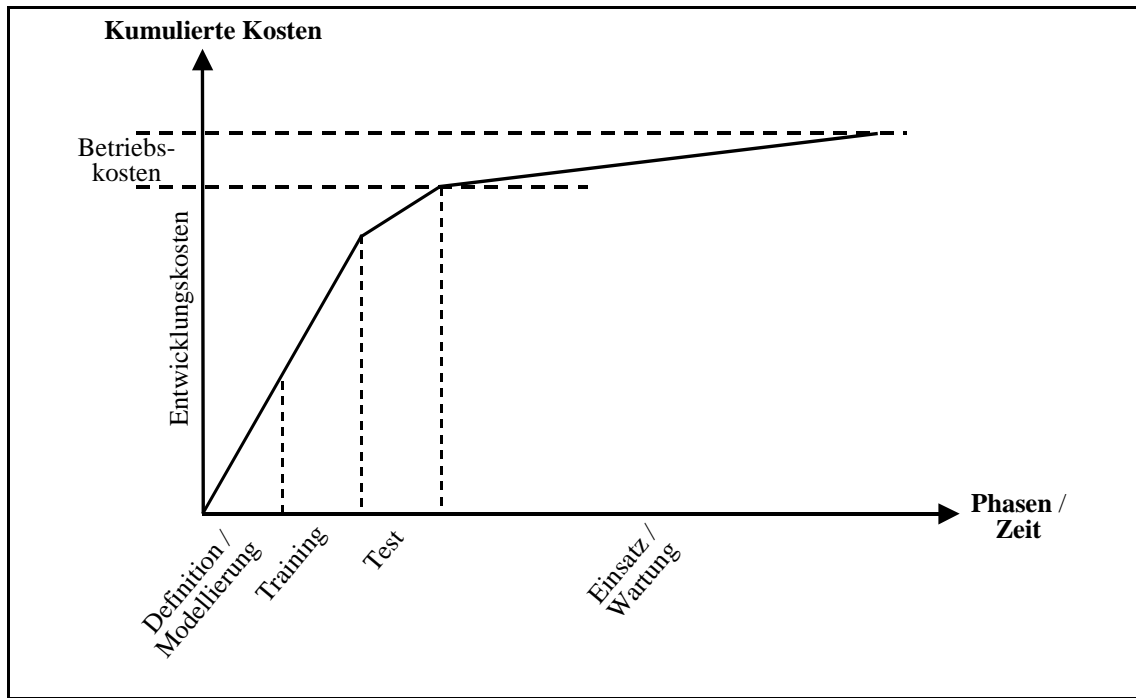


Abbildung 17: Kostenverlauf in den Lebensphasen eines neuronalen Netzes⁹⁶

Bei der Betrachtung der Kostenverlaufskurve in

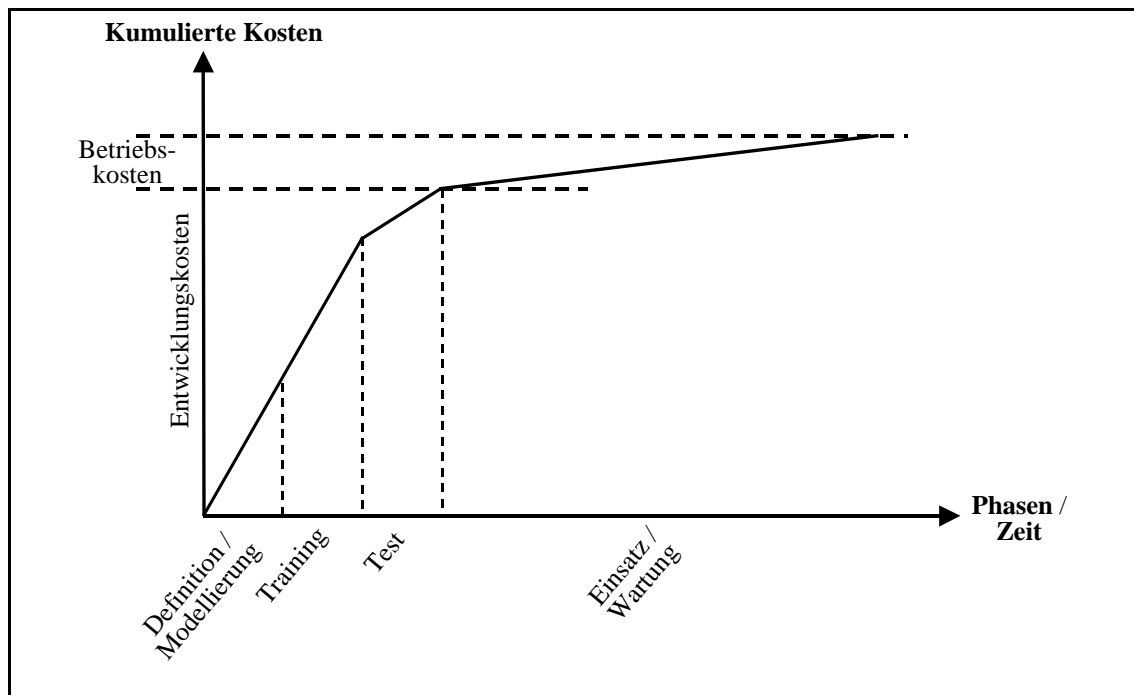


Abbildung 17 wird das Risiko nochmals deutlich. So setzen sich ca. 80% der Gesamtkosten in der Gesamt-Lebensphase eines neuronalen Netzes aus den Entwicklungskosten zusammen. In der Literatur finden sich sicherlich auch aus diesem Grund eine Reihe von „erfolgverspre-

⁹⁶ In Anlehnung an Lackes (München 2000), S. 69. Der Kostenverlauf im Diagramm ist weder allgemeingültig noch maßstabstreu und dient lediglich der groben Skizzierung der Kostenverteilung im Lebenszyklus eines neuronalen Netzes.

chenden“ Skizzen zur Vorgehensweise bei der problemspezifischen Entwicklung eines neuronalen Netzes zur Finanzprognose. Diese sind aber kritisch zu hinterfragen. So soll an dieser Stelle eine grobe Skizze der Entwurfsmethodologie aufgezeigt werden. Diese besteht aus 6 Phasen (s. Abbildung 18) und stellt

den Versuch dar, Entwicklungstheorien diverser Autoren zu vereinigen.⁹⁷ Bei dieser, aus theoretischer Sicht streng deterministischen Reihenfolge der Entwicklungsstufen, sind in der Praxis Rückschritte jedoch kaum zu vermeiden.⁹⁸

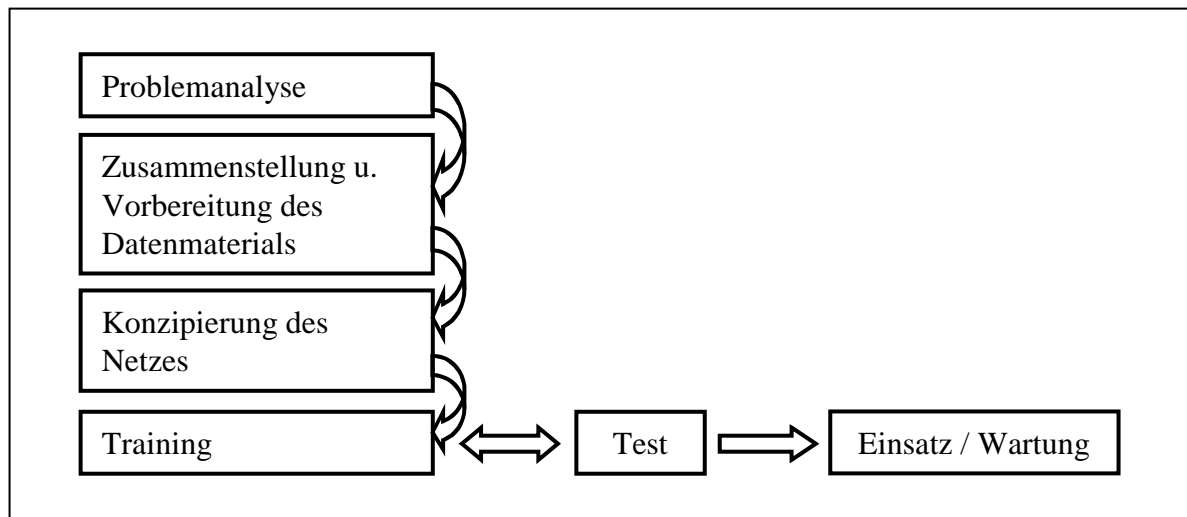


Abbildung 18: Entwicklungsphasen neuronaler Netze

Die Problemanalyse bildet den Ausgangspunkt bei der Entwicklung eines neuronalen Netzes. Hier ist die jeweilige Problemstellung zu klären. Dabei ist zu berücksichtigen, dass für wohlstrukturierte Probleme (Ursachen-Wirkungszusammenhänge sind einfach und übersichtlich) der Einsatz konventioneller Instrumentarien sinnvoller ist. Des Weiteren müssen bei den Daten die Regelmäßigkeiten überwiegen. Zufällige und chaotische Einflüsse dürfen also nicht dominieren.⁹⁹

Wie schon im vorherigen Kapitel erwähnt, sind die Ursachen-Wirkungszusammenhänge bei finanzwirtschaftlichen Zeitreihen keineswegs einfach und übersichtlich. Es liegen in der Regel nicht-lineare Wechselbeziehungen vor, was für den Einsatz neuronaler Netze spricht. Ferner ist die Stationarität auf dem Problemgebiet zu überprüfen. Herrschen immer die gleichen Gesetze am Markt oder ist mit Strukturbrüchen im Betrachtungszeitraum zu rechnen? Dies sind Fragen, welche in der Problem-Analyse-Phase zu beantworten sind. In der Finanzprog-

⁹⁷ Vgl. Corsten, H. (1996), Füser, K. (1995), Haun, M. (1998), Heuer, J. (1997), Poddig, D. (1999) und Baumgartner, P. (1998).

⁹⁸ Vgl. Füser (Wiesbaden 1995), S. 42.

⁹⁹ Vgl. Corsten (Wiesbaden 1996), S. 6.

nose stellt dies eine potentielle Gefahr für jedes konventionelle Prognose-instrument dar. Betrachtet man den enormen Wandel in der Wirtschaft (EU-Binnenmarkt, Währungsunion, Reform von Wirtschaftsgesetzen, Änderungen des Konsumentenverhaltens aufgrund der IKT) ist eher von der Varianz der zu prognostizierenden finanzwirtschaftlichen Zeitreihe auszugehen, was ein flexibles und adaptives Instrument erfordert. Neuronale Netze stellen durch ihre Lernfähigkeit ein solches Instrument dar.

Weiterhin sind in der Phase der Problemanalyse die relevanten Einflussvariablen (Inputvariablen) und die erwünschte Ausgabe (Outputvariablen) festzulegen. Hierzu wird in der Regel ein Experte benötigt, da ohne hinreichende Erfahrung und Grundwissen eine Analyse der realen Problemsituation nicht möglich ist. Sind einmal die Einflussvariablen und der grobe Zusammenhang zwischen Input und Output bekannt, wird prinzipiell keine weitere Information benötigt. Im Gegensatz zu den konventionellen Prognoseinstrumenten, welche ein wohldefiniertes Modell benötigen, leiten neuronale Netze aus den zugrundeliegenden Daten das Modell selbst her. Eine gänzliche Modellfreiheit ist natürlich nicht vorhanden, da die Auswahl der Daten schon einen Modell impliziert.¹⁰⁰ In der Finanzprognose muss aufgrund der Komplexität der Finanzmärkte mit einer sehr hohen Zahl von Inputvariablen gerechnet werden. So müssen bei der Dollar/DM-Wechselkursprognose um die 300 Einflussfaktoren verarbeitet werden. Während viele konventionelle Instrumente damit überlastet wären, können solch große Datenmengen in neuronalen Netzen durch die parallele Arbeitsweise innerhalb kürzester Zeit bewältigt werden.¹⁰¹ Jedoch hat es sich in der Praxis als sinnvoll erwiesen, Korrelationsanalysen durchzuführen. Es existieren direkte Zusammenhänge zwischen Einflussfaktoren, so dass deren Zeitreihen synchron verlaufen. Um den Rechenaufwand zu reduzieren, empfiehlt es sich, diese bis auf einige wenige repräsentative Einflussfaktoren wegzustreichen.

Anschließend sind die Datenquellen zu bestimmen, aus denen man für das spätere Training ausreichende Samples herausnehmen kann. Dies ist für die zukünftige Leistungsfähigkeit eines neuronalen Netzes entscheidend. Daher sind diese Daten auf Qualität und Quantität zu überprüfen. Die Qualität wird beeinflusst vom Anteil fehlerhafter, nicht repräsentativer Daten. Schlechte Qualität der Daten spiegelt sich in der Arbeitsweise des „falsch trainierten“ neuronalen Netzes in Form von falschen Ausgaben wider. Bezüglich der Quantität ist zu gewährleisten, dass eine ausreichende Menge an Datensätzen vorliegt. Eine unter Autoren verbreitete „Faustregel“ besagt, dass die Anzahl der Samples mindestens das 10-fache der Anzahl von

¹⁰⁰ Vgl. Baumgartner (Bamberg 1998), S. 93.

¹⁰¹ Vgl. Miller (Mannheim 1994), S. 23.

Verbindungsgewichten betragen sollte.¹⁰² Als nächstes muss dann ein Konzept entworfen werden, das die Eingabe der Datensätze in das neuronale Netz ermöglicht. Zum Aufgabenbereich der Datenpräsentation¹⁰³ gehört die Codierung der Daten. Da neuronale Netze nur mit reellen Zahlen arbeiten können, müssen die Eingangs- und Ausgangsinformationen in solche Zahlen transformiert werden.¹⁰⁴ Hierzu müssen die, dem Netz zuzuführenden Daten, innerhalb eines Intervalls abgebildet werden. Dies bedarf einer Skalierung der Daten. Die Datennormierung dient der symmetrischen Verteilung der Daten auf diesem Intervall, wodurch bessere Ergebnisse erzielt werden können.¹⁰⁵

Nach der Aufbereitung der Daten für die Präsentation kann mit dem Entwurf des neuronalen Netzes begonnen werden. Hier ist u.a. folgendes festzulegen:

- Netzwerkmodell mit Lernregel
- Netztopologie
- Anfangsinitialisierung der Verbindungsgewichte
- Neuronstruktur mit Aktivierungsfunktion

Für den problemspezifischen Aufbau eines optimalen neuronalen Netzes gibt es immer noch keine theoretisch untermauerte Vorgehensweise. Die Erfahrung und Kreativität des Designers ist in diesem Fall ausschlaggebend.¹⁰⁶ Es gibt aber mehrere Ansätze, um geeignete Netzwerkarchitekturen zu entwickeln. Auf die Darstellung soll aber nicht eingegangen werden, da es den Rahmen dieser Arbeit sprengen würde.

Nach der Festlegung der Netzwerkdimensionen kann mit dem Training begonnen werden. Hierfür wird der Gesamtdatenbestand aufgeteilt in Trainings- und Testdaten. Während dem Training lernt das Netz beim überwachten Lernen aus den vorgegebenen Eingangs- und Ausgangsinformationen. Ziel ist es dabei, eine hohe Generalisierungsfähigkeit des Netzes zu erreichen. „Die Generalisierungsfähigkeit beschreibt die Genauigkeit, mit der das Neuronale Netz Prognosen für Daten außerhalb der Trainingsmenge gestalten kann.“¹⁰⁷ Ein Problem stellt dabei das sog. Overfitting dar.¹⁰⁸ Beim Overfitting lernt das Netz auswendig und verliert dabei die Generalisierungsfähigkeit. Diese wird auch durch die Anzahl der verdeckten Neuro-

¹⁰² Vgl. Corsten (Wiesbaden 1996), S. 8.

¹⁰³ Die Daten werden dem neuronalen Netz präsentiert.

¹⁰⁴ Vgl. Baumgartner (Bamberg 1998), S. 101-104.

¹⁰⁵ Vgl. Zimmermann (München 1994), S. 21-23.

¹⁰⁶ Vgl. Füser (Wiesbaden 1995), S. 44.

¹⁰⁷ Baumgartner (Bamberg 1998), S. 124.

¹⁰⁸ Vgl. Poddig (Uhlenbruch 1999), S. 422-426.

nen beeinflusst. „Viele“ verdeckte Neuronen erlauben eine bessere Approximation der Zusammenhänge, wobei „zu viele“ tendenziell zum Overfitting führen.¹⁰⁹ Von dieser Problematik soll aber abgesehen und vielmehr das Problemlösungspotenzial durch das Datenhandling gezeigt werden. Um ein Optimum an Generalisierungsfähigkeit zu erreichen, wird das Konzept der Kreuzvalidierung eingesetzt. Hierzu werden meistens aus den Trainingsdaten zusätzliche Kreuzvalidierungsdaten entnommen. Zum eigentlichen Lernvorgang dienen die Trainingsdaten. Hier werden die Verbindungsgewichte verändert. Nach jeder Trainingsphase wird das Netz mit den Kreuzvalidierungsdaten gespeist, um die Generalisierungsfähigkeit zu bestimmen.¹¹⁰ Dieser Zyklus wird iterativ fortgesetzt. Kann keine weitere Verbesserung der Generalisierungsfähigkeit durch Training erreicht werden, so wird der Lernvorgang gestoppt. Bei nichtzufriedenstellender Netzleistung müssen dann vorherige Phasen wiederaufgegriffen werden.

Bei befriedigender Generalisierungsfähigkeit wird das Netz der Abschlussprüfung durch die Testdaten unterzogen. Läuft auch diese erfolgreich ab, kann das neuronale Netz eingesetzt werden. In diesem, als Operatingphase bezeichneten Lebensabschnitt, werden die Inputvariablen entsprechend den Kodierungsvorschriften transformiert und in das Netz eingegeben. Da neuronale Netze nicht starr sind, sondern sich im Laufe der Zeit verändern, kann sich die Prognosequalität u. U. verschlechtern. Daher ist eine laufende Kontrolle und Wartung unumgänglich.

II. Wechselkursprognosen

1. Problemstellung

Die Vorhersage von Wechselkursen ist ebenso schwierig wie notwendig. Besonders in einer Zeit, in der Globalisierung und Faktormobilität aufgrund der weltweiten Vernetzung unerwartete Ausmaße erlangen, ist für einen erfolgreichen Marktauftritt ein ausgeklügeltes Währungsmanagement mit zuverlässigen Instrumenten zur Wechselkursprognose unumgänglich. Dies gilt nicht nur für reine Investoren, welche aufgrund der unterschiedlichen Währungsrenditen durch Devisenhandel Gewinne realisieren wollen, sondern auch für produzierende Unternehmen, von denen viele sich heute schon als Global Player bezeichnen. Die Bedeutung der Wechselkurse für den internationalen Handel bedarf sicherlich keiner Begründung. Dies gilt insbesondere für exportstarke Länder, wie die Bundesrepublik Deutschland.

¹⁰⁹ Vgl. Uhlig (München 1995), S. 48-49.

¹¹⁰ Vgl. Wittkemper (Köln 1994), S. 52-54.

Jedoch ist die Prognose von Wechselkursen äußerst schwierig und kompliziert, so dass auf diesem Gebiet die Theorie effizienter Märkte plausibel erscheint. Das Prognoseproblem ist charakterisiert durch:¹¹¹

- hohe Anzahl von Einflussfaktoren
- komplexe Interdependenzen innerhalb der Einflussfaktoren
- viele Marktteilnehmer mit unterschiedlichen Verhaltensmustern und
- hoch-frequentierte Strukturbrüche und Störungen

Die vorhandenen Wechselkursatheorien sind äußerst kompliziert, können aber nicht mal annähernd den gesamten Zusammenhang modellieren, sondern berücksichtigen immer nur Teilaspekte. Daher eignen sich diese nur bedingt zur Wechselkursprognose. Da neuronale Netze prinzipiell keine theoretischen Kenntnisse über Wirkungszusammenhänge benötigen, stellen diese ein vielversprechendes Prognosewerkzeug dar. Um eine Übersicht über Einflussfaktoren und deren groben Zusammenhänge zu erhalten, sollten die wesentlichen Wechselkursatheorien dennoch betrachtet werden. Diese können in Partialmodelle und Strukturmodelle aufgeteilt werden. Bei Partialmodellen wird immer nur ein Einflussfaktor berücksichtigt. Zu den ältesten unter diesen zählt die Kaufkraftparitätentheorie (KKP). Hier wird der Einfluss der Güterpreise zwischen In- und Ausland als wesentliche Bestimmungsgröße für den Wechselkurs angesehen. Arbitragegewinne aufgrund von Preisdifferenzen werden danach durch Anpassungen der Wechselkurse hin zu einem Gleichgewicht verhindert.¹¹² Ein arbitragefreies Gleichgewicht unterstellt auch die auf Keynes zurückgehende Theorie der Zinsparitäten (ZPT). Wie der Name schon sagt, steht hier nicht der Gütermarkt, sondern der Kapitalmarkt mit dem Zinssatz als Preisgröße im Vordergrund.¹¹³

Strukturmodelle versuchen mehrere Einflussfaktoren und deren Interdependenzen darzustellen. Diese können wiederum untergliedert werden in Bestandsgrößenansätze und Stromgrößenansätze. Während bei den Bestandsgrößenansätzen mit den Größen der jeweiligen Einflussfaktoren (z.B. Höhe des Einkommens, Höhe des Vermögens usw.) kalkuliert wird, kommen bei den Stromgrößenansätzen die Veränderungen der Einflussfaktoren (z.B. Einkommensänderung, Vermögensänderung usw.) zum Einsatz. Auf eine detaillierte Darstellung soll aber verzichtet werden, da diese Theorien nur zur Aneignung des Grundwissens über Einflussfaktoren bei Wechselkursen und deren Zusammenhänge dienen. Als Fundament für die

¹¹¹ Vgl. Poddig/Wallem (München 1994), S. 293.

¹¹² Vgl. Shams (München 1985), S. 31-32.

¹¹³ Vgl. Grimm (Wiesbaden 1997), S. 73-80.

Erstellung von Wechselkursprognosen eignen sich diese weniger.¹¹⁴ Auch aus diesem Grund hielten Wissenschaftler Wechselkursprognosen auf Basis fundamentaler Wechselkursatheorien lange Zeit für nicht sinnvoll, da sie angeblich Zufallsprozessen folgen würden.¹¹⁵ Die Hypothese effizienter Märkte hatte lange Zeit ihre Hochburg auf dem Gebiet der Wechselkursproblematik.

Bezieht man die Theorie effizienter Märkte auf die Renditeberechnung beim Handel mit Devisen, so wären theoretisch keine systematischen Extragewinne erzielbar. Andernfalls wäre die Hypothese effizienter Märkte zu verwerfen. Folglich dienen auf solchen Problemgebieten die naive Prognose und das Buy & Holding-Konzept als Referenzmodelle zur Beurteilung der Prognosegüte. Beim Buy & Holding-Konzept wird in eine Anlage (in diesem Fall ist es eine Währung) investiert. Diese wird dann bis zum Ende einer Periode gehalten. Ziel dieser Strategie ist es, aus dem Aufwärtstrend eines gesamten Marktes zu profitieren.¹¹⁶ Daher wird diese Strategie oft auf Indizes oder eine weit gestreute Anzahl von markt-repräsentativen Anlagentiteln angewendet. Durch dieses Verfahren lässt sich die Performance von Prognoseinstrumenten selbst in depressiven Zeiten gut vergleichen und bewerten. So ist z.B. eine Anlagerendite von -20% natürlich unzufriedenstellend. Beim direkten Vergleich mit einer Buy & Holding-Strategie (auf den Marktindex), welcher im gleichen Zeitraum zu einer Anlagerendite von -40% geführt hat, tritt die Performance des Prognoseinstrumentes zum Vorschein.

Die naive Prognose ist ebenso primitiv, wie zuverlässig und stellt darüber hinaus ein sehr einfach realisierbares Prognoseinstrument dar. Es werden zwei Varianten unterschieden. Bei der einfachen (no change) Variante wird als Prognosewert der zuletzt beobachtete Wert genommen. Die modifizierte (same change) Variante stellt ebenfalls auf den zuletzt beobachteten Wert ab. Zusätzlich wird aber eine Trendkomponente mitkalkuliert. Hierfür wird die zuletzt beobachtete Veränderung fortgeschrieben.¹¹⁷

Kann durch ein Tradingsystem, basierend auf einem Prognoseinstrument, systematisch mehr Rendite erwirtschaftet werden als auf der Basis dieser beiden Prognoseverfahren, wird ein systematischer Gewinnertrag möglich und die Hypothese effizienter Märkte widerlegt. Dass neuronale Prognoseinstrumente diese Hypothese widerlegen können, zeigen die nachfolgenden, empirischen Studien.

¹¹⁴ Vgl. Poddig/Wallem (München 1994), S. 295-299.

¹¹⁵ Vgl. Meese, Rogoff und MacDonald, Taylor.

¹¹⁶ Vgl. Steiner/Bruns (Stuttgart 1994), S. 252.

¹¹⁷ Vgl. Hüttner (Stuttgart 1982), S. 75-76.

2. Modell von Poddig/Wallem¹¹⁸.

Ziel der Studie von *Poddig und Wallem* ist es, die zukünftige Ein-Tages-Differenz des Wechselkurses auf einen Tag zu prognostizieren. Wie bei den meisten Studien zur Wechselkursprognose, beschäftigen sich auch diese Autoren mit dem Verhältnis der heimischen Währung zum US-Dollar. Der Grund dafür ist, dass der US-Dollar international als Referenzwährung dient und ein großer Teil des internationalen Handels auf Dollarbasis abläuft. Im Gegensatz zu den meisten Modellen soll hier aber nicht ein 2-Währungsmodell realisiert werden, wodurch die Realität stark vereinfacht werden würde. Denn dadurch, dass unterschiedliche Währungen gleichzeitig gegeneinander gehandelt werden, hängen Wechselkurse auch vom Einfluss anderer Währungen ab. Es liegt also kein bilateraler, sondern multilateraler Zusammenhang vor.¹¹⁹ Dieser Tatbestand sollte in dem Ansatz mitberücksichtigt werden. Als vorläufige Einflussfaktoren auf die zu prognostizierende Größe wurden 24 tagesbasierte Rohvariablen aus der fundamentalen und der technischen Analyse genommen, welche sich neben USA und BRD auf weitere Länder (England, Frankreich, Japan und Schweiz) beziehen. Zu diesen Einflussfaktoren gehören die relevanten Wechselkurse, Geldmarktsätze, die jeweils wichtigsten Aktienindizes der berücksichtigten Länder und Währungsfutures aus dem Chicagoer International Money Market. Die für das neuronale Netz aus diesen Größen vortransformierten Eingangszeitreihen können in fünf verschiedene Cluster zusammengefasst werden (s. Abbildung 19).

¹¹⁸ Vgl. Poddig/Wallem (München 1994), S. 323-333.

¹¹⁹ Vgl. Poddig/Wallem (München 1994), S. 294.

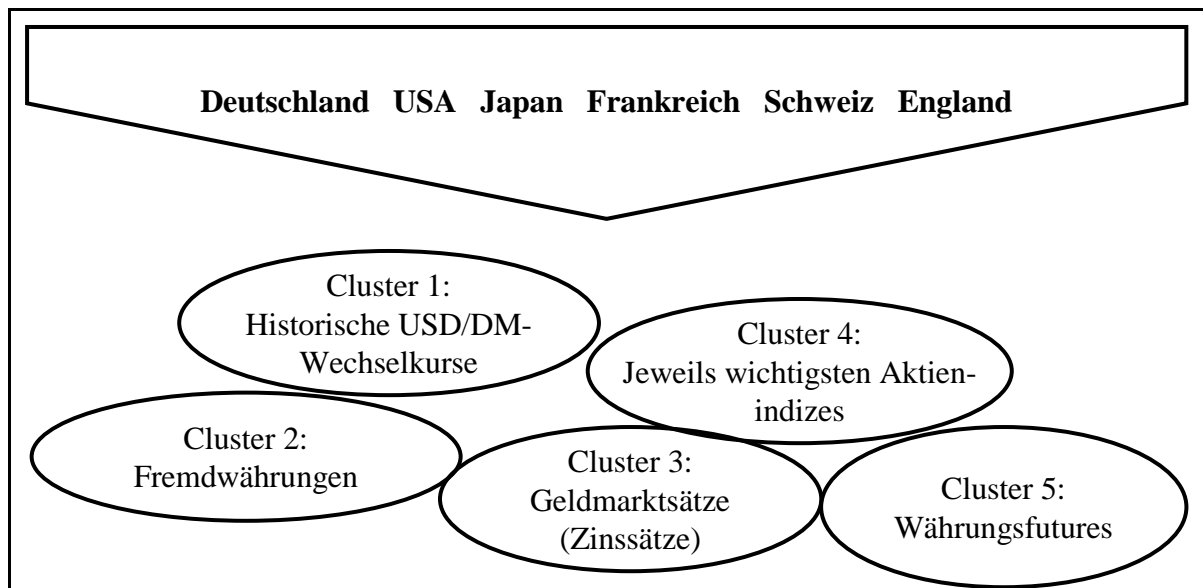


Abbildung 19: Rubriken der Einflussvariablen

Die so ausgewählten Datenreihen umfassen den Zeitraum von 1.1.1985 bis 12.1.1994. Nach einer Korrektur und Vervollständigung der Rohzeitreihen (z.B. Feiertage), ergaben sich insgesamt 2357 Datensätze. Jeder dieser Datensätze entspricht einem erfassten historischen Wechselkurs, welcher sich aus den vorher festgelegten Einflussfaktoren herausgeleitet haben soll. 2152 Datensätze aus dem Zeitraum vom 1.1.1985 bis 31.3.1993 wurden zum Training und die restlichen 205 Datensätze ab dem 1.4.1993 als Testdatenmenge eingesetzt. Aus der Trainingsdatenmenge wurde wiederum jedes vierte Muster chronologisch zur Kreuz-validierung herausgenommen. Diese 538 Datensätze wurden vom Training ausgeschlossen. Nach einer mathematischen und logischen Transformation der Rohdaten erfolgte eine Skalierung.

Nach der Bearbeitung der Zeitreihen wurden diese einer statistischen Filterung unterzogen. Die 24 Rohvariablen wurden zu Beginn durch 203 Zeitreihen dargestellt. Innerhalb von Zeitreihen existieren aber Korrelationen. Diese aufzudecken und entsprechende Eliminationen vorzunehmen, ist Zweck der Korrelationsanalyse. Anschließend blieben 177 Eingangszeitreihen übrig.

Entsprechend der Anzahl von Eingangszeitreihen enthielt das ausgewählte Netzmodell eine Inputschicht mit 177 Neuronen. Hierbei handelte es sich um ein vorwärtsgerichtetes Multilayer-Perceptron mit einer Eingabeschicht, einer verdeckten Schicht, bestehend aus 10 Neuronen und einer Ausgabeschicht mit nur einem Ausgabeneuron. Als Aktivierungsfunktion wurde für

die Inputschicht die lineare Funktion und für die verdeckte Schicht die Tanh-Funktion¹²⁰ verwendet.

Für das Training kam der Backpropagation-Algorithmus zum Einsatz, wobei ein Modell mit dem gängigen Mean-Square-Error¹²¹ als Zielfunktion trainiert wurde und ein Vergleichsmodell mit der von Siemens-Nixdorf entwickelten Profmax-Funktion. Da aber die Auswahl der Zielfunktion das Endergebnis nur geringfügig beeinflusst hat, soll hier von der Differenzierung abgesehen werden. Zur Realisierung des Netzmodells wurde der von Siemens-Nixdorf zur Verfügung gestellte Simulator SENN eingesetzt. Zum Vergleich der Ergebnisse wurde die sog. Wegstreckenberechnung¹²² und die Rendite-berechnung verwendet. Da die Wegstreckenberechnung wenig Aussagekraft besitzt, soll im folgenden die Performance des entwickelten Netzes anhand des hypothetisch getätigten Kaufs und Verkaufs und die daraus resultierende Rendite dargestellt werden. Hierbei wurden von *Poddig und Wallem* alternativ auch die Transaktionskosten, welche bei jedem Positions-wechsel 5 bzw. 10 Basispunkte betragen, mitberücksichtigt (s. Abbildung 20). Der tägliche Einsatz betrug 1000 DM, wobei Gewinne und Verluste nicht miteinbezogen wurden. Die Berechnung der Rendite erfolgte für den (Test-) Zeitraum vom 1.4.1993 bis 19.4.1994 aus den Kauf- und Verkaufssignalen des neuronalen Netzes. Es ergaben sich folgende Renditen:

Transaktionskosten	KNN- Prognose ¹²³	Naive Prognose	Buy & Hold
0 Punkte	49,16 %	-2,43 %	5,91 %
5 Punkte	40,95 %	-10,55 %	5,85 %
10 Punkte	32,74 %	-18,75 %	5,78 %

Abbildung 20: Renditeberechnung unterschiedlicher Handelsstrategien

Selbst bei der Berücksichtigung von Transaktionskosten gelang es durch die Prognosen des neuronalen Netzes, die Renditen der naiven Prognose und der Buy & Hold-Strategie systematisch zu übertreffen. Daraus ist zu folgern, dass es sich beim Devisenmarkt nicht um einen effizienten Markt handelt. Hier könnte das Netz gewinnbringend eingesetzt werden. Obwohl die Autoren mit dem Hinweis auf den langen Test-Zeitraum die Zufälligkeit der Renditen verneinen, stellten sich bei diesem Vergleich der Renditen eine Reihe von Kritikpunkten auf. Diese können aber bei den meisten Arbeiten mit neuronalen Netzen zur Finanzprognose beobachtet werden und sollen daher am Ende des Abschnitts aufgegriffen werden.

¹²⁰ Vgl. Zell (Bonn 1994), S. 92.

¹²¹ Vgl. Poddig (Uhlenbruch 1999), S. 458-459.

¹²² Vgl. Baun (München 1994), S. 186.

¹²³ Es handelt sich hier um das Netzmodell, welches mit der MSN-Zielfunktion trainiert wurde.

3. Sonstige Studien zur Wechselkursprognose

Neben diesem Modell von *Poddig und Wallem* existieren noch viele weitere Studien zur neuronalen Wechselkursprognose. Diese mehr oder weniger erfolgreichen Ansätze befassen sich hauptsächlich mit dem US-Dollar. Es ist aber kaum möglich, diese untereinander zu vergleichen, da große Differenzen zwischen dem Prognosehorizont und den Vortransformationen und Vorselektionen der Einflussfaktoren existieren. Auch die Auswahl der Einflussfaktoren und der Benchmarks, mit denen die Ergebnisse des Netzes verglichen wurden, variieren stark. Es sollen daher im folgenden einige dieser Ansätze skizziert werden, wobei für interessierte Leser eine detaillierte Beschäftigung mit der jeweils angegebenen Primärliteratur angebracht erscheint.

Das in Kap. 2.2.2 vorgestellte Modell ist eine Weiterentwicklung der von *Rehkugler und Poddig* im Jahre 1990 erstmals veröffentlichten Studie zur Prognose des USD/DM-Wechselkurses bei einem Prognosehorizont von drei Monaten.¹²⁴ Hierbei wurden nur fundamentale Faktoren berücksichtigt. Als Netzwerkmodelle kamen das Single-Layer-Perceptron und das Multi-Layer-Perceptron zum Einsatz, wobei das erstere Netzwerkmodell die höhere Performance aufwies. Als Referenzmodell wurde die modifizierte naive Prognose mit einem einfachen Trading-System herangezogen. Während die naive Steigt/Fällt-Prognose zu einer Trefferquote von 58,3% bei einem ökonomischen Gewinn von -28,99% führte, erreichte das Single-Layer-Perceptron eine Trefferquote von 70%, woraus sich ein Gewinn von 46,96% ableiten ließ.

Auf dieser Grundlage erarbeiteten *Rehkugler und Poddig* eine weitere Studie, welche es zum Ziel hatte, Steigt/Fällt-Prognosen des USD/DM-Wechselkurses mit einem Prognosehorizont von 1, 3 und 5 Tagen zu treffen.¹²⁵ Als Einflussgrößen wurden verschiedene technische und fundamentale Indikatoren verwendet. Hier entschieden sich die Autoren für das Multi-Layer-Perceptron. Als Benchmark diente eine multivariate lineare Diskriminanzanalyse mit einer vorgeschalteten Faktorenanalyse (näheres dazu ist der Originalquelle zu entnehmen). Obwohl die Ergebnisse nicht sehr zufriedenstellend waren, konnte das Referenzmodell dennoch recht deutlich geschlagen werden. Es ließen sich Trefferquoten zwischen 61% und 63% realisieren. Einen ganz anderen Ansatz zur USD/DM-Wechselkursprognose liefern *Refenes und Zaidi*. Diese nehmen an, dass für jede Phase der Kursentwicklung optimale Prognoseinstrumente wie z.B. gleitende Durchschnitte oder Mittelwerte existieren. Ziel ihrer Strategie ist es, auf-

¹²⁴ Vgl. *Rehkugler/Poddig* (Bamberg 1990).

¹²⁵ Vgl. *Rehkugler/Poddig* (Heidelberg 1994), S. 1-24.

grund historischer Kursentwicklungen für ein gegebenes bzw. prognostiziertes Marktszenario ein ähnliches aus der Vergangenheit zu finden und dessen optimale Prognosetechnik auf das aktuelle Szenario anzuwenden.¹²⁶ Es handelt sich hier also um eine rein technisch fundierte Prognose. Als Eingaben dienen die jeweiligen Prognosen der Strategien, für die auch die gleitenden Durchschnitte bzw. Mittelwerte eingegeben werden. Die Moving-Average-Strategy (Gleitender Durchschnitt) zeigt besonders in trendbehafteten Phasen eine hohe Performance, während mit der Mean-Value-Strategy (Mittelwertstrategie) bei sich seitwärtsbewegenden und oszillierenden Märkten hohe Gewinne realisiert werden können. Weitere Inputvariable sind die mit den Strategien in den letzten zwei Tagen (hypothetisch) realisierten Gewinne und Verluste. Die letzte Inputvariable ist die beobachtete Veränderung des Wechselkurses über n Tage. Über den Output gibt das Netz dann an, welche Strategie (Moving-Average-Strategy, Mean-Value-Strategy oder die Stillhalterposition) zu nutzen ist. Als Netzwerkmodell fand ein Multi-Layer-Perceptron Verwendung mit:

- einer Inputschicht mit 9 Inputvariablen,
- zwei verdeckten Schichten mit 12 und 6 Neuronen und
- einer Outputsicht mit 3 Neuronen.

Basierend auf täglichen Trainingsdaten vom Zeitraum 1986 bis 1992 wurde im Testzeitraum (1986-1992) bei einem Einsatz von 1 Mio. Dollar folgendes Ergebnis erzielt:

	gesamter Ertrag (kumuliert in Mio. Dollar)	jährlicher Ertrag (prozentual)
MLP	1,45	18
Moving-Average-Strategy	0,98	12,3
Mean-Value-Strategy	1,19	13,1

Abbildung 21: Vergleich der Ergebnisse bei Refenes und Zaidi

Auch hier konnte die Performance zweier klassischer Prognosetechniken von einem neuronalen Netz deutlich überboten werden.

Es gibt noch eine ganze Reihe von Ansätzen zur Wechselkursprognose. Angesichts der recht turbulenten Entwicklung der Gemeinschaftswährung Euro, wird die Anwendung neuronaler Prognoseinstrumente auf die Euro-USD-Relation in Zukunft an Bedeutung gewinnen. Leider existieren noch keine (veröffentlichten) Studien dazu.

¹²⁶ Vgl. Refenes/Zaidi (England 1995), S. 213-220.

III. Aktienkurs- und Indexprognose

1. Problemstellung

Aktienkurse und Aktienrenditen sind bereits seit Beginn des letzten Jahrhunderts, speziell seit den 30er Jahren Gegenstand wissenschaftlicher Untersuchungen. Jedoch erfolgte erst Anfang der 60er Jahre eine intensivere Auseinandersetzung mit dieser Problematik. Auch hier drohte die Theorie effizienter Märkte die Forschungsbemühungen für die Entwicklung von Prognoseinstrumenten. Auf die Aktienkurse bezogen unterstellt diese Theorie, dass sämtliche verfügbaren Informationen in den Aktienkursen schon enthalten sind und folglich keine systematischen Vorteile gegenüber anderen Anlegern erreicht werden können. Dies entspricht aber, wie in den vorherigen Kapiteln schon dargestellt wurde, nicht der Realität.

Unabhängig davon versuchen Aktieninvestoren seit jeher, die Entwicklung von Aktienkursen mit dem Ziel zu prognostizieren, überdurchschnittliche Renditen zu erzielen. Hierbei werden Entscheidungen von allen Beteiligten unter Unsicherheit der zukünftigen Kursentwicklungen getroffen. Die Entwicklung von Wertpapierkursen ist das Produkt hochkomplizierter Interdependenzen. Die Komplexität wird erhöht durch einen hohen Grad an Rekursivität.¹²⁷ Vereinfacht sagt man dazu auch: „Der Aktienmarkt führt ein Eigenleben“.

Das Problem ist wie bei der Wechselkursbestimmung charakterisiert durch eine hohe Anzahl von Einflussfaktoren, zu denen natürlich auch die Erwartungen der Individuen gehören. Des Weiteren können sich die Einflussfaktoren gegenseitig beeinflussen, verstärken oder kompensieren. So kann eine Phase der Unsicherheit und des Pessimismus zu noch mehr Unsicherheit und Pessimismus führen, was unverständlicherweise nicht selten in Kurs-zusammenbrüchen endet. Dies führt so weit, dass eine große Anzahl von Anlegern im Zusammenhang mit der Börse schon von „Roulette“ sprechen. Als Beispiel sei auf die zur Zeit sehr prekäre Situation auf dem „Neuen Markt“ hingewiesen. Während in den letzten Jahren unerwartete und irrationale Kursexplosionen zu verbuchen waren, führt heute bei einem Datenbankspezialisten eine Gewinnsteigerung von „nur“ 40% anstelle von den erwarteten 50% zu einem Eintages-Kursverlust von 4%! Hier wird noch einmal die Relevanz der Erwartungswerte für die Kursbestimmung in deutlicher Form dargestellt. Wie kann aber deren Einfluss, mal abgesehen von deren Größe, quantifiziert werden? Ist das überhaupt möglich? So ist die Aktienkursprognose ein sehr heikles und risikobehaftetes Geschäft.

¹²⁷ Ganz geschweige von den nicht-linearen Zusammenhängen.

Dennoch ist das Interesse an der Anlageform Aktie speziell seit den Zinssenkungen 1993¹²⁸ stark gestiegen. Dagegen sind festverzinsliche Wertpapiere für viele Anleger auf das „Hinter-treppchen“ geraten. Dies gilt insbesondere für den privaten Anleger.

Aus den selben Gründen wie bei der Wechselkursproblematik, eignen sich auch hier künstliche neuronale Netze gut zur Prognose. Aufgrund gewachsener Nachfrage nach Börsengeschäften durch Privatanleger, werden in letzter Zeit verstärkt fertige und einfach handhabbare Software-Simulatoren mit (eingeschränkten) Modifikationsmöglichkeiten angeboten. Es wäre sicherlich sehr interessant zu wissen, wie solch ein Tool im Vergleich mit konventionellen Prognosemethoden oder im direkten Vergleich mit einem professionellen Börsenmakler abschneidet. Leider kann die Performance solcher Software-Simulatoren zur Zeit nur aus Werbebroschüren herausgelesen werden und ist daher für die Wissenschaft weniger nützlich.

Aus diesem Grund werden diese ausgelassen und statt dessen einige wissenschaftliche Studien zur Aktienkursprognose mit neuronalen Netzen skizziert werden.

2. Modell von Gehring et al.¹²⁹

Die meisten Studien zur Aktienkursprognose beschäftigen sich nur mit einzelnen Aktien oder mit der Prognose von Indizes und zielen somit nur auf Teilaspekte ab. Wie auch bei dem folgenden Modell basieren diese oftmals auf der technischen Analyse. Des weiteren ist die uneinheitliche und nicht selten wenig aussagekräftige Präsentation der Ergebnisse zu bemängeln. Der Grund für die Auswahl des Modells von *Gehring et al.* ist die Prognose von mehreren Aktien, welchen die Prognose des dazugehörigen Index vorgeschaltet ist. Darüber hinaus stellt diese Studie einen vollständigen und sehr umfangreichen Ansatz zur Aktienkursprognose dar und eignet sich daher auch für das Portfoliomanagement.

Gegenstand der Untersuchung ist der Deutsche Aktienindex DAX und einige der darin gewichteten Aktien. Wie es in der Praxis auch üblich ist, beschränken sich die Autoren auf eine bestimmte Menge von Aktien, dem Evoked Set. In diesem Fall sind es fünf DAX-Werte aus den Branchen Banken, Automobilhersteller, Chemie und Maschinenbau, zuzüglich der Aktie einer diversifizierten Holding.¹³⁰ Es sind also täglich Entscheidungen darüber zu treffen, welche und wie viele Wertpapiere aus dem Evoked Set in das Portfolio aufzunehmen sind und wie viele von welchen im Portfolio enthaltenen zu verkaufen sind. Als primäres Motiv für

¹²⁸ Möglicherweise ist auch dies der Grund für eine ganze Reihe von Veröffentlichungen zur Aktienkurs- und Indexprognose im Jahre 1994.

¹²⁹ Vgl. Gehring (Hagen 1999).

¹³⁰ Vgl. Gehring (Hagen 1999), S. 19.

diese Aktivität wird nicht die Renditesteigerung ausgewählt. Hier wird eine Relativierung der Rendite durch die Hinzunahme eines Anlagerisikos vorgenommen. Eine Gegenüberstellung von Renditepotenzial und Risiko wird durch die Risikonutzen-Funktion beschrieben.¹³¹ In dieser Risikonutzen-Funktion ist auch ein Präferenzparameter enthalten, wodurch die Risikobereitschaft mitberücksichtigt wird. Durch diesen Risikopräferenzparameter kann die Risikonutzen-Funktion entsprechend der Ausrichtung der Anleger optimiert werden.¹³²

Demgemäss ist primäres Ziel, den Risikonutzen entsprechend der Präferenzen des Anlegers für das Depot zu maximieren. Hierzu wird ein Entscheidungsunterstützungssystem entwickelt, das den Prozess von der DAX-Prognose bis zur Depotanpassung unterstützt. Dieser Prozess kann in vier Phasen gegliedert werden:¹³³

- (1) Prognose der allgemeinen Börsenentwicklung durch DAX-Prognose
- (2) Kursprognose für die Aktien aus dem Evoked Set mit Berücksichtigung der DAX-Prognose
- (3) Bestimmung des optimalen Portfolios durch Berechnung der Anteile der Wertpapierarten und der Barreserve
- (4) Depotanpassung durch Berechnung der zu kaufenden bzw. zu verkaufenden Wertpapierarten aus dem Evoked Set

Zur Vereinfachung werden nur die Gewinne und Verluste berücksichtigt, welche durch Kursänderungen entstehen. Dividenden, Transaktionskosten, Steuern und Zinseinnahmen durch die Barreserve werden aus der Berechnung ausgeschlossen.

Den Kern dieser Arbeit stellt die Prognose von DAX und Aktienkursen aus dem Evoked Set dar. Hierfür wurde ein zweistufiges, neuronales Netz herangezogen (s. Abbildung 22). Bei den beiden miteinander gekoppelten Netzen handelt es sich um vorwärtsgerichtete Multi-Layer-Perceptrons mit 3 Schichten und je einem Ausgabeneuron.

¹³¹ Die Herleitung der Risikonutzen-Funktion ist den Seiten 11 und 12 zu entnehmen.

¹³² Vgl. Gehring (Hagen 1999), S. 11-13. Hier wird in drei Klassen unterteilt: Sicherheitsstreben, mittlere Risikopräferenz und Renditestreben.

¹³³ Vgl. Gehring (Hagen 1999), S. 2-3.

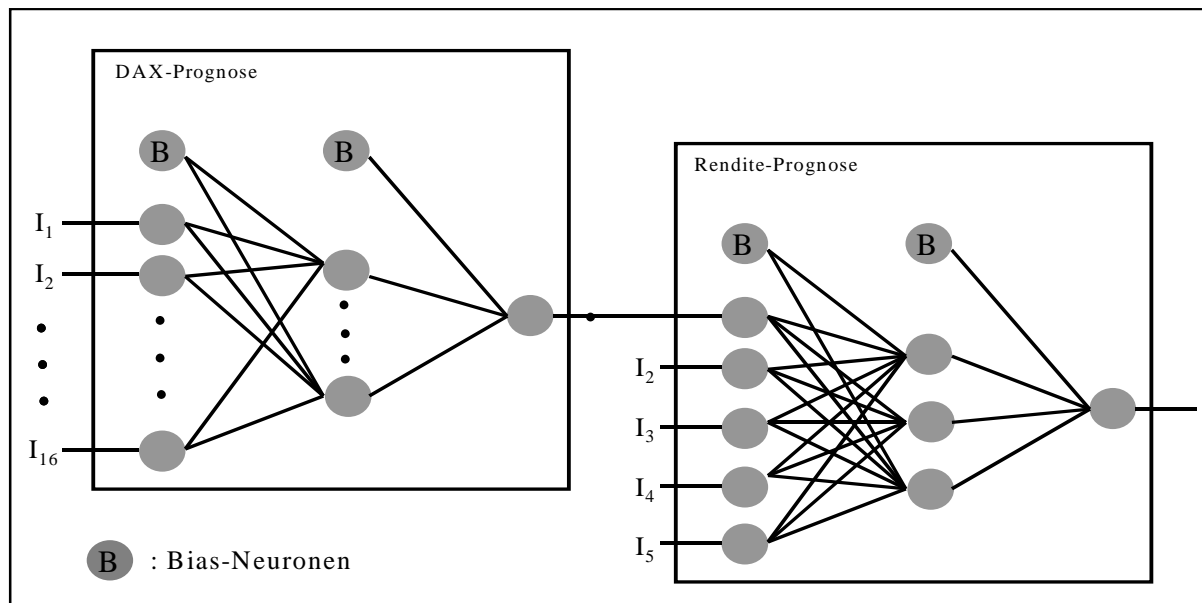


Abbildung 22: Zweistufiges neuronales Netz zur Aktienkursprognose ¹³⁴

Das Netz zur DAX-Prognose enthält acht verdeckte Neuronen, während das Netz für die Renditeprognose der Aktien 3 verdeckte Neuronen enthält. Zusätzlich enthalten beide Netze je zwei Bias-Neuronen in der verdeckten Schicht und in der Ausgabeschicht. Bias-Neuronen werden zur externen Modifikation der Schwellenwerte herangezogen. Die Anzahl der Inputneuronen beträgt entsprechend den für die DAX-Prognose relevant gehaltenen Indikatoren 16. Diese sind ausschließlich technische Indikatoren. Das neuronale Netz zur Renditeprognose enthält insgesamt fünf Inputneuronen, wobei das Prognoseergebnis für den DAX einen Input darstellt. Die markttechnischen Indikatoren werden in der Literatur explizit dargestellt. ¹³⁵

Das besondere an dieser Studie ist, dass für das Training der neuronalen Netze ein genetischer Algorithmus eingesetzt wurde. Genetische Algorithmen lösen komplexe Optimierungsaufgaben, welche, wie im vorliegenden Fall, mit konventionellen Methoden kaum zu bewältigen wären, nach dem Vorbild der biologischen Evolution und der molekularen Genetik. ¹³⁶ Auf eine detaillierte Beschreibung der Anwendung auf die dargestellten neuronalen Netze wird verzichtet. ¹³⁷

Für die Trainings- und die Testphasen wurden Daten bezüglich DAX (Eröffnungs-, Hoch-, Tief- und Schlusskurse) und Aktien (Schluss- und Kassakurse) von insgesamt 1856 Börsentagen eingesetzt. Diese Daten wurden in acht Phasen zu je 232 Börsentagen unterteilt, wobei

¹³⁴ Vgl. Gehring (Hagen 1999), S. 5.

¹³⁵ Vgl. Gehring (Hagen 1999), S. 7-8.

¹³⁶ Eine ausführliche Darstellung findet sich in Schöneburg (Bonn 1994).

¹³⁷ Auf S. 13-17 wird das Training mit dem genetischen Algorithmus beschrieben.

unabhängig von der zeitlichen Folge je 2 Phasen zum Training und 6 Phasen für den Test genutzt wurden. Insgesamt ergaben sich durch die Phasenverschiebungen und durch die unterschiedliche Risikoneigung etc. 108 neuronale Entscheidungsunterstützungssysteme. Als Bewertungsmaßstab diente, wie bereits erwähnt, der Risikonutzen. Als Referenzstrategie verwendeten die Autoren auch hier die schon betrachtete Buy & Holding-Strategie. Dabei wurde nur in Aktien aus dem Evoked Set investiert. Die Autoren begnügten sich folglich mit einer Gegenüberstellung der erreichten Risikonutzen zur Demonstration der Performance von neuronalen Netzen.

Als Ergebnis wurden die prozentualen Differenzen der Risikonutzen für jedes optimierte Netz in einer Tabelle aufgelistet. Hierzu wurden jeweils die Mittelwerte der Beträge für die Trainings- und die Testphasen der 108 unterschiedlichen Anlagesysteme eingesetzt. Anstelle der detaillierten Darstellung der Ergebnisse für die 108 Anlagesysteme¹³⁸ soll im folgenden für jede Risikoneigung die durchschnittlich erhaltene Risikonutzen-Differenz in der Testphase interpretiert werden (s.

Abbildung 23).

Risikoneigung	Ø Risikonutzen-Differenz	Ø Anzahl gekaufter Aktien	ØKapitaleinsatz
Sicherheitsorientiert	2,8%	2,3	29,6%
Mittlere Risikopräferenz	5,4 %	1,8	58,5%
Renditestreben	15,5%	2,7	70,6%

Abbildung 23: Ergebnisse der neuronalen Anlagesysteme

Unabhängig von der Risikoneigung und den unterschiedlichen Testphasen kann eine signifikante Überlegenheit der neuronalen Anlagesysteme beobachtet werden. Mit steigender Risikopräferenz nimmt auch die durchschnittliche Risikonutzen-Differenz (der nach Risikoklasse unterteilten 108 Anlagesysteme) im Verhältnis zur Buy & Hold-Strategie von 2,8% bis 15,5% zu. 92% der Anlagesysteme konnten über alle Testphasen gemittelt besser abschneiden als die Referenzstrategie. 42% konnten sogar über alle Testphasen die Referenzstrategie schlagen. Mit steigender Risikoneigung ist verständlicherweise der Kapitaleinsatz gestiegen (s.

Abbildung 23). Verwunderlich ist jedoch, dass keine signifikante Korrelation zwischen Diversifikation (d.h. Anzahl der gekauften Aktien aus dem Evoked Set) und dem Ergebnis festzustellen ist. Nur selten wurde durch das System in mehr als drei Wertpapiere investiert. Ebenso

¹³⁸ Vgl. Gehring (Hagen 1999), S. 23-24.

wenig lässt sich eine signifikante Korrelation zwischen Risikopräferenz und Diversifikation feststellen, wobei es erstaunlich ist, dass die durchschnittlich höchste Diversifikation nicht beim sicherheitsorientierten Anlagesystem, sondern bei der höchsten Risikoneigung beobachtet werden konnte.

Insgesamt konnte auch hier die Performance von neuronalen Netzen eindrucksvoll demonstriert werden. Besonders lobenswert ist die Vollständigkeit und die Geschlossenheit des entwickelten Anlagesystems, was u.a. auch die vielfältigen und weitreichenden Nutzungsmöglichkeiten neuronaler Netze verdeutlicht. Nachzuholen wäre hier die Erweiterung der Funktionalität zur Berücksichtigung von Steuern, Dividenden, Transaktionskosten und der Zinseinnahmen aus der Barreserve, was sicherlich keine größeren Probleme darstellen dürfte, aber für den kommerziellen Einsatz Voraussetzung ist.

3. Sonstige Studien

Fast ein Jahrzehnt früher entwickelten *Kimoto et al.*¹³⁹ einen ähnlich umfangreichen und ausgereiften Ansatz. Ihre Absicht war es, die Entwicklung des TOPIX (Tokyo Stock Exchange Price Index) mit einem Prognosehorizont von 1 Monat vorherzusagen. Hierfür wurde ein Handelssystem namens SMPS (Stock Market Prediction System) entwickelt, welches selbstständig Kauf- und Verkaufssignale erzeugte. SMPS besteht aus mehreren neuronalen Netzen, deren Zielsetzung es ist, optimale Kauf- und Verkaufszeitpunkte in Abhängigkeit von technischen Indikatoren zu finden. Das mit einem modifizierten Backpropagation-Algorithmus trainierte Netzwerk-Komplex wurde über einen Zeitraum von 33 Monaten zwischen Januar 1987 und September 1989 getestet. Während in diesem Intervall mit einer Buy & Hold-Strategie eine Rendite von 67% erreicht worden wäre, führten die auf der Basis der Handelssignale des SMPS getroffenen Transaktionen zu einer Rendite von 98%.

Von der SGZ-Bank AG (Südwestdeutsche Genossenschafts-Zentralbank AG) gibt es einige nennenswerte Studien zur Prognose mittels neuronaler Netze. Das sind die Arbeiten von *Baun* (1994) und *Graf* (1991) bzw. *Hillmer und Graf* (1994). *Graf* entwickelte ein System zur Vorhersage von optimalen Kaufs- und Verkaufszeitpunkten für den DAX mit Hilfe von technischen und ökonomischen Daten.¹⁴⁰ Hier gibt sich der Autor mit einem neuronalen Netz zufrieden, welches aus einer Eingabeschicht, zwei verdeckten Schichten und einer Ausgabeschicht besteht. *Füser* vermutet, dass als Trainingsverfahren ein modifiziertes Backpropagati-

¹³⁹ Vgl. Kimoto (San Diego 1990), S. 1-7.

¹⁴⁰ Vgl. Graf (Heidelberg 1991), S. 496-499.

on-Algorithmus verwendet wurde.¹⁴¹ Die Trainingsdatenmenge wurde aus dem Zeitraum zwischen 1.1.1988 und 8.10.1990 gebildet. Die daran anschließenden 10 Monate wurden als Testdatenmenge herangezogen. Die Eingangsinformationen setzen sich aus den Gruppen technische Indikatoren, ausländische Aktienindizes, Wechselkurse, Zinsen und Turnovers zusammen. Das neuronale Netz übertraf mit 17,7% Rendite die Buy & Hold-Strategie, welche zu einer Rendite von 11,26% führte.

In ihrer, im Jahre 1994 veröffentlichten Studie, versuchen *Hillmer und Graf*, mit Hilfe statistischer Verfahren und Fehler-Korrektur-Modellen eine Tendenzprognose für den DAX mit einem Prognosehorizont zwischen 6 und 12 Monaten zu erstellen.¹⁴² Bei dieser, als Weiterentwicklung des vorherigen Ansatzes zu verstehenden Studie, wurden als Einfluss-faktoren nicht nur technische, sondern auch fundamentale Daten herangezogen. Nach einer Selektion der Einflussfaktoren konzentrieren sich die Autoren auf die Inputgrößen Auftragseingang im verarbeitenden Gewerbe, DM/USD-Wechselkurs und Rendite von Bundesanleihen mit 10-jähriger Laufzeit. Das Netz enthielt folglich eine Eingabeschicht mit 3 Neuronen, eine verdeckte Schicht mit 3 Neuronen und einen Ausgabeneuron. Das Multilayer-Perceptron wurde mit dem Backpropagation-Algorithmus trainiert. Die Trainingsmenge bestand aus 24 Quartalswerten. Die Testphase verlief über einen Betrachtungszeitraum von 10 Jahren. Das Ergebnis war eine Trefferquote von 68,4% mit einer Rendite von 503,6%, was im Vergleich mit einer Buy & Hold-Strategie im gleichen Zeitraum zu einem Mehrgewinn von 130,9% geführt hätte.

Baun beschreibt in ihrer Studie, welche ebenfalls im Jahre 1994 veröffentlicht wurde, eine Vorgehensweise zu Erstellung von Prognosen mit Hilfe der von Siemens-Nixdorf hergestellten Simulationssoftware SENN. Auf dieser Grundlage versucht sie, die Ein-Tages-Rendite des DAX zu prognostizieren.¹⁴³ Das Netz besteht aus einer Eingabeschicht, einer verdeckten Schicht und einer Outputschicht mit einem Ausgabeneuron, welches die Renditeprognose für den nächsten Tag ausgibt. Auch hier wurde ein vorwärtsgerichteter Multilayer-Perceptron eingesetzt. Die für die Untersuchung verwendeten Daten sind 34 technische und fundamentale Indikatoren, deren Zeitreihen sich über ca. 10 Jahre erstrecken. Von den nach der Vorverarbeitung erhaltenen Daten wurden die vom Zeitraum zwischen 5.3.1984 und 4.1.1993 zum Training eingesetzt. Die nachfolgenden Daten für 237 Tage dienten dem Test, wobei wiederum auf der Basis hypothetischen Tradings ein Renditevergleich erstellt wurde. Dabei ergab

¹⁴¹ Vgl. Füsler (Wiesbaden 1995), S. 140.

¹⁴² Vgl. Hillmer/Graf (Heidelberg 1994), S. 149-182.

¹⁴³ Vgl. Baun (München 1994), S. 171-193.

sich für das neuronale Netz eine Rendite von 41%, für die naive Prognose 21% und für die Buy & Hold-Strategie eine Rendite von 34%.

Neben diesen Studien existieren viele weitere Ansätze zur Aktienkurs- und Indexprognose mit neuronalen Netzen, von denen die meisten zufriedenstellende Ergebnisse erzielen konnten. So wurden bei allen Studien Trefferquoten von durchschnittlich 60% erreicht, wobei Buy & Hold-Strategien und naive Prognosen fast immer übertroffen werden konnten. Jedoch stellt man bei der näheren Untersuchung dieser Ansätze fest, dass sehr viel auf ex-post-Wissen basiert und viele Ergebnisse verschönert sind, was sich oftmals am Ende der Studien in Form von „ist für den praktischen Einsatz nicht geeignet“ ausdrückt.

IV. Abschließende Betrachtung zur Finanzprognose

1. Sonstige Einsatzmöglichkeiten für Finanzprognosen

Ein weiterer interessanter Einsatzbereich für künstliche neuronale Netze ist die Zinsprognose. Für die oben geschilderte Wechsel- und Aktienkursprognose kann eine erfolgreiche Vorhersage von Zinssätzen als Basis dienen. Zwischen Aktienkursen, Wechselkursen und Zinssätzen existieren komplexe Interdependenzen. Aus diesem Grund stellt die Zinsprognose einen sehr wichtigen Aufgabenbereich in der Finanzwirtschaft dar. Jedoch ist hier zuerst einmal zu klären, was für ein Zinssatz prognostiziert werden soll. Zu den wichtigsten unter diesen gehören u.a. Emissions- und Umlaufrenditen der verschiedenen festverzinslichen Wertpapiere, Diskont- und Lombardsatz und die unterschiedlich gefristeten Geldmarktsätze.¹⁴⁴ Bei den unterschiedlichen Theorien zur Erklärung von Zinsen wird auf eine derartige Differenzierung verzichtet, wobei der „Zins“ entsprechend unterschiedlich definiert sein kann.

¹⁴⁴ Vgl. Poddig (München 1994), S. 214-215.

Studie	Datenmaterial/ Analysemethode	Zielvariable/ Prognosehorizont	Netzwerkmodell	Ergebnis
Matthes, 1994	Monatsdaten techn./fund.	Kapitalmarktzins, 3 Monate	MLP	Trefferquoten von ca. 70%
Rieß, 1994	Monatsdaten überw. fund.	Kapitalmarktzins, 1, 3, 12 Monate	MLP	„enttäuschend“ ¹⁴⁵
Poddig, 1994	Monatsdaten techn./fund.	3-Monats-EURO- Geldmarktsatz, 3-Monate	MLP GRNN ¹⁴⁶	Ranking: 1. MLP 2. GRNN 3. naive Prognose
Baun/Storm, 1992	Monatsdaten techn./fund.	Kapitalmarktzins, 3, 6, 12 Monate	MLP	Trefferquoten: 70% (3 Monate) 82% (6 Monate) 82% (12 Monate)

Abbildung 24: Studien zur Zinsprognose mit neuronalen Netzen¹⁴⁷

Die meisten Studien zur Prognose mit künstlichen neuronalen Netzen beschäftigen sich mit den Emissions- und Umlaufrenditen der verschiedenen festverzinslichen Wertpapiere, wofür auch der Begriff Kapitalmarktzins verwendet wird. Einige dieser Studien sind in

Abbildung 24 zusammengefasst. Diese basieren auf technischen und fundamentalen Analysemethoden, wobei oft Kombinationen eingesetzt werden. Ein Problem der fundamentele Analyse ist die Verfügbarkeit von Daten bzw. die langen Zeitabstände der Informationsgewinnung. Während für technische Analysen die Zeitreihen auf bis zu Sekundenabständen vorliegen, bleiben für die Fundamentalanalyse oft nur Monats- bzw. Quartalsreihen.¹⁴⁸ Somit liegt in der Regel eine sehr „grobkörnige“ Darstellung der realen Vorgänge vor. Unter Berücksichtigung der aktuellen Dynamik der Märkte mit ihren Strukturbrüchen und sonstigen Varianzen wird die Problematik bei der Zinsprognose leicht ersichtlich. Festzustellen ist, dass mit neuronalen Netzen fundamentale Analysemethoden überwiegend für längerfristige Prognosehorizonte eingesetzt werden, während technische Ansätze eher für kurzfristige Prognosen herangezogen werden.¹⁴⁹

Im Gegensatz zur Zinsprognose ist der Einsatz neuronaler Netze für Derivate bisher nur spärlich erforscht. Jedoch stellt der Terminhandel mit seinen derivativen Finanzinstrumenten (aus anderen Produkten abgeleitete Wertobjekte) einen in letzter Zeit stark nachgefragten und zu-

¹⁴⁵ Vgl. Rieß (Aachen), S. 153.

¹⁴⁶ Vgl. Poddig (München 1994), S. 272-276.

¹⁴⁷ Vgl. Füser (Wiesbaden 1995), S. 165-192 und Poddig (Uhlenbruch 1999), S. 377.

¹⁴⁸ Vgl. Poddig (München 1994), S. 211.

¹⁴⁹ Vgl. Füser (Wiesbaden 1995), S. 165.

kunftsträchtigen Geschäftsbereich dar. Eine wesentliche Ursache hierfür sind die gestiegenen Volatilitäten an den Märkten, was zu höheren Risiken für Anleger und Unternehmen führt. Derivative Instrumente, insbesondere Financial Futures, können unabhängig vom aktuellen Marktzustand zur Risikoabsicherung (Hedging) herangezogen werden.¹⁵⁰ Des Weiteren erlauben derivative Instrumente u.U. das Umgehen von staatlichen Regelungen, wie Kapitalverkehrsbeschränkungen und das effektivere Ausschöpfen steuerlicher Gegebenheiten.¹⁵¹ Ein besonderer Anreiz zum Handel mit Derivaten ist das immense Gewinnpotenzial aufgrund der Hebelwirkung, dem natürlich ein ebenso großes Verlustrisiko entgegensteht.

Neuronale Netze können in diesem Kontext zur Kursprognose der zugrundeliegenden Basiswerte (Underlying) verwendet werden, um Gewinnchancen zu erhöhen oder um ein erfolgreiches Risikomanagementsystem zu entwickeln bzw. zu untermauern. Bisherige Studien zum Einsatz neuronaler Netze für Derivate beschäftigen sich überwiegend mit Financial Options und Financial Futures. Häufiges Hauptaugenmerk liegt leider nur in der Approximation schon vorhandener Bewertungsformeln, wie jene, die 1973 von *Black und Scholes*¹⁵² zur Optionsbewertung, entwickelt wurden. Hier liegen aber oftmals wesentliche Vereinfachungen vor, wie z.B. die Nicht-Berücksichtigung von Volatilitäten der zugrundeliegenden Aktien.¹⁵³ So stellt *Hanke* 1998 ein Studie zur Approximation der Black/Scholes-Bewertungsformel mit einem 3-schichtigen neuronalen Netz auf.¹⁵⁴ Als Untersuchungsgegenstand dienen DAX-Optionen. Auch in diesem Ansatz wird auf die hohe Komplexität der Problematik und die unzureichende Performance konventioneller Analysemethoden hingewiesen. Als Ergebnis wird verlautet: „*Neuronale Netze eignen sich zur Approximation von Optionsbewertungsformeln für Optionstypen und Bewertungsmodelle, die einer analytischen Lösung nicht zugänglich sind.*“¹⁵⁵

Der zukünftige Einsatz neuronaler Netze für Derivate muss in engem Zusammenhang mit der Eignung dieses neuen Prognoseinstrumentes für die Vorhersage der jeweiligen Basiswertkurse gesehen werden. Es bestehen sicherlich größere Einsatzpotenziale als die alleinige Funktionsapproximation vorhandener Berechnungsformeln. Wie in den vorherigen Kapiteln aufgezeigt wurde, existieren zahlreiche Studien zur neuronalen Prognose von Wechsel-, Zins- und

¹⁵⁰ Vgl. Schierenbeck/Hölscher (Stuttgart 1998), S. 694-698.

¹⁵¹ Vgl. Schmidt (Stuttgart 1999), S. 2.

¹⁵² Vgl. Schierenbeck/Hölscher (Stuttgart 1998), S. 667-672.

¹⁵³ Vgl. Schierenbeck/Hölscher (Stuttgart 1998), S. 669.

¹⁵⁴ Vgl. Hanke (Frankfurt a. M. 1998), S. 108-138.

¹⁵⁵ Hanke (Wien 1998), S. 139.

Aktienkursen. Es versteht sich von selbst, dass Termingeschäfte auf diese Basiswerte von der Prognosegüte dieser Basiswerte selbst abhängt.

Es ist zu erwarten, dass in Zukunft durch das Zusammenwirken mehrerer neuronaler Netze eine ganze Reihe von Zielen gleichzeitig verfolgt werden wird. Hierbei ist die Homogenität eingesetzter Werkzeuge auch vorteilhaft.

So könnte z.B. ein neuronaler Netzwerkkomplex aufgebaut werden, mit der zuerst Zinssätze und Wechselkurse prognostiziert werden. Mit diesen Informationen und zusätzlichen, relevanten Prognosen, wie z.B. die Vorhersage von Aktienkursvolatilitäten, die zum Teil auf den vorangegangenen aufbauen, können dann Kurse bestimmter Aktien prognostiziert werden. Auch Optionen ließen sich dann problemlos einordnen. Solch umfangreiche Studien könnten als Entscheidungsunterstützungssystem im Rahmen eines Portfoliomanagements eingesetzt werden.

Neuronale Netze bieten in diesem Kontext ein großes Potenzial. Der Einsatz dieses neuen Instrumentes birgt aber auch Risiken. Aus diesem Grund soll im folgenden eine Bewertung neuronaler Netze zur Prognose von Finanzmarktdaten erfolgen.

2. Kritische Würdigung

Neuronale Netze stellen als eine innovative Prognosetechnik ein großes Potenzial für die Finanzwirtschaft dar. Varianzen und multilaterale, nicht-lineare Interdependenzen zwischen den Einflussfaktoren des zugrundeliegenden Prognoseproblems, welche mit konventionellen Vorgehensweisen kaum zu bewältigen sind, werden auch in Zukunft konnektionistischen Systemen einen hohen Stellenwert einräumen.

Ein besonderer Vorteil neuronaler Netze ist in der Modellfreiheit zu sehen. Eine gänzliche Modellfreiheit liegt natürlich nicht vor, da zumindest die relevanten Variablen auf Modellen basieren. Es ist aber keine detaillierte Kenntnis über diese, oftmals sehr komplizierten und dennoch die Realität stark vereinfachenden Modelle nötig. Denn neuronale Netze leiten ihre Modelle aus vorhandenen Daten selbst her. Die Daten bilden den kritischen Bereich bei der Entwicklung neuronale Netze. Die Qualität der Daten und der Datenspeisung bestimmt letztendlich die Performance des neuronalen Netzes. Die Lern- und Adaptionfähigkeit kann durch ein gleitendes Fenster der zum regelmäßigen Training verwendeten Zeitreihen verwendet werden, um die Problematik der in der Finanzwirtschaft öfters auftretenden Varianzen und sonstiger Strukturbrüche zumindest teilweise zu entschärfen. Diese Vorgehensweise wurde in den meisten der vorgestellten Studien eingesetzt. Jedoch wurden von den Autoren unterschiedliche Vorverarbeitungen und sonstige Modifikationen verwendet, so dass ein Vergleich

kaum möglich ist und somit eine optimale Datenapplikation nicht herausselektiert werden kann. Die Datenbeschaffung stellt hierbei ein nicht zu unterschätzendes Problem dar. Obwohl mit zunehmenden informationstechnischen Integrationsprozessen die Datenbeschaffung erleichtert wurde,¹⁵⁶ zeigte es sich als außerordentlich schwierig, regelmäßig aktuelle und homogene Daten zu erhalten, wobei Informationslags Prognosen zusätzlich erschweren. Wie im vorherigen Kapitel aufgezeigt wurde, spitzt sich diese Problematik speziell bei fundamentalen Ansätzen zu, da hier in der Regel mit Monats- bzw. Quartalsreihen gearbeitet werden muss.

Ein weiteres großes Problem ist die Vorgehensweise zur Entwicklung eines neuronalen Netzes zur Prognose. Es gibt keine Theorie, welche für jeweilige Aufgabenstellungen die entsprechende Netzarchitektur und Netzparameter bestimmt. Dieses Problemlösungswissen basiert leider immer noch auf Intuition, Erfahrung und Zufall.

Einen wesentlichen Kritikpunkt stellt die „Black-Box“-Eigenschaft neuronaler Netze dar. Es muss berücksichtigt werden, dass neuronale Netze keine Erklärungskraft besitzen. Die Ursachen für das entsprechende Handelssignal durch das Netz sind für den gewöhnlichen Anwender nicht ersichtlich. Vielleicht werden in Zukunft Verfahren entwickelt, welche eine Transparenz der inneren Abläufe gewährleisten.

Trotz dieser Makel und einem immensen Forschungsbedarf ist es den meisten Autoren gelungen, die Eignung neuronaler Netze für die Prognose von Finanzmarktdaten empirisch nachzuweisen. Die Theorie effizienter Märkte, welche sich überwiegend aus der Unzulänglichkeit konventioneller Prognoseinstrumente herausleitete, konnte durch den Einsatz neuronaler Netze zumindest stark relativiert werden. So zeigten Studien, dass systematische Extragewinne sehr wohl möglich sind. Gängige Referenzmodelle konnten in den meisten Studien signifikant geschlagen werden, wobei nicht selten Trefferquoten von über 70% erreicht wurden.

Problematisch ist der Vergleich dieser Studien, welche meistens im Auftrag von Finanzinstituten erarbeitet wurden. Eine direkte Gegenüberstellung ist kaum möglich, da unterschiedliche Vorverarbeitungen von Variablen und unterschiedliche Ergebnisformulierungen vorliegen. Es ist davon auszugehen, dass viele Wissenschaftler für die Erstellung der neuronalen Netze ex-post-Wissen herangezogen haben. So konnte in einigen Studien beobachtet werden, dass relevante Einflussfaktoren, welche erst in der Testphase spezifiziert wurden und bei der Netzwerkmodellierung nicht vorlagen, im nachhinein hinzugefügt wurden. Manche Autoren gehen sogar noch weiter und benutzen einige Musterdaten aus der Testphase, um ihr Netz zu

¹⁵⁶ Diese Tendenz wird sich in Zukunft sicherlich fortsetzen.

trainieren. Hier ist es natürlich kein Wunder, dass das Netz sehr hohe Trefferquoten erzielt. Bei Trefferquoten um die 80% sollte dieser Verdacht immer bestehen.

Weiterhin ist zu kritisieren, dass bei vielen Studien die Autoren ihre Vorgehensweise, die zu den jeweiligen Ergebnissen führte, nur spärlich dokumentieren. Ein Grund hierfür kann die Geheimhaltung kritischer Bereiche durch die Finanzinstitute sein, die um ihre Erkenntnisse und Investitionen fürchten. Folglich ist auch davon auszugehen, dass eine ganze Reihe von Studien bisher überhaupt nicht veröffentlicht wurden. Möglicherweise stellen die veröffentlichten Studien nur den Gipfel eines Eisbergs dar. Jedenfalls drosselt dies durch das Verhindern von Synergieeffekten den Fortschritt auf dem Gebiet neuronaler Netze erheblich.

C. Klassifikationsanwendungen neuronaler Netze im Finanzbereich

I. Grundlagen der neuronalen Klassifikation

1. Problemstellung und Ziele der neuronalen Klassifikation

Die Klassifikation ist neben der Prognoseanwendung der wichtigste Anwendungsbereich neuronaler Netze. Hierbei können unterschiedliche Ziele verfolgt werden. Meistens besteht die Aufgabe der Klassifikation darin, Objekte anhand von Merkmalen bestimmten, vordefinierten Klassen zuzuordnen. Neue Kreditfälle werden folglich auf der Grundlage der Erfahrung aus vergangenen Kreditfällen bearbeitet. Somit handelt es sich bei der neuronalen Klassifikation um rein empirisch-induktive Analysemethoden. Ein anderer Aufgabenbereich ist die selbstständige Segmentierung von Objekten. In diesem Fall existieren keine vordefinierten Klassen. Für diesen Einsatzbereich werden hauptsächlich Kohonen-Netze,¹⁵⁷ welche auch als „selbstorganisierende Merkmalsebenen-Netze“ bezeichnet werden, verwendet. Diese sind analog zur Clusteranalyse in der Lage, selbstständig Klassen zu bilden. Jedoch soll von einer Behandlung dieses Anwendungsgebietes abgesehen werden.

In der Finanzwirtschaft existieren bereits eine ganze Reihe von Einsatzarten neuronaler Netze als Klassifikatoren. Sieht man die Kreditgewährleistung auch als Investition an, kann festgestellt werden, dass neuronale Netze zur Klassifikation überwiegend von Großinvestoren, wie Banken und Near-Banks eingesetzt werden. Ziel ist es hierbei, den Kapitaleinsatz zu optimieren, wofür u.a. das Risiko von Störungen bei der Rückzahlung reduziert werden muss. Folglich ist eines der Hauptanwendungsgebiete neuronaler Netze bei der Klassifikation die Bonitätsprüfung von Privatkunden und Unternehmen. Dabei wird beabsichtigt, potentiell illiquide Kunden zu identifizieren. Weiterhin wird die neuronale Klassifikation häufig herangezogen, um im Rahmen einer erfolgreichen Marktsegmentierung, die Behandlung von Kunden, sei es die Betreuung oder die Entwicklung einer kundenspezifischen Anlagestrategie, zu optimieren. Dabei könnten neuronale Netze helfen, den Zielkonflikt zwischen Spezialisierung und Kundenorientierung einzudämmen.

Die Klassifizierung von Unternehmen kann natürlich auch zur Bewertung von Aktien und sonstigen Anlagemöglichkeiten herangezogen werden. Für diese existieren aber auch speziel-

¹⁵⁷ Vgl. Füsler (Wiesbaden 1995), S. 104-106.

le Klassifikationsanwendungen. Eines davon ist die Beurteilung von langfristigen Anleihen, den sog. Bonds.

Das Prinzip der Anwendung neuronaler Netze ist bei diesen genannten Problematiken ähnlich. Ob es sich um die Bonitätsprüfung von Unternehmen, Privatkunden oder ob es sich um die Beurteilung von Bonds handelt, spielt letztendlich nur für die Differenzierungs-kriterien¹⁵⁸ eine Rolle, kaum aber für das Prinzip der Informationsverarbeitung in einem neuronalen Netz. Zu den konventionellen Instrumenten der Klassifikation zählen die multivariate Diskriminanzanalyse und die Regressionsanalyse. Diese werden im nächsten Kapitel näher dargestellt. Die eingesetzten neuronalen Netze orientieren sich oftmals an deren Schemata. Jedoch ist hier, genauso wie bei den Prognoseanwendungen, der Hauptvorteil neuronaler Netze in der Berücksichtigung von nicht-linearen Wechselwirkungen zu sehen, von denen auch hier auszugehen ist. Des weiteren ist die Robustheit neuronaler Netze gegenüber lücken- und fehlerhafter Informationen, die speziell im Kreditwesen üblich sind, besonders nützlich.¹⁵⁹

Die nachfolgende Abbildung zeigt ein vereinfachtes Schema der neuronalen Klassifikation.

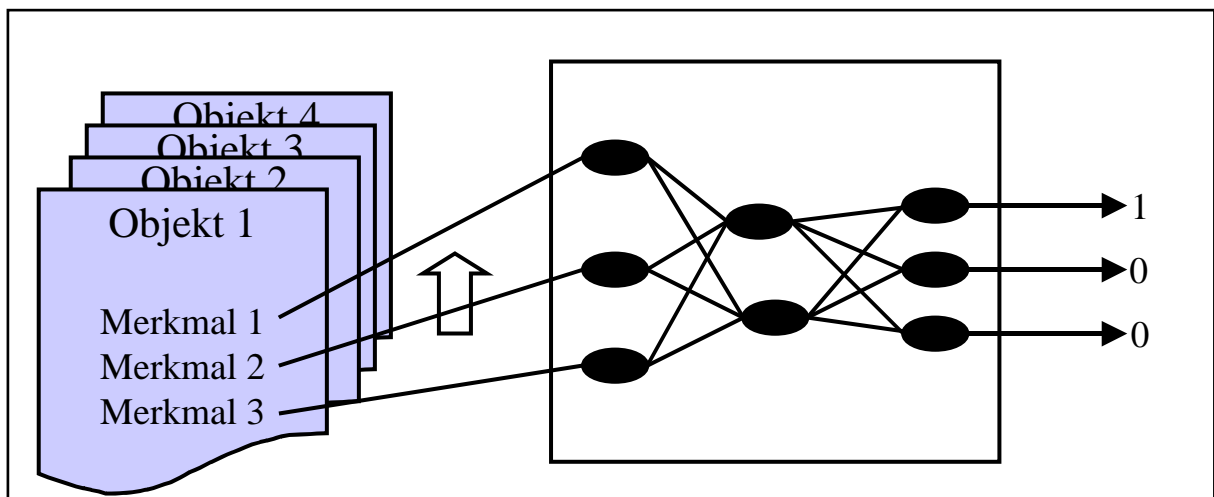


Abbildung 25: Prinzip der neuronalen Klassifikation

Die Anzahl der ausgesuchten Differenzierungskriterien determiniert in der Regel die Anzahl der Eingangsneuronen. Die Anzahl der Ausgangsneuronen wird bei der Simulation multivariater Diskriminanzanalysen durch die Anzahl der vordefinierten Klassen, denen die Objekte (Kunden, Unternehmen, Bonds etc.) zugewiesen werden sollen, bestimmt. Folglich wird hier durch das neuronale Netz einem Eingabevektor ein Ausgabevektor zugewiesen, welcher eine Eins und sonst nur Nullen enthält. Die Stelle, an der die Eins liegt, bestimmt die Zugehörig-

¹⁵⁸ Eingangsvariablen, welche die Zuordnung zu einer Klasse determinieren.

¹⁵⁹ Vgl. Schmidt-von Rhein/Rehkuglerei (München 1994), S. 503.

¹⁶⁰ Vgl. Wilbert (Frankfurt a. M. 1996), S. 82 und S. 126-127.

keit zu einer Klasse.¹⁶⁰ Bei der Simulation von Regressionsanalysen besitzt das neuronale Netz in der Regel nur einen Ausgabeneuron mit metrischem Output. Hier werden die Klassen durch Intervalle bestimmt, innerhalb derer sich der Output befindet. Problematisch ist die zeitliche Veränderung von den ausgewählten Kriterien. Diese können (noch) nicht berücksichtigt werden. Daher ist es ratsam, Differenzierungskriterien auszuwählen, die relativ stabil sind und nur geringfügige Volatilitäten aufweisen.

Auf eine detaillierte Darstellung der Problematik neuronaler Klassifikation soll an dieser Stelle jedoch verzichtet werden, da in den nachfolgenden Kapiteln einige Beispiele hierzu aufgezeigt werden. Vorher sollen aber die wichtigsten konventionellen Klassifikatoren dargestellt werden.

2. Konventionelle Ansätze

Die gängigsten Verfahren für die finanzwirtschaftliche Klassifikation sind die multivariaten statistischen Methoden. Zu diesen empirisch-induktiven Verfahren gehören die Regressionsanalyse und die multivariate Diskriminanzanalyse (MDA). Die Regressionsanalyse wird für statistische Fragestellungen relativ häufig herangezogen. Der Grund hierfür ist u.a. die hohe Flexibilität. So können Regressionsanalysen für jegliche Fragestellungen verwendet werden, bei denen eine abhängige Variable durch eine oder mehrere unabhängige Variablen beschrieben werden soll.¹⁶¹ Weiterhin sind die Anwendungsvoraussetzungen im Gegensatz zu anderen statistischen Analysemethoden, wie z.B. die multivariate Diskriminanzanalyse, weniger restriktiv. Die Regressionsanalyse beschäftigt sich im wesentlichen mit dem funktionalen Zusammenhang zwischen der abhängigen Variable (Regressant), welche die Klassenzuordnung festlegt und den unabhängigen, erklärenden Variablen (Regressoren). Allgemein betrachtet sieht die (lineare) Regressionsfunktion folgendermaßen aus:¹⁶²

$$y = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \dots + \beta_j * x_j$$

y = Variable für die Klassenzugehörigkeit

j = Anzahl der erklärenden Merkmale

β_j = Funktionskoeffizienten (-parameter)

¹⁶¹ Vgl. Pytlik (Frankfurt a. M. 1995), S. 86.

¹⁶² Vgl. Pfeifer (Frankfurt a. M. 1998), S. 88.

Die wesentliche Schwierigkeit bei Regressionsanalysen liegt in der richtigen Schätzung der Funktionsparameter. Selbstverständlich ist für die Bestimmung der grundlegenden Form der Funktion eine vorangegangene Analyse des Sachverhaltes notwendig, um eine Vorstellung über wesentliche Zusammenhänge zu erhalten.¹⁶³ Die Koeffizienten werden häufig mit der Methode der kleinsten Quadrate bestimmt.¹⁶⁴ Der Vorteil von Regressionsanalysen gegenüber vielen anderen statistischen Methoden ist die Fähigkeit, auch nicht-lineare Zusammenhänge erfassen zu können. Die oben dargestellte allgemeine Form der Regressionsfunktion entspricht der linearen Regression. Daneben gibt es aber die Möglichkeit, nicht-lineare Regressionsmodelle aufzustellen, wodurch diese Analysemethode für Wirtschaftswissenschaften höhere Relevanz erlangt.¹⁶⁵

Bei der multivariaten Diskriminanzanalyse wird eine nominal-skalierte Variable durch metrisch-skalierte Variablen bestimmt, während bei der Regressionsanalyse nur metrisch-skalierte Variablen zum Einsatz gelangen. Ansonsten existieren kaum Unterschiede zur Regressionsanalyse. Daher wird die Diskriminanzanalyse vereinfacht als Regressionsanalyse für nominal-skalierte, abhängige Variablen bezeichnet. Es werden im wesentlichen zwei wichtige Ausprägungen unterschieden: die einfache und die mehrfache Diskriminanzanalyse. Bei der einfachen Diskriminanzanalyse werden die Objekte im Gegensatz zur mehrfachen Diskriminanzanalyse in 2 Gruppen aufgeteilt.¹⁶⁶ So z.B. bei der Bonitätsprüfung oder der Insolvenzprognose, wo zu entscheiden ist, ob ein Subjekt kreditwürdig ist oder nicht, bzw. ob von einer zukünftigen Insolvenz auszugehen ist oder nicht. In diesem Zusammenhang ist zu erwähnen, dass in empirischen Untersuchungen mit der Diskriminanzanalyse in der Regel bessere Ergebnisse erzielt wurden als mit der formal sehr ähnlichen Regressionsanalyse, so dass bei der Bonitätsprüfung der Diskriminanzanalyse eine höhere Performance zugerechnet wird.¹⁶⁷ Die Diskriminanzanalyse konnte speziell zur Klassifizierung von Unternehmen anhand von Jahresabschlüssen vielfach erfolgreich eingesetzt werden und kann als Standardverfahren für derartige Anwendungsfälle betrachtet werden. Jedoch gilt die Diskriminanzanalyse trotz der zuverlässigen und empirischen Bestätigung als ausgereizt, was damit verbunden wird, dass mit dieser Methode Gruppen nur linear getrennt werden können (s. Abbildung 26).¹⁶⁸

¹⁶³ Vgl. Wilbert (Frankfurt a. M. 1996), S. 82 und S. 127-128.

¹⁶⁴ Eine ausführliche Darstellung findet sich in: Litz (München 2000), S. 92-97.

¹⁶⁵ Vgl. Litz (München 2000), S. 37-43.

¹⁶⁶ Vgl. Litz (München 2000), S. 351-355 und S. 365-366.

¹⁶⁷ Vgl. Pfeifer (Frankfurt a. M. 1998), S. 90.

¹⁶⁸ Vgl. Pytlik (Frankfurt a. M. 1995), S. 121.

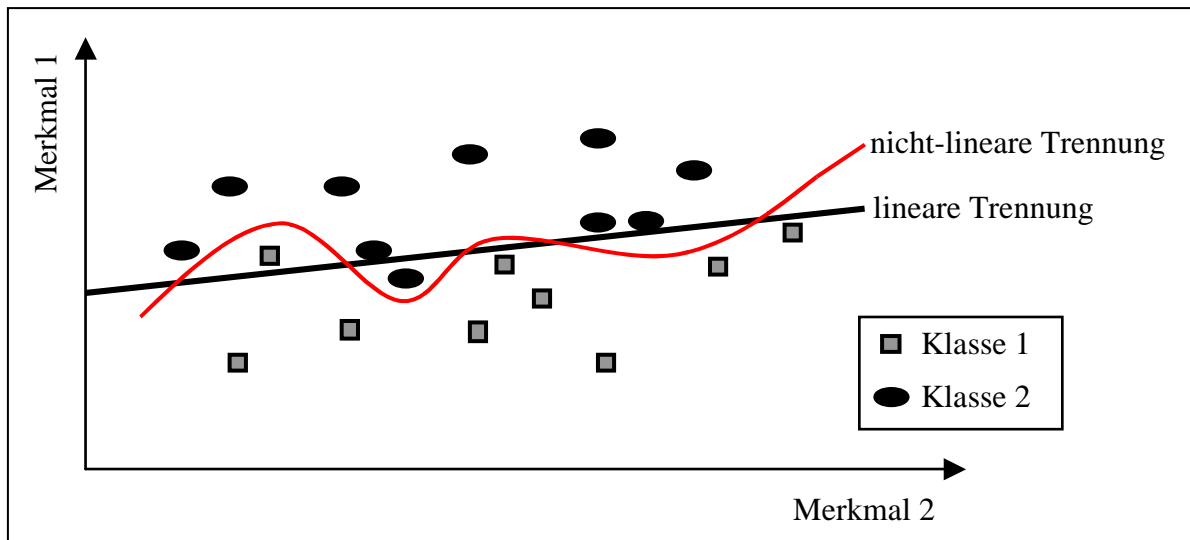


Abbildung 26: Lineare versus nicht-lineare Trennung

Wie auch aus Abbildung 9 zu sehen ist, bieten neuronale Netze in dieser Hinsicht eine vielversprechende Alternative, deren Klassifikationspotenzial praktisch unbegrenzt ist. In den folgenden Kapiteln soll anhand von Beispielen dargestellt werden, wie der konkrete Einsatz neuronaler Netze zur Klassifikation aussehen kann.

II. Kreditwürdigkeitsprüfung von Konsumentenkrediten

1. Problemstellung

Ein wichtiges Einsatzgebiet von neuronalen Netzen ist die Kreditwürdigkeitsprüfung von Konsumentenkrediten. Obwohl der Begriff „Kreditwürdigkeitsprüfung“ selbsterklärend ist, findet sich in der Literatur keine einheitliche Definition hierfür. Allgemein formuliert, wird mit der Prüfung der Kreditwürdigkeit beabsichtigt, klarzustellen, ob der Kreditnachfrager alle Voraussetzungen erfüllt, um seine Raten- und Zinsverpflichtungen erfüllen zu können.¹⁶⁹ Dies erfordert die Berücksichtigung aller relevanten Einflussfaktoren auf die Kreditwürdigkeit. Der Kreditgeber, welcher in der Regel eine Bank oder auch eine Near-Bank ist, sucht folglich alle aufschlussreichen Informationen über personelle und materielle Gegebenheiten des Antragstellers.

Abbildung 27 gibt eine Übersicht der potentiellen Einflussfaktoren. Selbstverständlich sind die personenbezogenen Daten am jeweiligen Kreditrahmen zu relativieren, was sich in der Spalte „Kreditspezifische Einflussfaktoren“ widerspiegelt.

¹⁶⁹ Vgl. Feulner (Frankfurt a. M. 1980), S. 5-6.

Personelle Einflussfaktoren	Materielle Einflussfaktoren	Umfeldbedingte Einflussfaktoren	Kreditspezifische Einflussfaktoren
Charakterliche Eigenschaften <ul style="list-style-type: none"> • Alter • Familienstand • Unterhaltspflichtige Personen • Wohndauer • Zahlungsverhalten • SCHUFA Berufliche Situation <ul style="list-style-type: none"> • Qualifikation • Art der Tätigkeit • Beschäftigungsdauer Physische Situation <ul style="list-style-type: none"> • Attest 	Einkommen <ul style="list-style-type: none"> • Verfügbarkeit • Nachhaltigkeit Vermögen Sicherheiten	Schuldnermoral Konjunktur <ul style="list-style-type: none"> • Arbeitslosenquote • Branchenrating Wohngegend Soziales Milieu	<ul style="list-style-type: none"> • Kreditbetrag • Laufzeit • Rate • ...

Abbildung 27: Potenzielle Einflussfaktoren auf die Kreditwürdigkeit¹⁷⁰

Bei diesen Einflussfaktoren ist zu berücksichtigen, dass nicht-triviale und komplexe Zusammenhänge existieren, welche die Schätzung des Kreditrisikos erschweren.¹⁷¹ Der Kreditgeber versucht im Rahmen der Kreditwürdigkeitsprüfung das Risiko zu „kontrollieren“, da dieser letztendlich der Risikoträger ist. *Keysberg* unterscheidet folgende Risikoarten im Kreditgeschäft:¹⁷²

- Verlustrisiko
- Liquiditätsrisiko
- Besicherungsrisiko

Das Liquiditätsrisiko ist ein Maß, das die Gefahr nicht rechtzeitig erfolgreicher Zahlungen beschreibt, was zu Zinseinnahmeverlusten führen kann, während das Besicherungsrisiko die Wahrscheinlichkeit der Abwertung von Kreditsicherheiten des Kreditnehmers darstellt. Die bedeutendste Risikoart ist aber das Verlustrisiko. Wenn der Kreditbetrag nur teilweise oder

¹⁷⁰ Vgl. Schnurr (Wiesbaden 1997), S. 28.

¹⁷¹ Vgl. Schnurr (Wiesbaden 1997), S. 28-30.

¹⁷² Vgl. Keysberg (Münster 1989), S. 10.

überhaupt nicht mehr zurückgezahlt werden kann, ist dieser Fall eingetreten. In der Regel zielt die Kreditwürdigkeitsprüfung ausschließlich auf die Minimierung dieser Risikoart ab.¹⁷³ Im Rahmen der Kreditwürdigkeitsprüfung werden anhand der gewonnenen Informationen Rückschlüsse auf diese Risikoart gezogen, auf deren Grundlage die Entscheidung getroffen wird, ob der beantragte Kredit bewilligt oder abgelehnt wird. Hierbei wird die klassische Kreditwürdigkeitsprüfung anhand von Fragebogen und persönlicher Einschätzung durch einen Experten immer mehr durch standardisierte und automatisierte Verfahren ersetzt. Ursache ist u.a. die verstärkte Konkurrenz und somit der steigende Kostendruck. So führt die Rationalisierung der Kreditwürdigkeitsprüfung zur maschinellen Bearbeitung von Routine-tätigkeiten und somit zur Reduzierung von Personalaufwand. Weiterhin wird dadurch eine Objektivierung der Kreditnachfrage anvisiert, wodurch Gleichbehandlung und die Reduzierung „menschlichen Versagens“ erhofft wird.¹⁷⁴

Hierfür existieren bereits eine ganze Reihe von praxisbewährten Werkzeugen. Zu diesen gehören regelbasierte Expertensysteme und mathematisch-statistische Verfahren, von denen überwiegend die multivariate Diskriminanzanalyse eingesetzt wird. Ein ebenfalls weit verbreitetes Instrument, das zu den mathematisch-statistischen Verfahren gezählt wird, sind die Credit-Scoring-Modelle.¹⁷⁵

Im Gegensatz zur Bonitätsprüfung von Unternehmen anhand von Jahresabschlüssen,¹⁷⁶ ist die Kreditwürdigkeitsprüfung von Personen stark von qualitativen Persönlichkeitsdaten abhängig, welche in der Regel nur unvollständig und unscharf erfasst werden können. Hier sollten die Stärken von neuronalen Netzen liegen.

Inwieweit neuronale Netze als relativ neues Mustererkennungsverfahren der Künstlichen Intelligenz der Rationalisierung und Objektivierung der Kreditwürdigkeitsprüfung von Konsumentenkrediten dienen können, soll im Folgenden anhand von empirischen Studien von Experten auf diesem Gebiet erörtert werden.

2. Modell von Enache¹⁷⁷

Für den Einsatz neuronaler Netze erweist sich die Auswahl der Stichprobe als besonders problematisch. Da das neuronale Netz letztendlich aus den präsentierten Mustern lernt, spielt die

¹⁷³ Vgl. Enache (Köln 1998), S. 68.

¹⁷⁴ Vgl. Krause (Düsseldorf 1993), S. 11-12.

¹⁷⁵ Eine detaillierte Darstellung findet sich in: Urbatsch/Kunath (Mittweida 1999), S. 3-11.

¹⁷⁶ Auf diese Thematik wird später noch eingegangen.

¹⁷⁷ Vgl. Enache (Köln 1998), S. 67-88.

Zusammensetzung der Trainingsmuster aus „schlechten“ und „guten“¹⁷⁸ Kreditfällen eine wichtige Rolle für die Prognosegüte des neuronalen Netzes. In der folgenden Arbeit ist diese Problematik dadurch neutralisiert, dass das Verhältnis zwischen schlechten und guten Kreditfällen in der Validierungs- und Testmenge beibehalten wird.

Den Datensatz von insgesamt 38176 Kreditfällen mit jeweils 180 Variablen lieferte für die Studie von *Enache* eine, auf Konsumentenkredite spezialisierte Bank. Von diesen, aus den Jahren 1991 und 1992 stammenden Kreditfällen, wurden 8163 Stichproben ausgewählt. Das Verhältnis zwischen guten und schlechten Kreditfällen betrug ca. 1:1, was natürlich nicht repräsentativ sein kann, da der Anteil schlecht ausgehender Kredite erfahrungsgemäß viel niedriger ist. Durch die Verwendung eines Pseudo-Zufallsgenerators wurden diese Stichproben gleichmäßig in drei gleichgroße Datensätze (Trainings-, Validierungs- und Testmenge) aufgeteilt. Die abhängige Variable wird hier als TYP bezeichnet und ist eine binäre Variable, welche die Werte 0 und 1 annimmt. Mit dieser Variable wird gemessen, wie ein Kredit verlaufen ist. Der Wert 1 wird für Kreditfälle genommen, bei denen keine Probleme bei der Rückzahlung aufgetreten sind, während mit der Null Kreditfälle gekennzeichnet werden, bei denen mindestens ein Problem aufgetreten sind. Als Beispiel hierfür gibt der Autor den Zahlungsrückstand an, wobei leider nicht näher spezifiziert wird. Es ist aber davon auszugehen, dass es sich bei diesen „schlechten“ Kreditfällen um häufige und anhaltende Zahlungsrückstände handelt. Die nachfolgende Tabelle gibt die relative und absolute Häufigkeit von diesen schlechten und guten Kreditfällen für die Trainings-, Validierungs- und Testdatensätze an.

TYP	0	1	Σ
Trainingsdatensatz	1294 (48,2%)	1390 (51,8%)	2684
Validierungsdatensatz	1320 (48,3%)	1414 (51,7%)	2734
Testdatensatz	1299 (47,3%)	1446 (52,7%)	2745
Σ	3913 (47,9%)	4250 (52,1%)	8163

Abbildung 28: Häufigkeiten der guten und schlechten Kreditfälle

Von den ursprünglich 180 beschreibenden Variablen waren viele nur sehr schwach oder überhaupt nicht besetzt, so dass diese für das Modell kaum von Nutzen waren und folglich ausgemustert wurden. Nach einer detaillierten Analyse der 36 übriggebliebenen Variablen entschied sich der Autor für die nachfolgend abgebildeten 7 Klassifikationsmerkmale.

¹⁷⁸ Diese sind je nach Autor unterschiedlich definiert.

Variable	Bezeichnung	Skalierung
Geschlecht	GESCH	Binär (0, 1)
Beschäftigungsdauer	SEIT	metrisch
Alter	GEBJAHR	metrisch
PKW-Besitz	PKW	Binär (0, 1)
Telefon-Besitz	TEL	Binär (0, 1)
Familienstand	FAM	nominal
Kredithöhe	BAR	metrisch

Abbildung 29: Ausgewählte Klassifikationsvariablen und deren Skalierung

Als Netzwerkmodell kam das MLP mit einer verdeckten Schicht zum Einsatz. Um den Einfluss der Anzahl von verdeckten Neuronen herauszufinden, wurde die Zahl dieser in 2er Schritten zwischen 1 (lineares Modell) und 23 (nicht-lineares Modell) variiert. Als Aktivierungsfunktion wurde für die verdeckten Neuronen die standardisierte, logistische Verteilungsfunktion verwendet, um hier die Ausgabewerte auf den Intervall $[0;1]$ zu skalieren.¹⁷⁹ Für den Ausgabeneuron wurde die lineare Aktivierungsfunktion verwendet.

Zum Training dieser Netze kam neben dem gängigen Backpropagation-Algorithmus (BA-MLP) eine Modifikation dieser, das Competitive Backpropagation-Algorithmus (CBA-MLP), zum Einsatz.¹⁸⁰ Mit diesem Verfahren zielt der Autor auf eine Beschleunigung des Optimierungsprozesses ab. Wie im nachfolgenden zu sehen sein wird, kann damit aber auch eine Verbesserung der Klassifikationsgüte erreicht werden.

Nach abgeschlossenem Training und Kreuzvalidierung wurden in Abhängigkeit der verdeckten Neuronen folgende Resultate für den Testdatensatz erzielt:

¹⁷⁹ Vgl. Enache (Köln 1998), S. 54.

¹⁸⁰ Eine detaillierte Darstellung dieses Verfahrens wird vom Autor separat aufgeführt (S. 45-46).

Anzahl verdeckter Neuronen	Gesamttrefferquote [%]	
	BA-MLP	CBA-MLP
1	55,3	60,1
3	59,3	62,9
5	56,6	61,1
7	59,9	61,1
9	61,7	61,9
11	59,9	61,9
13	60,5	61,9
15	60,4	61,8
17	60,1	61,8
19	61,3	63,0
21	60,2	61,9
23	60,5	62,7

Abbildung 30: Trefferquoten des Ausgangsmodells

Aus dieser Tabelle wird deutlich, dass das CBA-MLP durchgehend bessere Ergebnisse liefert. Hierbei kann festgestellt werden, dass das Netz mit 19 verdeckten Neuronen die beste Klassifikationsperformance aufweist, was den Verdacht nicht-linearer Interdependenzen erhärtet. Bei dieser Anzahl von verdeckten Neuronen lieferte das BA-MLP für die schlechten Kreditfälle eine Trefferquote von ca. 70,9%. Jedoch konnte es nur 52,6% der guten Kreditfälle richtig erkennen. Das CBA-MLP lieferte indes ein ausgewogeneres Ergebnis. Hier konnten 62,3% der schlechten und 63,5% der guten richtig erkannt werden.

Im Anschluss an diese, doch recht enttäuschenden Ergebnisse wurde vom Autor ein verbessertes Modell erarbeitet, wobei es sich lediglich um Veränderungen der Inputvariablen handelt. Erstaunlich ist die Erkenntnis, dass die Höhe des Kredites keinen Einfluss auf die Klassifikationsergebnisse ausübte. Folglich wurde diese weggelassen und anstelle der metrischen Skalierung für GEBJAHR und SEIT die „metrisch klassifizierte“ Skalierung verwendet, was nichts anderes als die Ordinalskalierung ist. Wie in der

Abbildung 31 zu sehen ist, konnten hierdurch signifikante Verbesserungen erzielt werden.

Anzahl verdeckter Neuronen	Gesamttrefferquote [%]	
	BA-MLP	CBA-MLP
1	61,1	65,5
3	64,2	65,9
5	63,7	65,3
7	63,1	67,2
9	64,8	67,2
11	63,3	67,5
13	64,6	67,1
15	64,7	66,4
17	64,4	66,4
19	65,9	67,6
21	64,1	66,7
23	64,4	67,1

Abbildung 31: Trefferquoten des verbesserten Modells

Auch hier lieferte das Netz mit 19 verdeckten Neuronen das beste Resultat. Das CBA-MLP erzielte mit einer Trefferquote von 66,5% für schlechte und 68,6% für gute Kreditfälle ein ausgewogenes Ergebnis, während sich beim BA-MLP eine Umkehrung abspielt. So wurde hier für die guten Kreditfälle eine bessere Trefferquote (69,0%) erzielt als für die schlechten (62,4%). Als Benchmark verwendete der Autor die Diskriminanzanalyse. Damit konnten 66,0% der schlechten und 68,1% der gute Kreditfälle erkannt werden, wobei die Gesamttrefferquote auf 67,1% beziffert wird. Somit war die Diskriminanzanalyse dem neuronalen Ausgangsmodell überlegen.

Erst das verbesserte Modell, welches mit dem Competitive Backpropagation-Algorithmus trainiert wurde, konnte die Ergebnisse der Diskriminanzanalyse leicht übertreffen. Dennoch konnte der Autor mit dieser Studie nicht von der Überlegenheit neuronaler Netze gegenüber konventionellen Verfahren überzeugen.

3. Weitere Studien zur Kreditwürdigkeitsprüfung von Konsumentenkrediten

Mit der Kreditwürdigkeitsprüfung beschäftigte sich der vielzitierte Experte auf dem Gebiet neuronaler Netze, *Rehkugler*, bereits zu Beginn der 90er. Nach einer Anlauf- und Experimentierphase im Jahre 1992, wurde zusammen mit *Schmidt-von Rhein* eine Studie zur Kredit-

würdigkeitsprüfung von Konsumentenkrediten mit Hilfe zweier unterschiedlicher Netzwerkmodelle veröffentlicht.¹⁸¹

Eingesetzt wurde neben dem MLP ein sog. Learning Vector Quantizer.¹⁸² Jedoch konnte mit diesem relativ neuen Netztyp keine Verbesserungen erreicht werden. Daher soll von dieser abgesehen werden.¹⁸³ Als Untersuchungsstichproben dienten 836 Datensätze zu Ratenkrediten, welche von einer großen deutschen Bank zur Verfügung gestellt wurden. Diese stammen überwiegend aus den Jahren 1988-1990. Als Differenzierungskriterien wurden 19 soziodemographische (Familienstand, Alter, Beruf, etc.) und wirtschaftliche (SCHUFA, Kontoführung, Sparrate, Zahlungsverpflichtungen, etc.) herangezogen, die binär, metrisch und nominal skaliert wurden. Aus der Skalierung ergab sich eine Anzahl von 24 Eingangsneuronen. Die Zahl der Neuronen in der einzigen verdeckten Schicht wurde variiert, während in der Ausgabeschicht ein Neuron die Kreditwürdigkeit bestimmte.

Besonderes Augenmerk setzte die Autorengemeinschaft auf die fehlenden Daten, welche als Missing Values (MV) bezeichnet werden. Bei einer Veränderung der Zusammensetzung der Trainingsmenge aus schlechten und guten Kreditfällen (1:1 und 1:4) kamen die Autoren bei durchschnittlichen Trefferquoten von ca. 70% zu folgendem Schluss:

„Der Einfluss unterschiedlicher MV-Behandlungsverfahren auf die Klassifikationsqualität ist zwar erkennbar, aber auf wenige Prozentpunkte begrenzt. Auffällig ist die starke Abhängigkeit der Netzwerkstrukturen und Klassifikationsergebnisse von der Stichprobenzusammensetzung.“¹⁸⁴ ... „Resultat unserer Untersuchungen ist, dass mit MLP... im Vergleich zur MDA systematisch bessere Klassifikationsergebnisse erzielt werden konnten. Der Vorsprung beträgt bei den jeweils besten Klassifikationsergebnissen ca. 3 bis 6 Prozentpunkte.“¹⁸⁵

Im gleichen Zeitraum stellte *Lohrbach* eine Studie zur Prüfung der Eignung neuronaler Netze für die Kreditwürdigkeitsprüfung auf.¹⁸⁶ Zum Vergleich wurden konventionelle Klassifikationsinstrumente wie das Credit-Scoring und die (lineare) MDA gegenübergestellt. Der Autor analysiert hierfür die Performance mehrerer Netzwerkmodelle. Neben dem Counter-propagation-Netzwerk und dem Learning Vector Quantizer kamen auch andere Netzwerkmo-

¹⁸¹ Vgl. Schmidt-von Rhein/Rehkugler (München 1994), S. 491-545.

¹⁸² Für einen Überblick vgl. Zell (Bonn 1994), S. 171-178.

¹⁸³ Für interessierte Leser sei auf die Textstelle verwiesen: Schmidt-von Rhein/Rehkugler (München 1994), S. 531-539.

¹⁸⁴ Schmidt-von Rhein/Rehkugler (München 1994), S. 531.

¹⁸⁵ Schmidt-von Rhein/Rehkugler (München 1994), S. 543.

¹⁸⁶ Vgl. Lohrbach (Göttingen 1994), S. 99-157.

delle zum Einsatz, welche jedoch eine weitreichenden Grundkenntnis erfordern und daher ausgelassen werden sollen. Als Datensatz dienten hier 1000 Kredite einer süddeutschen Regionalbank aus den Jahren 1973 bis 1975. Diese setzen sich zusammen aus 300 schlechten und 700 guten Krediten und wurden auch für Studien anderer Autoren herangezogen. Hier wurden neben den wirtschaftlichen und soziodemographischen Daten zusätzlich kreditspezifische Differenzierungsmerkmale wie Laufzeit, Kreditbetrag und Verwendungszweck berücksichtigt. Eine Neuheit in dieser Arbeit stellt die Berücksichtigung der Folgekosten der Fehlentscheidung. Hier unterscheidet der Autor zwischen dem Fehler 1. Art, was die fälschliche Kreditzahlung an eine zukünftig insolvente Person darstellt und dem Fehler 2. Art. Bei diesem Fehler wird der Kreditgesuch trotz der Bonität des Kreditnachfragers abgewiesen. Die durch Fehler 1. und 2. Art entstehenden Kosten wurden in der Kostenfunktion in Abhängigkeit von der angenommenen Gewichtung durch das neuronale Netz mitberücksichtigt.¹⁸⁷ In Abbildung 32 sind die Trefferquoten der beiden neuronalen Netze und der konventionellen Verfahren für die Testmenge angegeben.

Verfahren	Trefferquote [%]
CP-N	62
LVQ	72
Credit-Scoring	42
MDA	72

Abbildung 32: Trefferquoten unterschiedlicher Verfahren

Am besten schneiden hier mit 72% das LVQ und die MDA ab, während mit Credit-Scoring nur eine Trefferquote von 42% erreicht wurde. Das Counterpropagation-Netzwerk kann mit einer Trefferquote von 62% der konventionellen MDA nichts entgegensetzen. Auch hier kann keine eindeutige Überlegenheit neuronaler Netze gegenüber konventionellen Verfahren in der Kreditwürdigkeitsprüfung festgestellt werden.

Mit den gleichen Datensätzen befasste sich auch *Schnurr* in seiner, im Jahre 1997 veröffentlichten Dissertation. Hier wurde bei einer veränderten Zusammensetzung der Datensätze die Klassifikationsperformance von MLPs mit unterschiedlicher Anzahl verdeckter Neuronen geprüft.¹⁸⁸ Gegenüber der Arbeit von *Lohrbach* konnten kaum Steigerungen der Trefferquoten verbucht werden. Auch hier konnten keine eindeutigen Verbesserungen gegenüber der Diskriminanzanalyse erreicht werden, woraus der Autor das Dominieren linearer Zusammenhänge ableitet. Die wesentlichen Ergebnisse wurden folgendermaßen zusammengefasst:

¹⁸⁷ Vgl. Lohrbach (Göttingen 1994), S. 112-113.

¹⁸⁸ Vgl. Schnurr (Wiesbaden 1997), S. 187-216.

„Durch die verschiedenen Netzarchitekturen, die unterschiedlich komplexe Zusammenhänge modellieren können, konnten kaum unterschiedliche Klassifikationsquoten erzielt werden. Die dargestellten Ergebnisse bewegen sich auf einem sehr einheitlichen Niveau zwischen 70 und 75%.“¹⁸⁹

III. Sonstige Klassifikationsanwendungen neuronaler Netze

1. Kundenklassifizierung

Wie zu Beginn schon erwähnt wurde, können neuronale Netze als Klassifikatoren für jede Art von Anwendungen eingesetzt werden, bei denen anhand von mehreren Kriterien eine Einteilung in bestimmte Klassen erzielt wird. Besonders für intransparente Problemfälle, wo aber eine Vielzahl von Mustern vorliegen, erscheint der Einsatz neuronaler Netze sinnvoll. Da die Entwicklung und Implementierung neuronaler Netze keineswegs billig ist, erwartet der Anwender hierdurch Kosteneinsparungen und/oder Gewinnsteigerungen.

Die Kundenklassifizierung ist ein wichtiges und kostenintensives Aufgabengebiet. Die Kunden bilden einen heterogenen Markt bezogen auf die unterschiedlichsten Produkte. Zu diesen Produkten gehören u.a. auch Versicherungsverträge, Kredite und Dienstleistungen wie Anlageberatungen etc.. Schlagwörter der gegenwärtigen Wirtschaftsphilosophie sind Begriffe wie Marktsegmentierung, Customer Care, Spezialisierung und Rationalisierung. Dadurch erhofft man in einer Zeit harter Konkurrenz einerseits die Steigerung von Marktanteilen durch individuell zugeschnittene Produkte und andererseits die Reduzierung von Kosten durch Automatisierung von Routinetätigkeiten und Aussondern von nicht-profitablen Geschäften. Da Geschäfte mit Kunden getätigt werden, bedarf es folglich einer effizienten und zuverlässigen Einschätzung dieser.

Als ein Beispiel sei die Lebensversicherung genannt. Die Zielvariable hierfür ist die Wahrscheinlichkeit für das Eintreten eines Todesfalls. Dafür verlangt der Versicherer vom Kunden Auskunft über den körperlichen Zustand, welche in der Regel über dessen Ärzte eingeholt werden. Des Weiteren fließen individuelle Eigenschaften und Verhaltensmuster in die Kalkulation der zu zahlenden Prämie ein. So z.B. interessiert den Versicherer, ob die Person Zigaretten raucht oder Extremsport betreibt. Dadurch entsteht eine große Menge an Informationen, welche analysiert und ausgewertet werden muss. Wie diese Merkmale auf die Lebenszeit wirken, kann nur anhand von Statistiken abgeschätzt werden, wobei in den Statistiken in der Regel nur einzelne Merkmale berücksichtigt werden. Es existieren aber auch hier wenig er-

¹⁸⁹ Schnurr (Wiesbaden 1997), S. 215.

forschte, nicht-lineare Zusammenhänge sowohl unter den Merkmalen selbst (Korrelationen) als auch zwischen diesen und der Zielvariable, so dass eine Einschätzung außerordentlich schwierig wird. Vorstellbar ist die Verarbeitung solcher Informationen in einem neuronalen Netz. Dazu könnten Fragebögen herangezogen werden, von denen eine vom Hausarzt des zu Versichernden und eine vom Versicherungsnachfrager selbst ausgefüllt wird. Zusätzlich können die subjektiven Aufnahmen des Bearbeiters mitberücksichtigt werden. Das Ergebnis im Output würde dann, ähnlich wie die Schadenfreiheitsklasse bei der Kraftfahrzeugversicherung, eine Einstufung des Kunden angeben, anhand derer die Prämie in Abhängigkeit von den Vertragsparametern (Versicherungshöhe etc.) bestimmt werden kann.

Leider existieren noch keine veröffentlichten Studien zu dieser Problematik, dafür aber prinzipiell ähnliche. Dazu gehört die Anlegerklassifizierung aus dem Bankensektor zur standardisierten Anlageberatung bei einer aktiven Depotverwaltung. Ziel ist es hierbei, die Anlegermentalität mit Hilfe eines neuronalen Netzes zu bestimmen, um anhand der dadurch gewonnenen Information über die Risikoneigung eine optimale Depotbetreuung zu gewährleisten.

In der Zeitschrift „Sparkasse“ erläutert die Autorengemeinschaft *Loistl/Füser/Schmidtmeier* ein Verfahren, mit der eine Klassifizierung der Anleger in die Gruppen konservativ (Typ A), dynamisch (Typ B) und progressiv (Typ C) vorgenommen wird.¹⁹⁰ Hierzu wurde ein Fragebogen mit 21 Attributen erstellt. Der Anleger hatte die Wichtigkeit dieser Attribute aus seiner Sicht auf einer Skala von 1 bis 5 zu markieren. Als Netzwerkmodell entschieden sich die Wissenschaftler hier für einen MLP mit 3 Schichten. Die Inputschicht besitzt entsprechend der Anzahl von Attributen 21 Eingangsneuronen. Hierzu wurden die Antworten der Anleger auf das Intervall [0;1] transformiert. Für die verdeckte Schicht wurden ebenfalls 21 Neuronen verwendet. Die Zuordnung des Anlegers findet in einem einzigen Outputneuron statt, dessen metrische Ausgabe sich auf dem Intervall [0,1] befindet. Somit wird durch dieses Netz eine Regressionsfunktion simuliert, welche nicht-lineare Zusammenhänge abbilden kann. So sollten auf die Anlegertypen A, B und C idealerweise die Ausgaben 0, 0,5 und 1 erfolgen. Zum Training standen 30 Muster bereit, von denen je 10 zu einem Anlegertyp gehörten. Das neuronale Netz wurde mit dem Backpropagation-Algorithmus trainiert. Als Ergebnis lassen die Autoren verlauten: „*Die Trennschärfe des neuronalen Netzes ist anhand der Trainingsfälle überzeugend.*“

Leider verzichten die Autoren darauf, das neuronale Netz mit ungelerten Testdaten zu prüfen, dem sie folgendermaßen entgegenreten: „*Neue Datensätze, insbesondere auch reale Bei-*

¹⁹⁰ Vgl. Loistl/Schmidtmeier/Füser, S. 183-189.

*spiele, lassen sich mit diesem solcherart trainierten neuronalen Netz nach unseren Erfahrungen problemlos einordnen.*¹⁹¹

2. Bonitätsanalyse von Unternehmen

Im Gegensatz zur Klassifizierung von Personen existieren für die Klassifizierung von Unternehmen mit neuronalen Netzen zahlreiche Studien. Während die Klassifizierung von Personen von qualitativen Daten abhängig sind, welche oft nur unvollständig und unscharf erfasst werden können, liegen für die Klassifizierung von Unternehmen umfangreiche und in der Regel zuverlässige Daten vor. Dies wird garantiert durch das externe Rechnungswesen eines jeden Unternehmens, welche nach gesetzlichen Vorschriften in festgelegten Intervallen Bilanzen und Jahresabschlüsse aufzustellen haben.

Aufgrund der stetig zunehmenden Zahl von Insolvenzen, speziell in der BRD, ist eine zuverlässige Bonitätsanalyse unerlässlich. Mit der Bonität eines Unternehmens ist die Fähigkeit gemeint, zukünftigen Zins- und Tilgungszahlungen fristgerecht nachzukommen.¹⁹² Folglich beschreibt das Bonitätsrisiko die Gefahr von verspäteten Zahlungseingängen oder des totalen Kreditausfalls. Um dieses Risikopotenzial kalkulieren zu können, bedarf es einem genauen Einblick in die Situation des Unternehmens im Rahmen einer Bonitätsprüfung. Auch hier spielen qualitative Daten wie z.B. Betriebsklima oder Qualifikationen der Beschäftigten eine gewisse Rolle.¹⁹³ Diese sind aber sicherlich nicht so bestimmend wie bei der Kreditwürdigkeitsprüfung von Konsumentenkrediten. Die quantitativen Daten, welche überwiegend aus den Jahresabschlüssen entnommen werden können, dürften hier ausschlaggebend sein. Es geht also bei der Bonitätsprüfung von Unternehmen in erster Linie um die Analyse des Jahresabschlusses. Diese Bilanzen enthalten Auskünfte, welche für die Einschätzung eines Unternehmens ausreichen dürften. Neben diesen Rohkennzahlen, die direkt aus der Bilanz zu entnehmen sind (Abschreibungen, Sachanlagen, Forderungen, Fremdkapital etc.), verwenden Analysten häufig Kennzahlen in komprimierter Form, wie z.B. Cash-Flows, wirtschaftliches Eigenkapital und Rentabilitätskennzahlen.

Zu den konventionellen (empirisch-induktiven) Verfahren der Jahresabschlussanalyse zählen auch hier die verschiedenen Ausprägungen der Diskriminanzanalyse. Die nachfolgende

¹⁹¹ s.o. S. 185.

¹⁹² Vgl. Everding (Hamburg 1996), S. 8.

¹⁹³ Vgl. Pytlik (Frankfurt a. M. 1995), S. 25-27.

Abbildung 33 gibt eine Übersicht der veröffentlichten Studien zur Bonitätsprüfung von Unternehmen mit neuronalen Netzen. In der Spalte „Ranking“ sind die jeweiligen Verfahren nach der Höhe ihrer Trefferquote geordnet.

Studie	Daten	Prognosehorizont	Netzwerkmodell	Ranking
Odom/Sharda, 1990	129 Abschlüsse (1975-1982), 5 Kennzahlen	1 Jahr	MLP	1. MLP 2. MDA
Erxleben et al., 1992	3539 Abschlüsse (1973-1987) 72 Kennzahlen	1-3 Jahre	MLP	1. MDA 2. MLP
Krause, 1993	6667 Abschlüsse (1973-1986) 73 Kennzahlen	1-3 Jahre	CPN, MLP	1. MLP 2. MDA 3. CPN
Kerling/Poddig, 1994	900 Abschlüsse (1985-1990) 45 Kennzahlen	1-3 Jahre	MLP	1. MLP 2. MDA
Altmann et al., 1994	7824 Abschlüsse (1982-1992)	1 Jahr	MLP	1. MLP 2. MDA
Pytlík, 1995	600 Unternehmen (1990) 45 Kennzahlen	1-3 Jahre	MLP, LVQ	1. MLP 2. MDA 3. LVQ

Abbildung 33: Studien zur Bonitätsprüfung von Unternehmen

Bei einer detaillierten Analyse und Gegenüberstellung dieser Studien stellt sich heraus, dass mit neuronalen Netzen durchschnittlich bessere Ergebnisse erzielt werden konnten als mit einem konventionellen, linearen Instrument. Jedoch lässt die Steigerung der Performance angesichts der wenigen Prozentpunkte viel zu wünschen übrig. Einen „Sprung“ konnte keine der Studien verbuchen. Dabei ist zu berücksichtigen, dass als Differenzierungskriterien in den meisten Fällen diejenigen der Diskriminanzanalyse verwendet wurden, von denen man im vornherein Kenntnis über einen signifikanten Zusammenhang hatte. Folglich waren die neuronalen Netze in den meisten Studien benachteiligt. Es hat sich aber dennoch gezeigt, dass neuronale Netze mindestens genauso gut für Bonitätsprüfungen von Unternehmen eingesetzt werden können, wie die ausgereiften und altbewährten Diskriminanzanalysen.

3. Klassifizierung von Bonds

Regierungen, Länder und Unternehmen geben zur langfristigen Kreditaufnahme größeren Umfangs am in- und ausländischen Kapitalmarkt in der Regel festverzinsliche Inhaberschuldverschreibungen aus, in denen die Rechte der Gläubiger verbrieft sind. Diese werden als An-

leihen oder Bonds bezeichnet.¹⁹⁴ Das Risiko von Bonds setzt sich im wesentlichen zusammen aus:

- Zinssatzrisiko
- Inflationsrisiko
- Kursänderungsrisiko (bei Fremdwährungsanleihen)
- Verlustrisiko (Defaultrisiko)

Während die 3 erstgenannten Risikokomponenten nicht direkt vom Kreditnehmer abhängen, wird das Defaultrisiko im Hinblick auf die Qualität des Emittenten ermittelt. Da der vom Gläubiger geforderte Zinssatz auch eine Prämie für das Defaultrisiko, d.h. das Risiko, dass die Emission nicht zurückgezahlt wird, beinhaltet, bedarf es einer Bewertung dieses Risikofaktors.¹⁹⁵

Als anerkanntes Bewertungssystem können hierfür die Notensysteme der Rating-Agenturen herangezogen werden. Zu den bekanntesten unter diesen gehören Standard&Poor`s und Moody`s, welche mit Buchstabenkombinationen wie AAA, und Aaa das Defaultrisiko klassifizieren.¹⁹⁶ Dabei wird weniger theoriegestützt vorgegangen. Vielmehr findet das Rating auf der Basis von Erfahrungswerten, welche mit Regressions- und Diskriminanzanalysen erarbeitet werden, statt.¹⁹⁷ Da die Benotungssysteme der Rating-Agenturen als Richtschnur für die Zinssätze von Neuemissionen gelten, beschäftigen sich Autoren schon seit längerer Zeit mit den Rating-Prozessen dieser, welche strenger Geheimhaltung unterliegen. Weiterhin existieren nicht für alle Unternehmen Ratings, und die vorhandenen werden oft in unregelmäßigen Zeitabständen aktualisiert.¹⁹⁸

Die Vermutung nicht-linearer Zusammenhänge bei einer immens großen Anzahl von Einflussfaktoren bewegte bereits 1988 *Dutta und Shekhar* zur Entwicklung neuronaler Netze für das Bond-Rating. In diesem frühen Ansatz ging es lediglich darum, zu testen, ob neuronale Netze die Einstufung von Unternehmen seitens der Rating-Agenturen Standard&Poor`s und Moody`s in die Kategorie „AA“ oder „nicht AA“ nachvollziehen können. In ihrer Studie wurden 2- und 3-schichtige Netze einem linearen Regressions-Modell gegenübergestellt.¹⁹⁹

¹⁹⁴ Gabler-Wirtschaftslexikon, A-E., 13., vollständig überarbeitete Auflage, Wiesbaden 1992, S. 139.

¹⁹⁵ Vgl. Gallati (Wiesbaden 2000), S. 16-19.

¹⁹⁶ Eine Darstellung der Rating-Definitionen dieser Agenturen findet sich in Gallati (Wiesbaden 2000), S. 56.

¹⁹⁷ Vgl. Wilbert (Frankfurt a. M. 1996), S. 82 und S. 138.

¹⁹⁸ Vgl. Moody/Utans (London 1995), S. 291-292.

¹⁹⁹ Vgl. Dutta/Shekhar (San Diego 1988), S. 443-450, aus: Füser (Wiesbaden 1995), S. 293-298 und Moody/Utans (London 1995), S. 302-303.

Die mit dem Backpropagation-Algorithmus trainierten Netze wurden mit 10 bzw. 6 Eingangsvariablen versehen, zu denen auch qualitative Merkmale gehörten. Von den aus dem Jahre 1986 stammenden Datensätzen von 47 Unternehmen wurden 30 zum Training und 17 zum Testen verwendet. Die Endfälligkeit der Datensätze lag zwischen 1998 und 2003. Für die recht kleine Testmenge erzielte die Regressionsanalyse eine Trefferquote von 64,7%, während mit den neuronalen Netzen Trefferquoten von 82,4% (2-schichtig) bzw. 88,3% (3-schichtig) realisiert werden konnten. Somit wiesen in dieser Studie neuronale Netze auf Anhieb eine höhere Klassifikationsperformance auf als die lineare Regression. Die Autorengemeinschaft folgert daraus das Vorherrschen nicht-linearer Interdependenzen mit stark subjektiven Einflussgrößen, welche mit neuronalen Netzen besser zu korrelieren wären als mit linearen Verfahren.

Auch *Moody/Utans* stellen in ihrer Studie der linearen Regression ein 2-schichtiges neuronales Netz mit unterschiedlichen Modifikationen gegenüber. Ihnen ging es in dieser Studie vor allem um die Erprobung einiger heuristischer Strategien zur Entwicklung optimaler Netzwerke. Das Bond-Rating stellt für sie lediglich einen interessanten Untersuchungsgegenstand dar.²⁰⁰ Hier wurde nicht die Zugehörigkeit eines Unternehmens zu einer Klasse, sondern die direkte Zuordnung zu einer Rating-Kategorie der Standard&Poor`s durchgespielt.²⁰¹ Dazu wurden die einzelnen Abstufungen des Notensystems im Verhältnis zueinander mit ganzen Zahlen skaliert, so dass z.B. der Kategorie B- die Zahl 3 zugewiesen wurde und der Kategorie AAA die Zahl 18. Diese und die ausgewählten fundamentalen Daten der Unternehmen wurden dann auf das Intervall [0,1] gebrochen. Ohne die Musterdaten und die Netzwerkparameter näher zu beschreiben, kamen die Autoren zu einem positiven Ergebnis, welches in *Abbildung 34* zusammengefasst ist. Mit einem neuronalen Netz konnte 36,2% der Ratings von Standard&Poor`s richtig prognostiziert werden, während fast 80% in einem Schwankungsbereich von maximal einer Kategorie lag. Mit der linearen Regression wurde lediglich 21,4% der Ratings richtig erkannt, wogegen ca. 80% der Ratings in einem Schwankungsbereich von 2 Kategorien lagen.

²⁰⁰ Vgl. Füser (Wiesbaden 1995), S. 299.

²⁰¹ Vgl. Moody/Utans (London 1995), S. 292-300.

Abstand zur Zielkategorie [in Kategorien]	neuronales Netz		lineare Regression	
	[%]	Cum[%]	[%]	Cum[%]
0	36,2	36,2	21,4	21,4
1	42,3	78,5	37,2	58,6
2	14,8	93,3	21,9	80,5
> 2	6,7	100	19,5	100

Abbildung 34: Vergleich der Trefferquoten von linearer Regression und neuronalem Netz²⁰²

Somit gelang es der Autorengemeinschaft auch hier, ein neuronales Modell zu entwickeln, welches die Ratingprozesse besser nachvollziehen konnte als ein altbewährtes, empirisch-induktives Verfahren.

Mit einem Zitat dieser beiden Autoren soll gleichzeitig eine Brücke zu den anderen Klassifikationsanwendungen errichtet werden:

*„At a very basic level, our bond rating study has shown that nonlinear networks can outperform a linear regression model on an interesting financial application. ... While we tested these conclusions only on the bond rating example, the results presented here suggest that our approach merits consideration for other application problems.“*²⁰³

IV. Kritische Würdigung

Das künstliche neuronale Netz ist ein relativ neues Klassifikationsinstrument. Dennoch haben die empirischen Studien gezeigt, dass bei dem gegenwärtigen Wissensstand diese erfolgreich für die unterschiedlichsten Klassifikationsaufgaben eingesetzt werden können.

Das Einsatzgebiet erstreckt sich von der Marktsegmentierung bis zur Klassifizierung von Wertpapieren. Aus theoretischer Sicht hatte man sicherlich bessere Resultate für die empirischen Tests erwartet. Die Performance neuronaler Netze im Umgang mit verrauschten und fehlerhaften Daten konnte kaum ausgeschöpft werden. Insbesondere hatte man sich dadurch ein besseres Ergebnis bei den, durch qualitative Daten geprägten Klassifikationsproblemen, wie es bei der Kreditwürdigkeitsprüfung im Konsumentengeschäft der Fall ist, erhofft. Statt sprunghafter Verbesserung der Klassifikationsgüte, konnten durchschnittlich nur den linearen Instrumenten ebenbürtige Klassifikationsergebnisse verzeichnet werden.

Ähnlich sieht dieser Tatbestand auch für die Klassifikation von Unternehmen anhand von Jahresabschlüssen aus, obwohl hier qualitative Daten beinahe überhaupt nicht verwendet wurden. Möglicherweise liegt es daran, dass in den meisten Studien die konventionellen Klas-

²⁰² Vgl. Moody/Utans (London 1995), S. 293 und 298.

²⁰³ Moody/Utans (London 1995), S. 299.

sifikationsinstrumente mehr als nur zum Ergebnisvergleich dienen. So beschäftigten sich die Wissenschaftler in den meisten Fällen zuerst mit den konventionellen Instrumenten (hauptsächlich die lineare Diskriminanzanalyse). Manche verwenden sogar erfolgreiche Diskriminanzergebnisse von früheren Studien zur Gegenüberstellung. Diese Diskriminanzfunktionen beinhalten die wesentlichen Einflussfaktoren auf die Klassifizierung, von denen in der Regel ein gewisser Grad an Linearität bekannt ist. Zur Bestimmung der Inputgrößen für die neuronalen Netze orientierten sich die Wissenschaftler im Wesentlichen an diesen Einflussfaktoren. Folglich wurden die neuronalen Netze im Wesentlichen zur Imitation dieser konventionellen Instrumente eingesetzt. Manche Autoren beschäftigten sich sogar ausdrücklich damit. Somit war für die neuronalen Klassifikationsinstrumente durch die Annahme linearer Zusammenhänge für konventionelle Instrumente ein fester Rahmen vorgegeben. Diese konventionellen Instrumente sind ausgereift und weisen eine recht hohe Performance auf, werden aber als ausgereizt angesehen. Einen großen Vorteil haben diese Instrumente in ihrer Transparenz, welche bei neuronalen Netzen kaum gegeben ist. Die künftige Aufmerksamkeit sollte vor allem der Entwicklung von besseren Erklärungskomponenten und der höheren Nachvollziehbarkeit der Ergebnisse gelten.

Problematisch beim Einsatz neuronaler Netze ist der hohe Trainingsaufwand, welcher beispielsweise höher ist als der Aufwand zur Aufstellung einer Diskriminanzfunktion. Dies dürfte sich aber mit der zunehmenden Reifung der Algorithmen und dem Einsatz schnellerer und leistungsstarker Rechner reduzieren lassen. Folglich könnten neuronale Netze in Zukunft einen breiteren Einzug in die Finanzwelt erfahren.

Leistungsfähigere Algorithmen, mit denen in Abhängigkeit von der jeweiligen Problemstellung Netzarchitekturen und -Parameter optimal bestimmt werden könnten, würden die Wahrscheinlichkeit hierfür erhöhen. Jedoch ist genauso wie bei der Prognoseanwendung die Vorgehensweise zur Entwicklung neuronaler Klassifikationsinstrumente einem trial-and-error-Prozess unterworfen, was die Akzeptanz sicherlich negativ beeinflusst. So bedarf es zur Entwicklung neuronaler Netze sehr viel Erfahrung und Know-How auf diesem Gebiet, was selbst bei größeren Unternehmen oft nicht vorhanden ist. Eine Standardisierung für die Entwicklung neuronaler Netze könnte zu einer Substitution der konventionellen Analysemethoden zu Gunsten der Künstlichen Intelligenz führen.

Zusammenfassung und Ausblick

„Neuronale Netze in der Finanzwirtschaft“: Heute können sich die meisten darunter nicht viel vorstellen. In naher Zukunft dürfte sich dies aber ändern. Während in Deutschland die Entwicklung auf diesem Gebiet, wie oft in der IKT-Branche, doch recht zurückgeblieben ist, sind neuronale Netze in Japan und den USA kaum mehr wegzudenken. Aber auch heimische Wissenschaftler, deren Studien in dieser Arbeit bevorzugt dargestellt wurden, haben die Überlegenheit neuronaler Analyseinstrumente für finanzwirtschaftliche Zwecke anhand zahlreicher Tests belegen können.

Neuronale Netze gehören zu den konnektionistischen Systemen. Dieser Terminus steht für einen Forschungszweig, der sich aus mehreren wissenschaftlichen Disziplinen, wie z.B. Informatik, Mathematik und Biologie zusammensetzt. Aufbau und Funktionsweise dieser Netze orientieren sich am menschlichen Gehirn. Wie beim biologischen Vorbild, ist eine Vielzahl künstlicher Nervenzellen, welche als Mikro-Entscheidungsträger zu sehen sind, miteinander verkoppelt. Dadurch entsteht ein parallel arbeitendes Informationsverarbeitungs-komplex, welches beliebig komplexe und nicht-lineare Zusammenhänge aus den präsentierten Musterdaten selbstständig erforschen kann. Dadurch wird die Kenntnis über Zusammenhänge und eine wohldefinierte Modellvorstellung, welche für konventionelle Analyseinstrumente eine Notwendigkeit sind, entbehrlich.

Hier wird der Schwerpunkt auf die Auswahl geeigneter Musterdaten verschoben, aus denen das Netz die Zusammenhänge selbstständig erkennen und daraus lernen soll. Bei der Entwicklung des Netzes werden in der sog. Trainingsphase die Verbindungsgewichte zwischen den Neuronen so lange angepasst, bis es dem Netz gelingt, die vorgegebenen Ergebnisse auf der Trainings- und Testdatenmenge möglichst fehlerfrei nachzubilden. In den meisten Studien wird eine zusätzliche Kreuzvalidierungsmenge gebildet, um die Generalisierungsfähigkeit zu garantieren. All diese Datenmengen bilden die Grundlage des Netzes und sollten möglichst repräsentativ für die Daten sein, die später mit dem Netz zu untersuchen sind.

Neuronale Netze werden in erster Linie durch Topologie und das verwendete Trainingsverfahren charakterisiert. Zu den bekanntesten und in der Finanzwirtschaft am häufigsten verwendeten Vertretern neuronaler Netze gehören die Multi-Layer-Perceptrons mit dem Backpropagation-Algorithmus als Trainingsverfahren.

Ziel dieser Arbeit war es, aufzuzeigen, wie, wofür, weshalb und mit welchen Resultaten neuronale Netze in der Finanzwirtschaft eingesetzt werden können. Es hat sich gezeigt, dass neuronale Netze in der Finanzwirtschaft sehr flexibel und universell einsetzbar sind. Das Anwendungsspektrum erstreckt sich von der Klassifikation von Privatkunden bis zur Prognose von

Wechselkursen. Der Einsatz neuronaler Netze ist besonders dann sinnvoll, wenn aus einer großen Anzahl von Eingangsinformationen eine Schlussfolgerung zu ziehen ist, welche nicht durch einfache arithmetische Operationen aus den Eingangsinformationen abgeleitet werden kann. Diese Kategorie von Problemfällen ist in der Finanzwirtschaft doch recht häufig angesiedelt. Dies zeigt das breite Anwendungsspektrum neuronaler Netze, welches in dieser Arbeit mit Sicherheit nicht gänzlich dargestellt werden konnte. Die veröffentlichten und teilweise in dieser Arbeit dargestellten Studien können als ein kleiner Ausschnitt angesehen werden.

Die wesentlichen Ergebnisse dieser Studien können folgendermaßen zusammengefasst werden: Die Leistungsfähigkeit neuronaler Netze für finanzwirtschaftliche Anwendungsbereiche ist mindestens genauso hoch wie die konventioneller Instrumente. Es hat sich gezeigt, dass durch dieses neue Instrument systematisch bessere Ergebnisse und Extragewinne erzielt werden können als mit den ausgereiften, aber veralteten und konventionellen Instrumenten. Bei den Prognoseanwendungen konnten neuronale Netze eine höhere Überlegenheit aufweisen als bei den Klassifikationsanwendungen. Die multivariate Diskriminanzanalyse und die Regressionsanalyse, als konventionelle Instrumente zur Klassifikation in der Finanzwirtschaft, scheinen doch sehr ausgereift zu sein. Hier konnten neuronale Netze trotz ihrer hoch gepriesenen Eigenschaften nicht recht überzeugen.

Der Einsatz neuronaler Netze hat aber auch einige Nachteile. Ein großes Risiko birgt die Entwicklungsphase neuronaler Netze. Hier existieren noch keine zuverlässigen Verfahren und Algorithmen. Vielmehr basiert die Entwicklung neuronaler Netze für eine bestimmte Einsatzart auf Erfahrung und Zufall. Weitere Nachteile sind die oft sehr langen Trainingszeiten und die fehlende Erklärungskomponente.

Dies sind aber keine „unsterblichen Handicaps“. Dem Einzug neuronaler Netze in die Finanzwirtschaft dürfte in Zukunft nichts im Wege stehen. Mit dem Fortschritt im Informations- und Kommunikationszeitalter ergeben sich durch größere und umfangreichere Datenpools und durch eine effizientere Informationsverarbeitung völlig neue Möglichkeiten. Die Mikroprozessoren werden immer leistungsstärker und schneller. Dadurch dürfte das Problem langer Trainingszeiten für neuronale Netze entschärft werden. Des Weiteren liegen Anstrengungen vor, neuronale Netze mit neueren Analyse- und Entwicklungsinstrumenten zu hybriden Ansätzen zu verkoppeln. Dazu gehören die genetischen Algorithmen und die Fuzzy-Algorithmen. Dadurch können Schwächen einzelner Verfahren, wie z.B. die fehlende Erklärungskomponente für neuronale Netze, kompensiert werden.

Eines ist sicher: Konventionelle Analyseinstrumente mit ihren einfachen arithmetischen Operationen sind an die Grenzen ihrer Evolution gelangt und werden von der Entwicklung der

„New Economy“ keine größeren Impulse mehr erfahren. Dagegen sind neuronale Netze wie das „erste Feuer“ in der Steinzeit, dessen wahrer Nutzen erst mit dem anhaltenden Fortschritt langsam aber sicher zum Vorschein getreten war.

Literaturverzeichnis

ADAM, D. ET AL.:

Entwicklung einer Heuristik für Travelling-Salesman-Probleme auf der Basis neuronaler Netze, Veröffentlichungen des Instituts für Industrie und Krankenhaus-betriebslehre der Westfälischen Wilhelms-Universität Münster, Nr. 39, Münster 1995.

ALEX, B.:

Künstliche neuronale Netze in Management-Informationssystemen. Grundlagen und Einsatzmöglichkeiten, Wiesbaden 1998.

BAUMGARTNER, P.:

Vergleich der Anwendung Neuronaler Netze und Genetischer Algorithmen zur Lösung von Problemen der Finanzprognose, Bamberg 1998.

BAUN, S.:

Neuronale Netze in der Aktienkursprognose, in: REHKUGLER, H. ET AL.: Neuronale Netze in der Ökonomie, München 1994.

BIENDARRA, G.:

Fallbasierte Entscheidungsunterstützung für Jahresabschlussprüfungen, Frankfurt a. M. 1998.

CORSTEN, H./MAY, C.:

Anwendungsfelder Neuronaler Netze und ihre Umsetzung, in: CORSTEN, H.; May, C.: Neuronale Netze in der Betriebswirtschaft, Wiesbaden 1996.

DIETERLE, L.:

Künstliche neuronale Netze in der Finanzprognose, Berlin 1999.

DORFFNER, G.:

Konnektionismus, Stuttgart 1991.

DÜSING, R.:

Betriebswirtschaftliche Anwendungsbereiche Konnektionistischer Systeme, Hamburg 1997.

DUTTA, S./SHEKHAR, S.:

Bond-Rating: A Non-Conservative Application of Neural Networks, in: IEEE Proceedings of the 2nd Annual ICNN, San Diego 1988, Bd. 2.

- ENACHE, D.:
Künstliche Neuronale Netze zur Kreditwürdigkeitsprüfung von Konsumentenkrediten, Köln 1998.
- EVERDING, S.:
Früherkennung von Kreditbetrug mit Hilfe bankmäßiger Kreditwürdigkeitsprüfung, Hamburg 1996.
- FEULNER, W.:
Moderne Verfahren bei der Kreditwürdigkeitsprüfung im Konsumentenkreditgeschäft, Frankfurt a. M. 1980.
- FISCHBACH, G. D.:
Gehirn und Geist, in: Spektrum der Wissenschaft, H. 11, Heidelberg 1992.
- FÜSER, K.:
Neuronale Netze in der Finanzwirtschaft. Innovative Konzepte und Einsatzmöglichkeiten, Wiesbaden 1995.
- GALLATI, R. R.:
Festverzinsliche Wertpapiere. Bewertung und Strategien, Wiesbaden 2000.
- GEHRING, H. ET AL.:
Ein Entscheidungsunterstützungssystem zur Aktienanlage auf der Basis eines genetisch lernenden neuronalen Netzwerks, in: Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft der Fern-Universität Hagen, Diskussionsbeitrag Nr. 268, Mai 1999.
- GRAF, J.:
Stock Market Prediction with Neural Networks, in: GRITZMANN, P.: Operations Research `91, Heidelberg 1991.
- GRIMM, G.:
Fundamentale Wechselkursprognose mit Neuronalen Netzen. Traditionelle versus neuere Ansätze zur Wechselkursbestimmung, Wiesbaden 1997.
- HANKE, M.:
Optionsbewertung mit Neuronalen Netzen, Frankfurt a. M. 1998.
- HAUN, M.:
Simulation Neuronaler Netze. Eine praxisorientierte Einführung, Renningen-Malmsheim 1998.

HERTZ, J. A. U.A.:

Introduction to the Theory of Neural Computation, Redwood City 1991.

HILLMER, M./GRAF, J.:

Aktienkursprognose mit Statistischen Verfahren und Neuronalen Netzen: Ein Systemvergleich, in: BOL, G. ET AL.: Finanzmarktanwendungen Neuronaler Netze und ökonometrischer Verfahren, Heidelberg 1994.

HINTON, G. E.:

Gehirn und Geist, in: Spektrum der Wissenschaft, H. 11, Heidelberg 1992.

HOCKING, S.:

Oberlechner, T.: Information and Decision-Making Processes in Foreign Currency Exchange. Examining the Financial News Media, Central Banks and Commercial Banks, Wien 1996.

HÜTTNER, M.:

Markt- und Absatzprognosen, Stuttgart 1982.

KEYSBERG, G.:

Die Anwendung der Diskriminanzanalyse zur statistischen Kreditwürdigkeitsprüfung im Konsumentenkreditgeschäft, Münster 1989.

KIMOTO, T. ET AL.:

Stock Market Prediction With Modular Neural Networks, in: Proceedings of the International Joint Conference on Neural Networks, Bd. 1, San Diego 1990.

KINNEBROCK, W.:

Neuronale Netze. Grundlagen, Anwendungen, Beispiele, 2. Aufl., München 1994.

KINNEBROK, W.:

Neuronale Netze: Grundlagen, Anwendungen, Beispiele, München 1992.

KRATZER, K.-P.:

Neuronale Netze. Grundlagen und Anwendungen, München 1991.

KRAUSE, C.:

Kreditwürdigkeitsprüfung mit Neuronalen Netzen, Düsseldorf 1993.

LACKES, R.:

Neuronale Netze in der Unternehmensplanung, München 2000.

- LITZ, H. P.:
Multivariate Statistische Methoden, München 2000.
- LOHRBACH, T.:
Einsatz von Künstlichen Neuronalen Netzen für ausgewählte betriebswirtschaftliche Aufgabenstellungen und Vergleich mit konventionellen Lösungsverfahren, in: BIETHAN, J., SCHUMANN, M.: Göttinger Wirtschaftsinformatik, Band 10, Göttingen 1994.
- LOISTL, O.:
Kapitalmarkttheorie, 3. Aufl., München 1994.
- LOISTL, O./SCHMIDTMEIER, S./FÜSER, K.:
Anlegerklassifizierung mit Hilfe eines neuronalen Netzes, in: Sparkasse. Zeitschrift des Deutschen Sparkassen- und Giroverbandes, Heft 4/1993.
- MÄRTENS, H./SAUER, J.:
Ein Ablaufplanungssystem auf Basis Neuronaler Netze, in: BIETHAN, J. ET AL.: Betriebswirtschaftliche Anwendungen des Soft Computing, Wiesbaden 1998.
- MINSKY, M. L./PAPERT, S. A.:
Perceptrons: An introduction to computational geometry, London 1988.
- MOODY, J./UTANS, J.:
Architecture Selection Strategies for Neural Networks: Application to Corporate Bond Rating Prediction, in: REFENES, A. N.: Neural Networks in the Capital Markets, London 1995.
- OBST, G. ET AL.:
Geld-, Bank- und Börsenwesen, 40. Aufl., Stuttgart 2000.
- PETERS, E.:
Fractal Market Analysis. Applying Chaos Theory to Investment & Economics, New York 1994.
- PFEIFER, A.:
Früherkennung von Unternehmensinsolvenzen auf Basis handelsrechtlicher Jahresabschlüsse, Frankfurt a. M. 1998.
- PODDIG, T.:
Handbuch Kursprognose. Quantitative Methoden im Asset Management, Uhlenbruch 1999.

PODDIG, T.:

Mittelfristige Zinsprognosen mittels KNN und ökonomischer Verfahren, in: REHKUGLER, H. ET AL.: Neuronale Netze in der Ökonomie, München 1994.

PODDIG, T.;

Wallem, A.: Wechselkursprognosen, in: REHKUGLER, H. ET AL.: Neuronale Netze in der Ökonomie, München 1994.

PYTLIK, M:

Diskriminanzanalyse und Künstliche Neuronale Netze zur Klassifizierung von Jahresabschlüssen: ein empirischer Vergleich, Frankfurt a. M. 1995.

REFENES, A.;

Zaidi, A.: Managing Exchange-Rate Prediction Strategies with Neural Networks, in: REFENES, A.: Neural Networks in the Capital Market, London, England 1995.

REHKUGLER, H./PODDIG, T.:

Entwicklung leistungsfähiger Prognosesysteme auf Basis Künstlicher Neuronaler Netzwerke am Beispiel des Dollars. –Eine Fallstudie–, in: Bamberger Betriebswirtschaftliche Beiträge, 76/1990, Bamberg 1990.

REHKUGLER, H/PODDIG, T.:

Kurzfristige Wechselkursprognosen mit Künstlichen Neuronalen Netzwerken. Finanzmarktanwendungen Neuronaler Netze und ökonomischer Verfahren, in: BOL, G. ET AL.: Ergebnisse des 4. Karlsruher Ökonometrie-Workshops, Heidelberg 1994.

RIEß, M.:

Theoretische und empirische Aspekte der Eignung Neuronaler Netzwerke zur Wirtschaftsprognose, Aachen 1994.

RIGOLL, G.:

Neuronale Netze, Renningen-Malmsheim 1994.

ROJAS, R.:

Theorie der neuronalen Netze. Eine systematische Einführung, Heidelberg 1996.

RUMELHART, D. E./HINTON, G. E./WILLIAMS, R. J.:

Error-Backpropagation, Cambridge 1986.

SCHIERENBECK, H./HÖLSCHER, R.:

Bankassurance. Institutionelle Grundlagen der Bank- und Versicherungslehre, 4. Aufl., Stuttgart 1998.

SCHIMPE, H.:

Wissensbasisrevision in Expertensystemen mit Anwendungen in der Kreditwürdigkeitsprüfung, Aachen 1997.

SCHMIDT, M.:

Derivative Finanzinstrumente. Eine anwendungsorientierte Einführung, Stuttgart 1999.

SCHMIDT-VON RHEIN, A./REHKUGLER, H.:

KNN zur Kreditwürdigkeitsprüfung bei Privatkunden-krediten, in: REHKUGLER, H. ET AL.: Neuronale Netze in der Ökonomie, München 1994.

SCHMITTER, E.-D.:

Neuronale Netze: Einführung, Programmierbeispiele, praktische Anwendungen, 1. Aufl., Holzkirchen 1991.

SCHNURR, C.:

Kreditwürdigkeitsprüfung mit Künstlichen Neuronalen Netzen. Anwendung im Konsumentenkreditgeschäft, Wiesbaden 1997.

SCHÖNEBURG, E. ET AL.:

Genetische Algorithmen und Evolutionsstrategien. Eine Einführung in Theorie und Praxis der simulierten Evolution, 1. Aufl., Bonn 1994.

SERAPHIN, M.:

Neuronale Netze und Fuzzy-Logik. Verknüpfung der beiden Verfahren, Vor- und Nachteile, Anwendungen, Simulationsprogramm, München 1994.

SHAMS, R.:

Wechselkursatheorie und -politik. Einführung, München 1985.

STEINER, M./BRUNS, C.:

Wertpapiermanagement, 4. Aufl., Stuttgart 1995.

THOMA, B.:

Chaostheorie, Wirtschaft und Börse, München 1996.

UHLIG, H.:

Finanzprognosen mit Neuronalen Netzen. Eine Einführung mit Anleitung, München 1995.

WIDDEL, G.:

Theorie und Praxis der Aktienspekulation. Strategien, Instrumente und Gewinnchancen, München 1996.

WILBERT, R.:

Interpretation und Anwendung Neuronaler Netze in den Wirtschaftswissenschaften, Frankfurt a. M. 1996.

WITTKEMPER, H.-G.:

Neuronale Netze als Hilfsmittel zur Rendite- und Risikoschätzung von Aktien, Köln 1994.

ZELL, A.:

Einführung in Künstliche Neuronale Netze, in: BIETHAN, J. U.A.: Betriebswirtschaftliche Anwendungen des Soft Computing, Wiesbaden 1998.

ZELL, A.:

Simulation Neuronaler Netze, Bonn 1994.

ZIMMERMANN, H. G.:

Neuronale Netze als Entscheidungskalkül. Grundlagen und ihre ökonomische Realisierung-, in: REHKUGLER, H. ET AL.: Neuronale Netze in der Ökonomie, München 1994.

Bisher in dieser Reihe erschienen:

- Band 1: *Hölscher, Reinhold / Kremers, Markus / Rücker, Uwe-Christian:*
Industrierversicherungen als Element des modernen Risikomanagements,
Ergebnisse einer empirischen Untersuchung, 1996
- Band 2: *Hölscher, Reinhold / Rücker, Uwe-Christian / Heller, Alexander /
Strohhecker, Marcus:*
Wirtschaftlichkeitsanalysen zu aeroben und anaeroben Verfahren bei der
Abwasserreinigung in der Weinwirtschaft, 1996
- Band 3: *Hölscher, Reinhold:*
Bankbetriebliche Marktpreisrisiken im Grundsatz I, 1998
- Band 4: *Dreher, Stefan:*
Cyber Money, Entwicklungstendenzen und Abwicklungstechniken im Internet,
1999
- Band 5: *Hölscher, Reinhold / Daferner, Stefan / Bonn, Rainer / Alsfasser, Jörg:*
Finanzierung von Existenzgründungen in Rheinland – Pfalz, 1999
- Band 6: *Bülent, Acig:*
Anwendungen neuronaler Netze in der Finanzwirtschaft, 2001

Die Studien zum Finanz-, Bank- und Versicherungsmanagement können unter folgender Adresse bezogen werden:

Technische Universität Kaiserslautern
Lehrstuhl für Finanzdienstleistungen und Finanzmanagement
Postfach 3049
67653 Kaiserslautern
Telefon: 0631 / 205-4109
Telefax: 0631 / 205-3621
E-Mail: lff@wiwi.uni-kl.de
URL: <http://lff.wiwi.uni-kl.de>