Diplomarbeit Nr.:	822
Prüfer:	Prof. Dr. A. Reuter
Betreuer:	Dr. A. Zell
Beginn:	15.04.91
Ende:	15.10.91
CR-Klassifikation:	F.1.1, I.5.1

Diplomarbeit

Günther Kubiak

Vorhersage von Börsenkursen mit neuronalen Netzen

Universität Stuttgart, Institut für Parallele und Verteilte Höchstleistungsrechner

Aufgabenstellung

Aufgabe dieser Diplomarbeit ist die Vorhersage von Börsenkursen mit neuronalen Netzen. Ziel ist die Prognose des kurzfristigen Kursverlaufs von Aktien ausgewählter bekannter Firmen (z.B. aus den Bereichen Automobilbau, Banken, Chemie, Computerhersteller) anhand des bisherigen Wertverlaufs, des Verlaufs von Konkurrenzwerten, allgemeiner Börsenindices und weiterer wirtschaftlicher Börsenparameter. Die Auswahl der vorhergesagten Aktienwerte und die Wahl der dafür relevanten Parameter ist dem Bearbeiter freigestellt.

Es sollen verschiedene Modelle neuronaler Netze für die Aufgabe untersucht werden. Interessant sind hierbei mehrstufige Feedforward-Netze mit verschiedenen Varianten von Backpropagation und Quickprop, sowie Adaline, Madaline, Counterpropagation, ART1 etc. Es sollen dabei nur bereits implementierte Verfahren zum Einsatz kommen, die evtl. für die Aufgabenstellung geringfügig modifiziert werden können. Es wird ein Vergleich von 3 verschiedenen Verfahren für die gleiche Aufgabenstellung angestrebt. Für die Anwendung sollten die relevanten Parameter für die einzelnen Netze variiert werden, um empirisch die 'beste' Topologie und das 'beste' Lernverfahren für die Aufgabenstellung unter einer Menge theoretisch guter Kandidaten zu finden. Der Lernaufwand und die Güte der Vorhersagen der einzelnen Modelle sind die eigentlich interessanten Ergebnisse dieser Arbeit.

Die Durchführung der Versuche soll bevorzugt mit dem am IPVR entwichelten Netzwerksimulator SNNS erfolgen.

Kurzfassung

Gegenstand dieser Arbeit ist die Vorhersage des kurzfristigen Verlaufs von Börsenkursen mit neuronalen Netzen. Ausgehend von verschiedenen Börsendaten wird versucht, die Kursentwicklung des nächsten Tages zu prognostizieren. Als Eingabedaten dienen dabei der bisherige Kursverlauf des Prognosetitels, der Verlauf von Konkurrenzwerten, sowie verschiedenste Börsenindices, die das Börsengeschehen auf unterschiedlichen Ebenen beschreiben. Es werden verschiedene Prognosearten durchgeführt. Zum einen wird der genaue Kurs, bzw. die Kursänderung, zum anderen nur die Tendenz der Kursentwicklung vorhergesagt. Die zur Verfügung stehenden Datenmengen umfassen zwei Zeiträume, in denen unterschiedliche Kurs- und Indexwerte vorhanden sind.

Es werden die Lernverfahren *Backpropagation, Quickprop* und *Counterpropagation* untersucht. Bei den Versuchen werden die Netztopologie, der Lern- und Vorhersagezeitraum, die Eingabeschemata und die Lernparameter variiert.

Die Versuche werden auf dem Stuttgarter Neuronale Netze Simulator SNNS durchgeführt.

Inhaltsverzeichnis

Tabellenverzeichnis

Abbildungsverzeichnis

Kapitel 1

Einleitung

Die Börse spielt eine wichtige Rolle in der Weltwirtschaft. Sie spiegelt nicht nur die Erwartungen der Kapitalanleger wieder, sondern ist auch Barometer für die zukünftige wirtschaftliche Gesamtentwicklung. Während für den letztgenannten Aspekt die langfristige Kursentwicklung von Bedeutung ist, ist im ersten Fall die kurzfristige Kursentwicklung von Intresse.

Die Prognose dieses kurzfristigen Kursverlaufs mit Hilfe neuronaler Netze ist Gegenstand dieser Arbeit. Ausgehend von verschiedenen, zu einem bestimmten Zeitpunkt zur Verfügung stehenden Börsendaten wird versucht, die Kursentwicklung des nächsten Tages vorherzusagen. Solche Börsendaten sind z. B. der bisherige Kursverlauf, die Entwicklung von Konkurenzwerten und Börsenindices auf Branchen- und gesamtwirtschaftlicher Ebene. Dabei liegt die Annahme zugrunde, daß es, trotz der Vielfältigkeit der Einflußfaktoren auf das Börsengeschehen, möglich ist aus diesen Daten Abhängigkeiten und Zusammenhänge zu erkennen, die zu einem gewissen Grad für den Verlauf der Kurse verantwortlich sind. Zu den nicht betrachteten Einflußfaktoren gehören beispielsweise politische und wirtschaftliche Ereignisse und Nachrichten in nationalem und internationalem Rahmen, die sich auf alle Aktienkurse auswirken, aber auch solche, die nur eine Branche oder eine einzelne Unternehmung betreffen. Oft sind auch nur Gerüchte oder Spekulationen Anlaß für Kursänderungen.

Eine der Eigenschaften neuronaler Netze ist die, daß sie aus den ihnen präsentierten Daten Zusammenhänge erkennen können, die nicht offensichtlich sind. Genau diese Fähigkeit macht die Anwendung neuronaler Netze für die Börsenkursprognose interesant. Zudem wurden in anderen Versuchen [Schö 90a, Schö 90b] bereits vielversprechende Ergebnisse erzielt.

Es werden die Lernverfahren *Backpropagation, Quickprop* und *Counterpropagation* auf verschiedene Netztopologien mit unterschiedlichen Lernparametern angewendet. Für die Versuche zufällig ausgewählt wurden die Aktientitel Degussa, Dresdner Bank und RWE. Diese werden noch durch Titel ergänzt, für die nur einzelne Versuche durchgeführt wurden. Die Prognosen sind zum einen der genaue Kurs, bzw. die Höhe der Kursänderung, zum anderen, auch abhängig vom Lernverfahren, nur die Richtung der Kursentwicklung. Sämtliche Versuche wurden mit dem Stuttgarter Neuronale Netze Simulator SNNS durchgeführt.

In Kapitel 2 erfolgt eine kurze Einführung in die Funktionsweise neuronaler Netze und eine Erläuterung der in dieser Arbeit verwendeten Lernverfahren. In Kapitel 3 wird kurz ein anderer Ansatz zur Börsenkursprognose mit neuronalen Netzen vorgestellt. Kapitel 4 enthält eine Beschreibung der zur Verfügung stehenden Datenmengen. Die Versuchsbedingungen werden in Kapitel 5 erläutert. Dazu gehören die Simulatorumgebung, die Erklärung der Art der verwendeten Eingabedaten und deren Skalierung, sowie die Bewertungskriterien für Aufwand und Erfolg eines Versuchs. Die Darstellung der Progoseergebnisse erfolgt in Kapitel 6. Ein Vergleich mit einfachen, konventionellen Prognoseverfahren folgt in Kapitel 7. In Kapitel 8 wird eine zusammenfassende Betrachtung durchgeführt, Kapitel 9 enthält einen Ausblick auf weitere Alternativen.

Kapitel 2

Neuronale Netze

2.1 Einführung

Die Erforschung von neuronalen Netzen hat des Ziel des Verstehens und Anwendens der Funktionsprinzipien des menschlichen Gehirns. Sie werden zum einen benützt um zu modellieren, wie Wissen gespeichert und verarbeitet wird, zum anderen sind sie ein Algorithmus zur Lösung von bestimmten Problemstellungen. Erfolgreiche Anwendungen gibt es bereits auf den Gebieten der Zeichenerkennung, der Spracherkennung und der Mustererkennung.

Neuronale Netze bestehen aus Einheiten, auch Zellen oder Neuronen genannt, die nur einfache Verarbeitungsschritte parallel ausführen können und die mehrere Eingänge, aber nur eine Ausgabe besitzen. Diese Zellen sind hochgradig untereinander verbunden. Jeder Verbindung zwischen zwei Zellen ist ein Gewicht zugeordnet, das die Stärke der Verbindung bestimmt. Jede Zelle berechnet aus den gewichteten Eingabewerten, die sie von den vorgeschalteten Zellen erhält, ihre Eingabe und daraus mit Hilfe der Aktivierungsfunktion ihre Aktivierung. Die Ausgabefunktion bestimmt aus dem Aktivierungswert die Ausgabe, die an die nachfolgenden Zellen weitergegeben wird. Abbildung 2.1 zeigt schematisch eine Zelle und ihre Einbindung in das Netzwerk.

Auf diese Weise wird zu einem Eingabemuster, das an die Eingabeebene des Netzes angelegt wird, die entsprechende Ausgabe des Netzes berechnet. Die Anwendung von neuronalen Netzen gliedert sich in zwei Phasen. In der ersten Phase, der Lernphase, wird dem Netz eine Menge von Trainingsmustern präsentiert aufgrund derer die Verbindungsgewich-

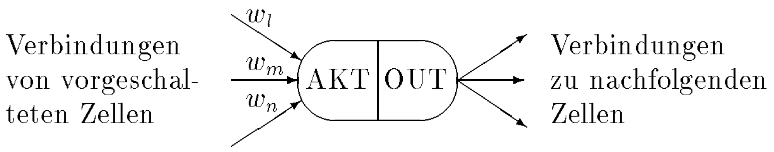


Abbildung 2.1: Zelle eines Netzes

te jeder Zelle so justiert werden, daß das Verhalten des gesamten Netzes dem gewünschten Verhalten entspricht. Die Gewichtsänderungen der einzelnen Verbindungen werden durch die Lernregeln des angewandten Lernverfahrens gesteuert. Während dieser Phase bildet das Netz interne Repräsentationen der Lernmuster, dabei werden aus den Mustern Zusammenhänge extrahiert, die für die Erzeugung der Ausgabemuster verantwortlich sind und die somit das Verhalten des Netzes bestimmen. So können später auch Muster richtig verarbeitet werden, die das Netz zuvor nicht gelernt hat. Dieses Wissen ist über die Verbindungen des gesamten Netzes verteilt. Ist im Trainingsmuster die geforderte Ausgabe des Netzes vorhanden, sodaß diese bei der Festlegung der Gewichte verwendet werden kann, so spricht man von überwachtem Lernen, enthält die Trainingsmenge nur die Eingabe, so handelt es sich um selbstständiges Lernen. In der zweiten Phase wird dann auf das gelernte Wissen zugegriffen um die gewünschte Aufgabe zu erfüllen.

2.2 Lernverfahren

2.2.1 Backpropagation

Das Backpropagation-Lernverfahren ist eine Form des überwachten Lernens. Es eignet sich zum Training von mehrstufigen Netzen. Diese bestehen aus einer Eingabeebene, beliebig vielen inneren Ebenen, das sind solche Ebenen, die nicht direkt mit den Ein- und Ausgabewerten des Lernmusterpaares in Berührung kommen, und einer Ausgabeebene. In der Eingabeebene findet keine Verarbeitung statt, sie hat nur die Funktion, die Eingaben an die Zellen der ersten inneren Schicht zu verteilen. Der Lernvorgang läuft in folgenden Schritten ab. Ein Musterpaar wird dem Netz präsentiert. Der Eingabevektor wird durch das Netz propagiert, d. h. die Netzausgabe wird berechnet. Für jede Zelle der Ausgabeebene wird die Abweichung zwischen der erzeugten und der gewünschten Ausgabe, die durch den Ausgabevektor im Musterpaar gegeben ist, gebildet. Dieser Fehler wird durch das Netz zurückpropagiert und die Gewichte werden so korrigiert, daß der Fehler minimiert wird. Die Justierung der Gewichte geschieht nach folgenden Gleichungen [Wass 89] :

$$\Delta w_{pq,k} = \eta \cdot \delta_{q,k} \cdot o_{p,j} \tag{2.1}$$

$$w_{pq,k}(n+1) = w_{pq,k}(n) + \Delta w_{pq,k} \tag{2.2}$$

wobei:

$\eta =$	Lernrate
$\delta_{q,k} =$	Beitrag der Zelle q in Ebene k zum festgestellten Fehler
$o_{p,j} =$	Ausgabe von Zelle p in Ebene j
$w_{pq,k}(n) =$	Gewicht der Verbindung von Zelle p zu Zelle q der nachfolgenden Ebene k im Schritt n
$w_{pq,k}(n+1) =$	dto. im Schritt n+1
j,k =	Indices der Ebenen, Ebene j liegt vor Ebene k, bei der Rückpropagierung wird also Ebene k zuerst erreicht.

Bei der Berechnung von $\delta_{q,k}$ für Zelle q in Ebene k ist zwischen Ausgabeebene und inneren Ebenen zu unterscheiden. Für die Ausgabeebene ist der Zielwert der Zelle durch den

Ausgabevektor im Musterpaar gegeben. Hier errechnet sich δ aus dem Produkt des Ableitungswertes der Aktivierungsfunktion und der Differenz zwischen Zielwert und tatsächlicher Ausgabe der Zelle.

$$\delta_{q,k} \; = \; act'(o_{q,k}) \cdot (Zielwert_{q,k} - o_{q,k}) \tag{2.3}$$

Für die inneren Ebenen gibt es keinen Zielwert für die Ausgabe einer Zelle, deshalb werden die bereits berechneten δ der Zellen der nachfolgenden Ebene benutzt, um die δ der Zellen der aktuellen Ebene zu berechnen.

$$\delta_{q,k} \; = \; act'(o_{q,k}) \cdot (\sum_r \delta_{r,l} \cdot w_{ql,r}) \tag{2.4}$$

wobei:	k,l =	Indices der Ebenen, Ebene k liegt vor Ebene l, für Ebene l liegen also die δ-Werte schon vor
	$\sum_r$ =	Summe über alle Zellen r einer Ebene

Dies ist das Zurückpropagieren des Fehlers, der dem Verfahren seinen Namen gegeben hat.

2.2.2 Quickprop

Beim Quickprop-Verfahren handelt es sich ebenfalls um überwachtes Lernen. Die Abfolge der Schritte ist dieselbe wie beim Backpropagation-Verfahren, jedoch wird bei der Berechnung der Gewichte anders vorgegangen. Es wird versucht die größtmögliche Gewichtsänderung zu bestimmen um zum Minimum der Fehlerkurve zu gelangen. Dem Quickprop-Verfahren liegen zwei Annahmen zugrunde [Fahl 88].

1. Die Fehlerkurve kann durch eine nach oben geöffnete Parabel approximiert werden.
2. Die Änderung der Steigung der Fehlerkurve, hervorgerufen durch die Änderung des betrachteten Gewichts, wird nicht durch alle anderen, gleichzeitig erfolgenden Gewichtsänderungen beeinflußt.

Die Steigung der Fehlerkurve im vorigen und aktuellen Schritt und die Gewichtsänderung im vorigen Schritt werden benutzt, um die Parabel zu bestimmen. Dann wird die neue Änderung des Gewichtes so bestimmt, daß man zum Minium der Parabel gelangt. Gleichung 2.5 beschreibt diesen Vorgang näherungsweise.

$$\Delta w(t) = \frac{s(t)}{s(t-1) - s(t)} \cdot \Delta w(t-1) \tag{2.5}$$

Dabei ist s=∂E/∂w, also die Steigung der Fehlerkurve. Ist die aktuelle Steigung kleiner als die alte Steigung, aber in derselben Richtung, so erfolgt eine Gewichtsänderung in dieselbe Richtung, deren Höhe von der alten Gewichtsänderung abhängig ist. Man bewegt sich somit weiter in Richtung des Minimums. Ist die aktuelle Steigung in entgegengesetzter Richtung wie die Steigung im vorigen Schritt, so ist das Minimum der Fehlerkurve bereits überschritten, der nächste Schritt führt dann wieder zurück. Wenn die alte und neue Steigung dieselbe Richtung besitzen, aber die neue größer als die alte ist, so bewegt man sich auf der Fehlerkurve nach oben, weg vom Minimum. Für diesen Fall wird ein neuer Parameter μ

eingeführt, der die maximale Gewichtsänderung bestimmt. Es wird festgelegt, daß keine Änderung eines Gewichtes größer sein darf als μ mal die vorige Änderung des betrachteten Gewichts. Um den korrekten Ablauf des Verfahrens zu gewährleisten, ist noch ein weiterer Parameter ν von Nöten. Sein Wert wird zur Steigung addiert und soll verhindern, daß zu große Gewichte errechnet werde, die zu einem Überlauf des Zahlenbereichs führen könnten. Dieser Parameter kann sehr klein gewählt werden.

Das Risiko des Verfahrens liegt in den Eingangs erwähnten Voraussetzungen. Treffen diese für das betrachtete Problem zu, so können erhebliche Verbesserungen des Lernaufwandes gegenüber Backpropagation erzielt werden.

2.2.3 Counterpropagation

Das Counterpropagation-Verfahren ist ein von Robert Hecht-Nielsen [Hech 90] entwickeltes Lernverfahren. Es verbindet die Prinzipen der *selbstorganisierenden Karten* von Kohonen, sowie des *outstar* von Grossberg miteinander. Die Counterpropagation-Netze bestehen aus drei Ebenen, einer Eingabeebene, die nur die Eingaben an die folgende Ebene verteilt, einer Kohonenebene, die die Eingabevektoren in Gruppen klassifiziert, und einer Grossbergebene, die zu jeder Klasse die gewünschte Ausgabe erzeugt. Die Kohonenebene arbeitet nach dem 'winner takes all'-Prinzip. Alle Zellen berechnen aus dem Eingabevektor und ihrem Gewichtsvektor ihre Eingabe, in Form des inneren Produkts der beiden Vektoren. Die Zelle mit dem höchsten Wert erzeugt als Ausgabe eine Eins, alle anderen Zellen geben eine Null aus. Während der Lernphase werden die Gewichte zwischen Eingabe- und Kohonenebene so justiert, daß alle Vektoren einer Klasse dasselbe Kohonenneuron aktivieren. Im Rahmen einer Vorverarbeitung werden die Eingabevektoren normalisiert. In der Lernphase wird die Eingabe jeder Kohonenzelle berechnet, wie oben beschrieben. Dann werden die Gewichte der Zelle verändert, deren Eingabe am höchsten ist, d.h. deren Gewichtsvektor dem Eingabevektor am nächsten liegt. Dieser Gewichtsvektor wird dem Eingabevektor noch ähnlicher gemacht. Dies geschieht nach folgender Formel:

$$w_{i,k}(n+1) = w_{i,k}(n) + \alpha \cdot (x_i - w_{i,k}) \tag{2.6}$$

wobei gilt:

$$\begin{aligned}
\alpha &= \text{Lernrate für Kohonenebene} \\
w_{i,k}(n) &= \text{Gewicht der Verbindung von Eingabezelle i zu} \\
& \text{Zelle k in Kohonenebene im Schritt n} \\
w_{i,k}(n+1) &= \text{dto. im Schritt n+1} \\
x_i &= \text{i-te Komponente des Eingabevektors}
\end{aligned}$$

Beim Training der Kohonenebene handelt es sich also um selbstständiges Lernen. Es wird kein Ausgabevektor benötigt, um die Gewichtsänderungen zu bestimmen.

Die Grossbergebene soll zu jeder, durch die Kohonenebene bestimmten Klasse die entsprechende Ausgabe erzeugen. Dazu werden während der Trainingsphase die Ausgaben der Zellen zunächst berechnet und die Differenz zu der gewünschten Ausgabe gebildet. Dann werden nur die Gewichte justiert, die einer Verbindung zugeordnet sind, die von einer aktivierten Kohonenzelle zum Grossbergneuron führt(Gleichung 2.7).

$$w_{i,j}(n+1) = w_{i,j}(n) + \beta \cdot (y_j - w_{i,j}(n)) \cdot k_i \tag{2.7}$$

Es gilt:

β	=	Lernrate für Grossbergebene
$w_{i,j}(n)$	=	Gewicht der Verbindung von Kohonenzelle i zu Grossbergzelle j im Schritt n, ist die Kohonenzelle aktiviert, so bestimmt dieses Gewicht die Ausgabe
$w_{i,j}(n+1)$	=	dto. im Schritt n+1
y_j	=	Ausgabe, die Grossbergneuron j erzeugen soll
k_i	=	Ausgabe des Kohonenneurons i, entweder 0 oder 1

Die Gewichte werden so angepaßt, daß bei jeder Aktivierung einer Grossbergzelle die gewünschte Ausgabe erzeugt wird. Dies entspricht dem *outstar*-Prinzip. Das Trainieren dieser Ebene erfolgt, im Gegensatz zur Kohonenebene, durch überwachtes Lernen. Dieser Umstand spiegelt sich im Namen des Verfahrens wieder.

In der Implementierung des Vefahrens im SNNS ist noch ein weiterer Parameter vorhanden. Er steuert die Auswahl einer neuen Kohonenzelle, und damit die Einteilung der Eingaben in Klassen.

Kapitel 3

Ansätze zur Börsenkursprognose mit neuronalen Netzen

Versuche zur Prognose des kurzfristigen Kursverlauf mit neuronalen Netzen wurden bereits von Schöneburg [Schö 90a, Schö 90b] durchgeführt. Er untersuchte Perceptron-, Adaline-, Madaline- und Backpropagation-Netzwerke. Nur mit letztgenanntem Verfahren ist die Prognose des genauen Kurses möglich. Mit den anderen Verfahren, mit denen nur die Tendenz der Kursänderung vorhergesagt wurde, wurden, abhängig von Aktientiteln, Netztopologie, Eingabedaten und anderen Parametern, Trefferquoten zwischen 40% und 80% erreicht. Es wurden die Kursverläufe der BASF-, Mercedes- und Commerzbank-Aktie prognostiziert. Als Trainingsdaten wurden die Notierungen von 42 Börsentagen verwendet. Die Prognose wurde über einen Zeitraum von maximal 56 Tagen durchgeführt.

Als Eingabedaten für die Backpropagation-Versuche wurden die Tageskurse der letzten 10 Börsentage verwendet. Als beste Netztopologie stellte sich ein Netz mit zwei inneren Ebenen mit je 10 Zellen heraus. Je 5 Zellen einer Schicht arbeiten mit der logistischen und der Sinus-Aktivierungsfunktion (s. Kapitel 5.1). Bei den Backpropagation-Ergebnissen zeigte sich, daß das Netz für sein Verhalten folgende Regel entwickelte [Schö 90a]: 'Sage für den nächsten Tag beinahe den gleichen Wert voraus wie für den aktuellen Tag, dann wird der Fehler relativ gering belieben, da eine Kurstendenzänderung unwahrscheinlicher ist, als eine Beibehaltung der Kurstendenz'. Dies führte dazu, daß die Prognosekurve dem tatsächlichen Kursverlauf nur zeitlich verschoben folgte. Das ist ein Verhalten, das bei herkömmlichen Verfahren, wie z.B. der exponentiellen Glättung auftritt. Verläuft der Kurs einer Aktie sehr beständig, d.h. eine eingeschlagene Richtung wird über mehrere Tage beibehalten, und es treten nur sehr selten Änderungen in der Richtung der Kursentwicklung auf, so lassen sich hiermit natürlich relativ hohe Trefferquoten erzielen. So betrug bei der Prognose der Mercedes-Aktie die Trefferrate 64%, was das beste erreichte Ergebnis war. Angaben zur Genauigkeit wurden nicht gemacht.

Von Schöneburg wurde nur das oben beschriebene Eingabeschema verwendet, das aus den 10 letzten Tagesnotierungen der Absolutkurse des zu prognostizierenden Titels bestand. Hier sind Erweiterungen möglich, die in dieser Arbeit vorgenommen werden. So werden beispielsweise Börsenindices verwendet, die das Börsengeschehen auf allgemeiner und branchenspezifischer Ebene beschreiben. Der Verlauf von Konkurrenzwerten wird mit berücksichtigt. Es wird nicht nur der Absolutkurs, sondern auch die Kursänderung prognostiziert. Bei den Versuchen, bei denen nur der bisherige Kursverlauf als Eingabe dient, werden Prognosen mit längeren Lern- und Vorhersagezeiträumen durchgeführt. Außerdem wird die Länge des Rückschauzeitraumes variiert. Die Güte der erreichten Ergebnisse wird anhand der durchschnittlichen Abweichung vom tatsächlichen Kurs beurteilt. Mit Hilfe der relativen Abweichung können die Ergebnisse der verschiedenen Titel miteinander verglichen werden.

Auch noch hier zu nennen, obwohl nicht direkt die Börsenkursprognose betreffend, sind die Versuche von Lapedes und Farber [Lape 88]. Sie prognostizierten den Verlauf einer chaotischen Zeitreihe, indem sie als Eingabedaten nur den bisherigen Verlauf der Kurve verwendeten. Sie benützten ein Netz mit zwei inneren Ebenen, das mit Backpropagation trainiert wurde und erreichten dabei hohe Genauigkeiten.

Kapitel 4

Verfügbare Datenbasis

Für die Versuche standen Daten aus zwei verschiedenen Zeiträumen zur Verfügung. Zum einen ist dies das gesamte Jahr 1988, das 247 Handelstage umfaßt. Diese Datenmenge enthält nur die Kursnotierungen der Aktientitel und des FAZ-Index. Die Wertebereiche der prognostizierten Titel und des Indexes sind Tabelle 4.1 zu entnehmen.

Titel	Kurs		Änderung	
	Min.	Max.	Min.	Max
Degussa	268.0	386.0	-20.0	+30.0
Dresdner Bank	215.0	315.0	-11.2	+15.0
RWE	187.0	245.2	-13.5	+15.5
Badenwerk	169.0	191.0	-6.0	+9.0
BHF-Bank	272.0	442.0	-29.0	+39.5
Hoechst	235.0	315.0	-13.4	+11.6
FAZ-Index	396.4	553.6	-14.1	+20.0

Tabelle 4.1: Wertebereiche 1988

Der zweite Zeitraum reicht vom 27.02.1991 bis zum 20.06.1991 und zählt 75 Handelstage. In dieser Menge sind vielfältigere Daten enthalten als in der ersten Datenmenge. Sie enthält die Notierungen des Deutschen Aktienindex DAX, des Dow-Jones-Index der New Yorker Börse, die branchenbezogenen Indices von Commerzbank und WestLB, sowie die Notierungen der benötigten Aktientitel an der Frankfurter Börse. Bei letztgenannten handelt es sich um die fortlaufenden Notierungen, die während des Handels nach jedem Abschluß eines bestimmten Umfangs neu festgelegt werden und von denen hier jeweils die Schlußnotierung erfaßt wurde. Für einen Titel ist auch die Kassennotierung vorhanden. Die Wertebereiche sind aus Tabelle 4.2 ersichtlich.

Titel	Kurs		Änderung	
	Min.	Max.	Min.	Max
Degussa	320.1	385.0	-14.6	+14.0
Dresdner Bank	347.0	398.8	-14.0	+10.0
RWE	367.5	423.3	-11.0	+19.0
DAX	1498.44	1715.80	-36.51	+54.06
Dow Jones	2843.02	3035.33	-75.54	+58.41
Commerzbank-I.				
Chemie	116.7	134.6	-3.0	+5.3
Banken	104.9	117.6	-3.3	+3.3
Energie	102.8	118.8	-3.3	+3.3
WestLB-Index				
Chemie	155.74	177.98	-3.98	+5.95
Banken	279.12	314.39	-6.84	+6.66
Versorgung	258.20	296.94	-6.78	+6.57

Tabelle 4.2: Wertebereiche 1991

Kapitel 5

Versuchsbedingungen

5.1 Simulatorumgebung

Die Versuche wurden mit dem Stuttgarter Neuronale Netze Simulator SNNS durchgeführt. Dieser Simulator wurde am Institut für Parallele und Verteilte Höchstleistungsrechner (IP-VR) der Universität Stuttgart, im Rahmen von studentischen Arbeiten, unter Leitung von Dr. Andreas Zell, seit 1989 entwickelt und wird ständig erweitert. Er besteht aus drei Komponenten:

Simulator-Kern Er verwaltet die Netze und stellt die nötigen Operationen darauf zur Verfügung.

Nessus-Compiler Mit dem Compiler können, mit Hilfe einer eigens entwickelten Beschreibungssprache, strukturierte Netze erzeugt werden.

Graphikoberfläche Die graphische Benutzeroberfläche steuert den Simulator-Kern, liefert die graphische Darstellung der Netze und erlaubt vielfältige Manipulationen der Netze.

Nähere Informationen sind [Zell 91] zu entnehmen.

Als Aktivierungsfunktion für die Zellen steht u. a. die sigmoide, logistische Funktion der Form

$$act = \frac{1}{1 + e^{-NET-\theta}} \tag{5.1}$$

zur Verfügung. NET bezeichnet die Eingabe einer Zelle, θ ist ein Schwellenwert(Bias), der während der Lernphase berechnet wird. Für die Versuche wurde zusätzlich die Sinus-Aktivierungsfuktion implementiert. Im Intervall [$-\pi/2$,$\pi/2$] nimmt sie den Verlauf der Sinusfunktion an, für Werte $< -\pi/2$ ist -1.0, für Werte $> \pi/2$ ist 1.0 der Funktionswert. In Abbildung 5.1 ist der Kurvenverlauf beider Funktionen dargestellt. Der Schwellenwert der logistischen Funktion ist dabei Null.

Der Wertebereich ist bei der Sinus-Aktivierungsfunktion größer, während die logistische Funktion auf einem größeren Definitionsbereich unterscheidbare Ausgaben liefert.

Der Simulator-Kern ermittelt nach jedem Lernzyklus einen Absolutfehler als Summe der quadrierten Differenzen zwischen Zielwert und errechneter Ausgabe, für alle Ausgabezellen, über alle gelernten Muster (Gleichung 5.2).

$$E = \sum_m \sum_z (Zielwert_z - Ausgabe_z)^2 \quad (5.2)$$

wobei: z : Zellen der Ausgabeebene
m : gelernte Muster

Zur Errechnung des durchschnittlichen Fehlers wird dieser Fehler durch die Zahl der Zellen der Ausgabeebene dividiert.

5.2 Schema und Skalierung der Eingabedaten

Die Versuche wurden mit verschiedenen Eingabeschemata durchgeführt. Einige, die häufiger Verwendung finden, werden hier beschrieben und mit einer Bezeichnung versehen, unter der sie im folgenden immer angesprochen werden. Es sind dies:

Schema A:	DAX-Tw	: Tageswert des DAX
	ΔDAX	: Änderung zum Vortag
	Dow Jones-Tw	: Tageswert des Dow Jones Index
	ΔDow Jones	: Änderung zum Vortag
	B-Index	: Tageswert eines Branchenindex
	ΔB-Ind.	: Änderung zum Vortag
	Konk.-Tw	: Tagesnotierung eines Konkurrenzwertes
	$\Delta K - Tw_1$	: Änderung zum Vortag
	$\Delta K - Tw_2$	: Änderung vom Vortag zum Tag davor
	$\Delta K - Tw_{3-4}$	: entsprechend
	Pr-Tw	: Tageswert der zu prognostizierenden Aktie
	$\Delta Pr - Tw_{1-4}$	: entspr. $\Delta K - Tw_{1-4}$
	Pr5-TD	: 5-Tagesdurchschnitt der zu prog. Aktie
Schema B:	Pr-Tw	: Tageswert der zu prognostizierenden Aktie
	$\Delta Pr - Tw_{1-7}$	: wie oben
	$\Delta Pr - Tw_{1,2}$	: Differenz von $\Delta Pr - Tw_1$ und $\Delta Pr - Tw_2$
	$\Delta Pr - Tw_{2,3}$	
	...	
	$\Delta Pr - Tw_{6,7}$	: entsprechend

Schema A ist nur mit den Daten von 1991 anwendbar, da nur hier die nötigen Werte vorhanden sind. Schema B läßt sich für beide Datenmengen verwenden. Die genauen Ausprägungen der Schemata, d.h. die verwendeten Aktien- und Indextitel, werden bei den Prognoseergebnissen beschrieben. Ebenfalls dort erläutert werden Variationen dieser Schemata und solche, die nicht häufiger zur Anwendung kommen.

Die verwendeten Daten wurden nach folgenden Formeln [Schö 90a] auf die Intervalle [0.1,0.9], bzw. [-0.9,0.9] skaliert.

$$Skalierung = \frac{untere\,Intervallgrenze}{maximaler\,Originalwert - minimaler\,Originalwert} \quad (5.3)$$

$$Offset = unt.\,Intervallgrenze - (min.\,Originalwert \cdot Skalierung) \quad (5.4)$$

$$Skalierter\,Wert = Originalwert \cdot Skalierung + Offset \quad (5.5)$$

Die Absolutkurse und die zu prognostizierende Kursänderung wurden auf das Intervall [0.1,0.9] abgebildet, da bei allen Versuchen die Zellen der Ausgabeebene mit der logistischen Aktivierungsfuktion arbeiten. Die anderen Kursänderungen, d.h. solche, die nur als Eingabedaten verwendet werden, wurden auf das Intervall [-0.9,0.9] skaliert.

5.3 Bewertungskriterien

Als Abbruchkriterium für den Lernvorgang gilt das Unterschreiten einer bestimmten Grenze für den durchschnittlichen Fehler, bzw. das Erreichen einer festgelegten Zyklenzahl. Die benötigte Zyklenzahl im ersten Fall, sowie die Höhe des erreichten Fehlers im zweiten, bilden die Maßstäbe für die Beurteilung des Lernaufwandes. Für Versuche mit den Daten von 1991 betrug die Fehlergrenze 0.00010 und die Zyklenzahl 40000, für Versuche mit Daten von 1988 wurden Werte von 0.00400 für den Fehler, und 70000 für die Anzahl der Lernzyklen festgelegt. Nähere Erläuterungen hierzu erfolgen bei der Darstellung der Versuchsergebnisse.

Die Bewertung der Ergebnisse erfolgt nach zwei Kriterien. Zum einen wird der Anteil der tendenziell richtigen Prognosen ermittelt. Hierfür werden die Fälle steigend, unverändert und fallend unterschieden. Da ein unveränderter Kurs, im Vergleich zu den anderen Fällen, nur sehr selten auftritt, im besten Fall in 12% aller Prognosen, wird dafür ein strenger Maßstab angelegt. Diese Prognosen werden nur dann als richtig gewertet, wenn der Prognosewert in einem engen Intervall von [x-0.1 DM,x+0.1 DM] um den tatsächlichen Wert x liegt. In den anderen Fällen genügt eine Änderung vom Ausgangskurs von größer als 0.10 DM in die richtige Richtung, um sie als tendenziell richtig zu bewerten. Die Genauigkeit der Vorhersagen wird mit Hilfe des zweiten Kriteriums bewertet. Es wird die durchschnittliche, betragsmäßige Abweichung zwischen prognostiziertem und tatsächlichem Wert im DM berechnet. Dieser Wert wird natürlich nur für die Vorhersagen von Kurswerten, bzw. Kursänderungswerten angegeben. Wird nur die Tendenz prognostiziert, so gilt sinngemäß nur das erste Kriterium, d.h. um den Ausgabewert 0.5, der unveränderten Kurs angibt, wird ein enges Intervall [0.499,0.501] gelegt, in dem die Prognosen, die unveränderten Kurs aussagen, liegen müssen. Für die Fälle fallend und steigend genügt ein Wert kleiner, bzw. größer dieser Intervallgrenzen.

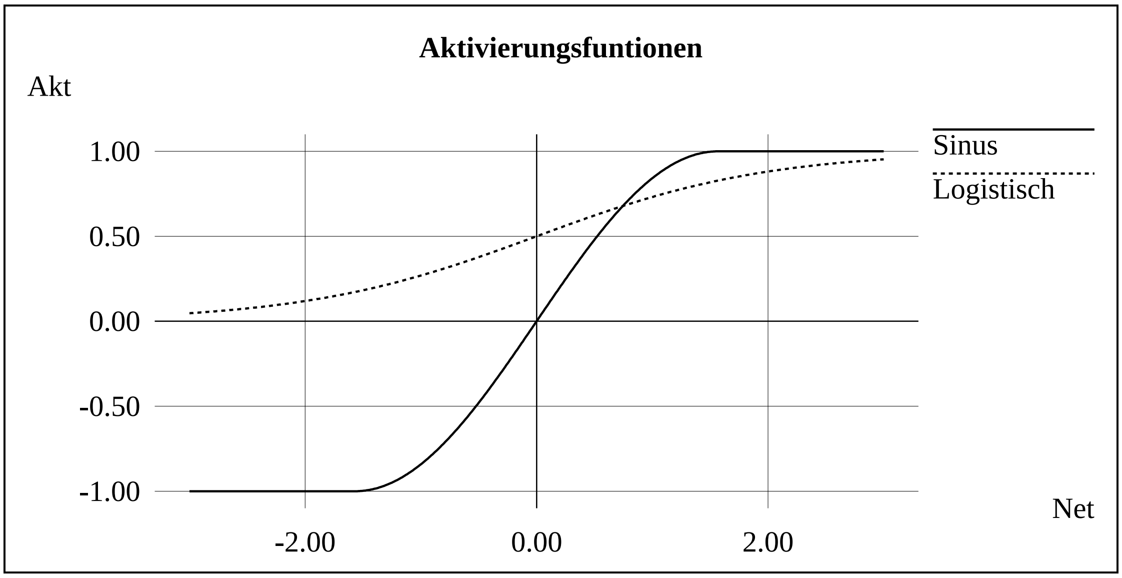


Abbildung 5.1: Aktivierungsfunktionen der Zellen

Kapitel 6

Prognoseversuche

Die Versuchsergebnisse sind in Tabellenform dargestellt. In der Kopfzeile der Tabelle ist die Netztopologie angegeben. Ein Eintrag der Form 17→10→1 beschreibt ein Netz mit 17 Eingabezellen, einer inneren Ebene mit 10 Zellen und einer Ausgabeebene mit einer Zelle. Des weiteren enthält die Kopfzeile die Ausprägung der Lernparameter, die Art der Prognose und die Musterauteilung in Lern- und Testmustermenge. Dabei beschreiben die Zusätze a und b die ursprüngliche Reihenfolge der Muster. Weitere Erklärungen erfolgen bei der Darstellung der Ergebnisse. Die Spalte *Quote* gibt den prozentualen Anteil der tendenziell richtigen Vorhersagen an, die Spalte *d. Abw* enthält die durchschnittliche betragsmäßige Abweichung des Prognoseergebnisses vom tatsächlichen Kurs in DM, die Spalte *rel. Abw* gibt die relative Abweichung an. Letztere wird im Verhältnis zum Mittelwert aus Höchst- und Tiefstkurs angegeben.

6.1 Versuche mit Backpropagation

Zunächst werden Versuche mit den Daten von 1991 betrachtet. In Schema A wird der jeweilige Branchenindex des Commerzbank-Index eingefügt. Für die Prognose der Degussa-Aktie wird als Konkurrenzwert BASF, für die Dresdner Bank die Deutsche Bank und für RWE die VEBA-Aktie verwendet. Es ergeben sich 70 Muster. Untersucht wird zuerst ein dreistufiges Netz mit zwei inneren Ebenen zu je zehn Zellen und einer Ausgabeebene mit einer Zelle. Jede Ebene ist vollständig mit der darauffolgenden Ebene verbunden. Alle Zellen arbeiten mit der logistischen Aktivierungsfunktion und der Identitätsfunktion als Ausgabefunktion. Prognostiziert wird der Absolutkurs. Mit einer Trainingsmenge von 40 Mustern und einer Testmenge von 30 Mustern ergeben sich für eine Lernrate von η=0.2 die Ergebnisse aus Tabelle 6.1.

Die Prognosen der Degussa- und der RWE-Aktie sind durchgehend zu niedrig. Dies trägt zu den hohen betragsmäßigen Abweichungen bei. Für die Prognose der Kursänderung erhält man die Resultate in Tabelle 6.2.

Auffallend ist hier die hohe Abnahme der Lernzyklenzahl gegenüber der Prognose der Absolutkurse. Die Zahl der tendenziell richtigen Vorhersagen bleibt unverändert niedrig. Die Genauigkeit nimmt bei RWE und Degussa zu, ist aber noch verhältnismäßig hoch, so daß keine zufriedenstellenden Ergebnisse erreicht werden. Einzelne Prognosen weisen Abweichungen von bis zu 20 DM auf. Die Wiederholung der Versuche mit einer Lernrate

Netz:17→10→10→1 η:0.2 Prognose:Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	28000	0.00010	0.42064	50.0%	12.75 DM	3.62%
Dresdner Bank	13000	0.00010	0.41549	53.3%	8.65 DM	2.32%
RWE	18900	0.00010	0.69301	53.3%	12.65 DM	3.20%

Tabelle 6.1: Ergebnis 1 Absolutkurs, Standardtitel 1991

Netz:17→10→10→1 η:0.2 Prognose:Kursänderung Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	10600	0.00010	0.84898	50.0%	9.35 DM	2.65%
Dresdner Bank	6700	0.00010	1.53001	50.0%	8.68 DM	2.33%
RWE	9200	0.00010	1.09709	53.3%	9.71 DM	2.46%

Tabelle 6.2: Ergebnis 1 Kursänderung, Standardtitel, 1991

von η=0.3 führt für die Prognose von Absolutkurs, bzw. Kursänderung zu den Ergebnissen der Tabellen 6.3, bzw. 6.4.

Netz:17→10→10→1 η:0.3 Prognose:Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	17100	0.00010	0.18728	63.3%	7.83 DM	2.22%
Dresdner Bank	8600	0.00010	0.35562	53.3%	6.77 DM	1.82%
RWE	16200	0.00010	0.67872	53.3%	11.88 DM	3.00%

Tabelle 6.3: Ergebnis 2 Absolutkurs, Standardtitel, 1991

In beiden Fällen ist, wie zu erwarten, ein Rückgang der Zyklenzahl gegenüber den entsprechenden Versuchen mit η=0.2 zu beobachten. Bis auf eine Ausnahme ergeben sich ähnliche Trefferquoten wie bei den vorigen Prognosen. Die Genauigkeit nimmt leicht zu. Lediglich bei der Degussa-Aktie erhält man deutliche Verbesserungen der Ergebnisse. Auch bei diesen Versuchen werden bei der Prognose der Kursänderungen erheblich weniger Lernzyklen benötigt, als bei der Prognose der Absolutkurse, obwohl die Eingabemuster in beiden Fällen dieselben sind. Der Grund hierfür liegt also in der zu erzeugenden Ausgabe. Die unterschiedliche Lage und Größe des jeweiligen Wertebereichs, die aus Tabelle 4.2 ersichtlich sind, können keine Auswirkungen haben, dies wird durch die Skalierung auf das Intervall [0.1,0.9] ausgeglichen. Die geringere Länge des Wertebereichs der Originaldaten von 24 bei den Kursänderungen, gegenüber ca. 50 bei den Absolutkursen, bedeutet sogar, daß bei einem Lernfehler gleicher Höhe die Genauigkeit der gelernten Muster, übertragen auf die Originaldaten, doppelt so hoch ist. Der Grund für die Abnahme der Lernzyklenzahl liegt also in der unterschiedlichen Aussage des Ausgabewertes, die zu einer anderen Werteverteilung führt. Tritt beispielsweise für drei Eingabemuster mit verschiedenen Tageskursen eine

Netz:17→10→10→1 η:0.3 Prognose:Kursänderung Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	6280	0.00010	0.75400	53.3%	8.93 DM	2.53%
Dresdner Bank	6200	0.00010	1.02764	46.7%	12.77 DM	3.42%
RWE	8500	0.00010	0.82034	53.3%	8.02 DM	2.03%

Tabelle 6.4: Ergebnis 2 Kursänderung, Standardtitel, 1991

Kursänderung um n DM auf, so müssen bei der Prognose von Absolutkursen für diese drei Muster unterschiedliche Ausgaben gelernt werden, während im anderen Fall die Ausgabe immer dieselbe ist.

Vertauscht man die Lern- und Testmustermengen miteinander, so ergeben sich für dasselbe Netz mit einer Lernrate von η=0.3 für die Prognose von Absolutkursen die in Tabelle 6.5 dargestellten Resultate.

Netz:17→10→10→1 η:0.3 Prognose:Absolutkurs Muster:30b/40a						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	17800	0.00010	0.32784	67.5%	9.72 DM	2.76%
Dresdner Bank	11100	0.00010	1.02176	55.0%	12.67 DM	3.40%
RWE	18300	0.00010	0.23912	55.0%	6.35 DM	1.61%

Tabelle 6.5: Ergebnis 3 Absolutkurs, Standardtitel, 1991

Obwohl die Lernmustermenge, statt bisher 40, nur noch 30 Muster beinhaltet, werden mehr Lernzyklen benötigt. Die Degussa-Aktie liefert wiederum das beste Ergebnis hinsichtlich der Zahl der tendenziell richtigen Prognosen, bis auf die Vorhersagen der RWE-Aktie nimmt jedoch die Genauigkeit ab. Die Werte aus Tabelle 6.6 resultieren aus der Prognose der relativen Kursänderungen bei vertauschten Mustermengen.

Netz:17→10→10→1 η:0.3 Prognose: Kursänderung Muster:30b/40a						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	8350	0.00010	0.48035	62.5%	6.03 DM	1.71%
Dresdner Bank	4800	0.00010	0.92880	70.0%	5.52 DM	1.48%
RWE	9900	0.00010	0.76435	57.5%	5.97 DM	1.51%

Tabelle 6.6: Ergebnis 3 Kursänderung, Standardtitel, 1991

Mit den Daten der Dresdner Bank wird das bisher beste Ergebnis aller Prognosen erzielt. Dies gilt sowohl für die tendenziell richtigen Prognosen, als auch für die Genauigkeit der Werte. Abbildung 6.1 veranschaulicht die Ergebnisse der Dresdner Bank graphisch. Es wird allerdings deutlich, daß noch viele Prognosen große Abweichungen aufweisen.

Außerdem erkennt man, daß häufiger ein steigender Kurs prognostiziert wird. Lediglich

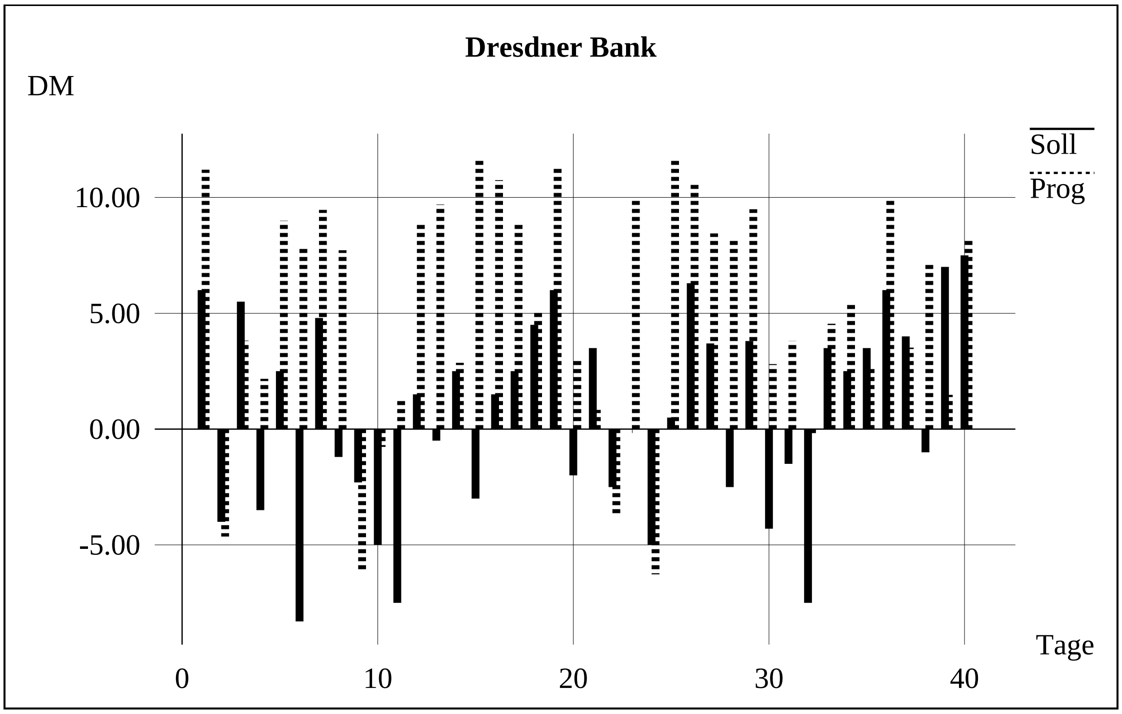


Abbildung 6.1: Ergebnis Kursänderung Dresdner Bank

sechs mal wird ein Fallen des Kurses vorhergesagt, diese Prognosen sind aber dann immer richtig. Anhand dieser Ergebnisse für die Dresdner Bank soll nun an dem trainierten Netz untersucht werden, inwieweit die einzelnen Komponenten des Eingabevektors das Prognoseergebnis beeinflussen. Dazu werden in allen 40 Musterpaaren der Testmenge die jeweils betrachteten Werte auf Null gesetzt und erneut dem Netz zur Prognose präsentiert, welches jedoch mit den kompletten 30 Mustern trainiert wurde. Setzt man die ΔDAX-Werte auf Null, so wird ausnahmslos eine Kurssteigerung zwischen 10 und 14 DM vorhergesagt. Da die Verteilung der Kursänderungen in der Testmenge 22 mal steigend, 17 mal fallend und 1 mal unverändert ist, ergäbe sich eine Trefferrate von 55%, die durchschnittliche Abweichung beträgt aber 11.66 DM. Fehlen nur die Absolutnotierungen des DAX, so ergeben sich ebenfalls, bis auf eine Ausnahme, steigende Prognosen, die Abweichung liegt aber nur noch bei 8.20 DM. Unterschiedliche Auswirkungen erhält man bei fehlenden Dow Jones Index-Werten. Ist der Absolutwert des Index zurückgesetzt, so zeigt das Netz ein Verhalten, das nur geringfügig von dem mit vollständigen Mustern abweicht. Die Trefferquote liegt bei 65%, die Abweichung beträgt 6.15 DM. Lediglich in zwei Fällen wird der anfangs richtig prognostizierte fallende Kurs nicht mehr vorhergesagt. Hier hat der Δ-Wert des Dow Jones-Index größeren Einfluß. Ist er auf Null gesetzt, so werden fast nur noch fallende Prognosen abgegeben, die um durchschnittlich 8.50 DM von der tatsächlichen Kursänderung abweichen. Geringeren Einfluß haben die Werte des Branchenindex und des 5-Tage-Durchschnitts der Absolutkurse des Prognosewertes. Lediglich ein Muster weniger wird richtig erkannt, die Abweichnungen liegen bei 6.44 DM, bzw. 5.98 DM. Auch die Werte von $\Delta Pr - Tw_1$

bis $\Delta Pr - Tw_4$ wirken sich geringer aus, als erwartet. Sind einzelne $\Delta Pr - Tw$-Werte auf Null gesetzt, so ergeben sich folgende Resultate:

	Quote	Abweichung
$\Delta Pr - Tw_1$:	65.0%	5.51 DM
$\Delta Pr - Tw_2$:	62.5%	5.80 DM
$\Delta Pr - Tw_3$:	62.5%	6.13 DM
$\Delta Pr - Tw_4$:	70.0%	5.57 DM

Wenn alle Δ-Werte zusammen Null sind, was nach der Definition der Daten bedeuten würde, daß sich der Kurs in den letzten 5 Tagen nicht geändert hat, so werden trotzdem 26 Muster = 65% richtig klassifiziert, die Abweichung beträgt 6.24 DM.

Da man bei den meisten Versuchen relativ große Abweichungen vom tatsächlichen Kursverlauf feststellt, bietet sich als Alternative, nur die Tendenz der Kursänderung vorherzusagen. Bei fallendem Kurs soll 0, bei unverändertem Kurs 0.5 und bei steigender Tendenz 1 ausgegeben werden. Die Versuche mit diesem Netz erbringen keine Verbesserung in der Zahl der richtigen Prognosen. Vielmehr ist zu beobachten, daß der Lernvorgang in der Mehrzahl der Fälle in einem lokalen Minimum bei einem Fehler von 0.5 endet. Somit werden nur Ergebnisse erreicht, die, bis auf wenige Ausnahmen, unter den bisherigen Prozentzahlen liegen.

Für die folgenden Versuche wird für alle Zellen der zweiten inneren Ebene die Sinusfunktion als Aktivierungsfunktion verwendet. Die Netztopologie bleibt unverändert. Die Prognose der Absolutkurse mit η=0.3 führt zu den Werten in Tabelle 6.7.

Netz:17→10→10(s) →1 η:0.3 Prognose: Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	9600	0.00010	0.44512	53.3%	13.26 DM	3.76%
Dresdner Bank	5000	0.00010	0.40311	46.6%	8.63 DM	2.31%
RWE	7000	0.00010	0.41000	60.0%	9.35 DM	2.36%

Tabelle 6.7: Ergebnis Absolutkurs, Standardtitel, 1991

Es werden erheblich weniger Zyklen benötigt, vgl. Tabelle 6.3, um auf einen Lernfehler derselben Höhe zu gelangen, als bei den entsprechenden Versuchen mit dem Netz ohne Sinus-Aktivierungsfunktion. Am deutlichsten ist der Rückgang bei der RWE-Aktie von 16200 auf 7000 Zyklen festzustellen. Die Entwicklung der Trefferquote und der Genauigkeit sind allerdings uneinheitlich. Bei Degussa und Dresdner Bank ergeben sich Verschlechterungen, während bei RWE leichte Verbesserungen erreicht werden. Ähnliches ist aus den Ergebnissen der Kursänderungsprognose in Tabelle 6.8 ersichtlich.

Auch hier werden weniger Lernzyklen benötigt als mit dem Netz, das ausschließlich die logistische Aktivierungsfunktion verwendet. Der größere Wertebereich der Funktionswerte der Sinusfunktion von [-1,1], gegenüber [0,1] der logistischen Funktion scheint hier eine Rolle zu spielen. Ein weiterer Unterschied der Aktivierungsfunktionen ist der steilere Verlauf der Sinusfunktion zwischen -$\pi/2$ und $\pi/2$, den Punkten, an denen die Grenzwerte -1 und 1 erreicht werden. Die logistische Funktion verläuft hier flacher, d.h. eine Änderung der Eingabe, aus der die Aktivierung berechnet wird, führt bei der Sinus-Aktivierungsfunktion zu

Netz:17→10→10(s) →1 η:0.3 Prognose: Kursänderung Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	4600	0.00010	0.68799	70.0%	8.36 DM	2.37%
Dresdner Bank	2700	0.00010	1.21113	50.0%	7.58 DM	2.03%
RWE	4160	0.00010	0.97819	53.3%	9.37 DM	2.40%

Tabelle 6.8: Ergebnis Kursänderung, Standardtitel, 1991

einer größeren Änderung der Aktivierung als bei der logistischen Funktion. Diese Umstände können es dem Netz ermöglichen, mit weniger Aufwand die Repräsentation der Muster intern zu erstellen. Die Ergebnisse für die RWE- und die Dresdner Bank-Aktie liegen im bisherigen Rahmen. Die Prognose der Kursänderungen der Degussa-Aktie erbringt das prozentual beste Ergebnis, die Genauigkeit ist jedoch, bei einer durchschnittlichen Abweichung von 8.36 DM, recht niedrig. In Abbildung 6.2 werden der tatsächliche Kursverlauf und die Prognoseergebnisse gegenübergestellt. Um die großen Schwankungen zu verdeutlichen, ist die Kursänderung mit den Tageskursen zum Absolutkurs verrechnet worden.

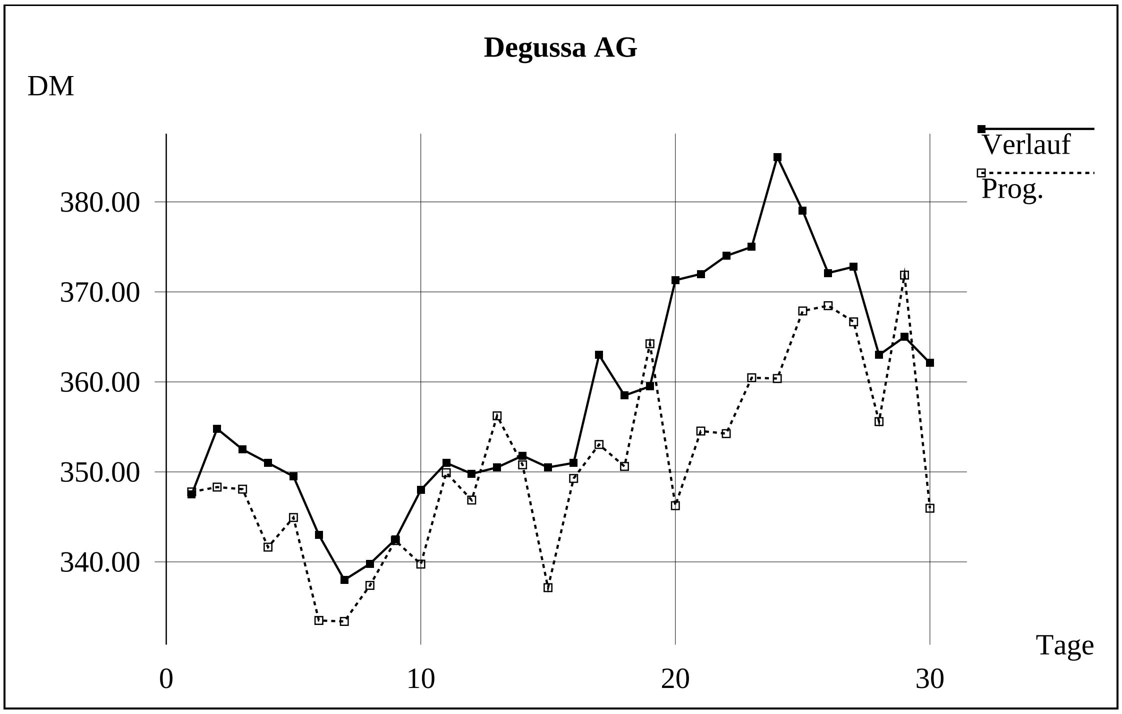


Abbildung 6.2: Ergebnis Kursänderung Degussa AG

Auch mit diesem Netz werden die Versuche mit vertauschten Mustermengen wiederholt. Die Resultate liegen im bekannten Rahmen. In der Regel werden wiederum mehr Zyklen benötigt, als mit der größeren Lernmenge, die Trefferraten liegen zwischen 52.5% und

57.5%. Einzige Ausnahme ist die Prognose der Absolutkurse der Degussa-Aktie. Es wird nach 8000 Lernzyklen ein Testfehler von 0.27824 erreicht. Der Anteil der richtigen Vorhersagen liegt bei 62.5%, die durchchnittliche Abweichung beträgt 7.93 DM, was einer relativen Abweichung von 2.25% entspricht. Die Prognose der Richtung der Kursentwicklung bringt auch mit diesem Netz keine Verbesserung des Ergebnisses. Es ist wieder festzustellen, daß häufig ein lokales Minimum auftritt. Aber auch dort, wo die Mindestfehlergrenze erreicht wird, bleiben die Trefferraten im bekannten Rahmen.

In der zur Verfügung stehenden Datenmenge von 1991 ist für die Aktie der Dresdner Bank, neben den bisher verwendeten Schlußkursen der fortlaufenden Notierung, auch die Kassennotierung vorhanden. Verwendet man diese Daten, so erreicht man mit diesem Netz verbesserte Resultate. Die Lernrate beträgt 0.3, die Lernmustermenge umfaßt 40, die Testmustermenge 30 Musterpaare. Bei der Prognose der Absolutkurse wird nach 5200 Lernzyklen die Fehlergrenze erreicht. Es werden 66.6% der Muster richtig vorhergesagt, die Genauigkeit liegt mit einer durchschnittlichen Abweichung von 8.06 DM allerdings nur im mittleren Bereich. Verwendet man, anstatt dem in diesen Mustern enthaltenen Commerzbank-Index, den Branchen-Index der WestLB, so beträgt die Trefferquote noch 63.3%. Die Abweichung ist mit 8.07 DM unverändert. Es ist ein Lernaufwand von 6000 Zyklen nötig, um die Fehlergrenze zu erreichen. Durch den WestLB-Index ergibt sich also keine Verbesserung der Ergebnisse. Dies untermauert die oben gemachte Beobachtung, daß der Branchen-Index, wider Erwarten, nur geringen Einfluß auf das Ergebnis hat.

Am Beispiel der RWE-Aktie wird ein Netz mit nur einer inneren Ebene mit 14 Zellen untersucht. Diese arbeiten mit der logistischen Aktivierungsfunktion. Es werden der Absolutkurs und die Kursänderung prognostiziert. Für eine Lernrate von η=0.3 ergeben sich folgende Resultate:

Netz:17→14→1 η:0.3 Prognose: Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw
RWE	19800	0.00010	1.09695	53.3%	15.16 DM	3.83%
Prognose: Kursänderung						
RWE	15150	0.00010	1.23167	60.0%	10.06 DM	2.54%

Tabelle 6.9: Ergebnis RWE, 1991

Es werden mehr Lernzyklen benötigt, als bei den Versuchen mit einem Netz mit zwei inneren Ebenen, deren Ergebnisse aus den Tabellen 6.3 und 6.4 zu entnehmen sind. Außerdem ist die Genauigkeit der Prognosen sehr gering. Dies spricht dafür, daß mindestens zwei innere Ebenen erforderlich sind.

Bisher wurden nur Netze mit einer Zelle in der Ausgabeebene betrachtet. Als Varianten davon werden nun Netze untersucht, deren Ausgabeebene aus mehren Zellen besteht. So wird zunächst der Absolutkurs und die relative Kursänderung zusammen in einem Netz vorhergesagt. Verwendet wird das Netz aus Tabelle 6.3, mit dem Unterschied, daß zwei Ausgabezellen vorhanden sind, je eine soll die Prognose der entsprechenden Art ausgeben. Mit einem Lernparameter von η=0.3 werden für die Prognose der Degussa-Kurse 24000 Zyklen benötigt, um die Fehlergrenze zu erreichen. In 53.3% der Fälle wird die Tendenz der Kursentwicklung richtig vorhergesagt. Die Genauigkeit ist jedoch nicht besonders hoch.

Es ist zu beobachten,daß die Tendenz bei beiden Ausgaben einheitlich ist, d.h. es sind entweder beide Prognosen richtig, oder beide falsch. Außerdem stimmt die Kursänderung, die sich aus der Differenz von Tageskurs und vorhergesagtem Absolutkurs ergibt, mit der prognostizierten Kursänderung, die die zweite Ausgabezelle erzeugt, in hohem Maße überein. Nur in 4 von 30 Fällen ergibt sich zwischen den Prognosen eine Abweichung von größer als 2 DM. Eine weitere Variante der bisherigen Prognosen ist die Vorhersage der Tendenz mit verfeinerten Intervallen. Es werden 5 Fälle unterschieden, dies sind stark steigend, steigend, unverändert, fallend und stark fallend. Die absoluten Kursänderungen werden nach ihrer Höhe in die entsprechenden Aussagen umgewandelt. Die Ausgabeebene enthält nun 5 Zellen, die jeweils dann aktiviert werden sollen, wenn der ihnen zugeordnete Fall vorhergesagt wird. Das Ergebnis bestimmt die Zelle, deren Aktivierung am höchsten ist. Untersucht wird die Netztopologie 17→10→10→5. Nach 18000 Zyklen wird mit einer Trainigsmenge von 30 Mustern ein absoluter Lernfehler von 0.00300 erreicht. Der exakte Fall wird nur bei 35% der Prognosen erkannt.

Im Eingabeschema A sind neben den Kursdaten noch andere Daten enthalten, die das Börsengeschehen beschreiben. An Schema B wird untersucht, inwieweit die zukünftige Kursentwicklung allein anhand des bisherigen Kusrverlaufs des Prognosetitels bestimmt werden kann. Dazu werden die Kursänderungen der letzten 7 Tage, sowie die Differenzen zwischen zwei aufeinanderfolgenden Änderungen betrachtet. Letztere bilden die zweite Ableitung der Kurskurve. Es wird wiederum ein Netz mit zwei inneren Schichten, zu je 10 Zellen, verwendet. Wegen es längeren Rückschauzeitraumes, ergeben sich für die Daten von 1991 noch 67 Muster. Die Prognosen werden mit der Degussa-Aktie durchgeführt. Eine Lernrate von η=0.2 führt zu den folgenden Resultaten in Tabelle 6.10.

Netz:14→10→10→1 η:0.2 Prognose: Absolutkurs Muster:37a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	40000	0.00049	0.19939	53.3%	9.25 DM	2.62%
Muster: 30b/37a						
Degussa	40000	0.00755	0.26248	56.8%	8.70 DM	2.47%

Tabelle 6.10: Ergebnis Degussa Absolutkurs, Schema B, 1991

Das auffallendste Ergebnis ist dies, daß mehr Lernaufwand erforderlich ist. Außerdem ist die Trefferrate deutlich geringer, als die besten Resultate für die Aktie der Degussa AG. Ähnliches ergibt sich auch bei der Wiederholung der Versuche zur Prognose der Kursänderung (Tabelle 6.11).

Die Genauigkeit von beiden Prognosearten liegt im unteren Bereich. Ausführlicher wird dieses Eingabeschema mit den Daten von 1988 untersucht. In dieser Menge sind die bisher verwendeten Indices nicht enthalten, sodaß auch noch andere Varianten getestet werden. Neben den drei Standard-Titeln werden einzelne Versuche auch mit drei weiteren Aktientiteln durchgeführt. Dies sind Hoechst, Badenwerk und BHF-Bank. Aus den 247 Handelstagen ergeben sich 239 Musterpaare, die zunächst in zwei Mengen von 139, bzw. 100 Musterpaaren aufgeteilt werden. Die ersten Versuche werden mit dem oben verwendeten Netz der Topologie 14→10→10→1 durchgeführt.Prognostiziert werden die Absolutkurse der drei zusätzlichen Aktientitel. Die Ergebnisse sind Tabelle 6.12 zu entnehmen.

Netz:14→10→10→1 η:0.2 Prognose: Kursänderung Muster:37a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	32700	0.00010	0.63953	56.6%	8.02 DM	2.27%
Muster: 30b/37a						
Degussa	40000	0.00039	1.45457	59.5%	11.25 DM	3.18%

Tabelle 6.11: Ergebnis Degussa, Kursänderung, Schema B, 1991

Netz:14→10→10→1 η:0.2 Prognose: Absolutkurs Muster:139a/100b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Badenwerk	70000	0.01085	0.41565	53%	2.18 DM	1.21%
BHF-Bank	70000	0.01852	0.05675	48%	5.53 DM	1.56%
Hoechst	70000	0.00874	2.05745	46%	12.58 DM	4.57%

Tabelle 6.12: Ergebnis zus. Titel, Absolutkurs, 1988

Zunächst ist festzustellen, daß in keinem Fall die vorgegebene Fehlergrenze von 0.00400 erreicht wird. Der im Vergleich zu den Versuchen mit den Daten von 1991 hohe Lernfehler, ist einerseits natürlich auf die wesentlich größere Anzahl von Lernmustern zurückzuführen, zum anderen wurde auch schon bei den letzten Versuchen mit Schema B, mit den Daten von 1991, beobachtet, daß die Konvergenz des Lernprozesses wesentlich langsamer verläuft, als mit Schema A. Um den Lernfehler zu verringern könnten zusätzliche Lernzyklen angesetzt, bzw. eine höhere Lernrate gewählt werden. Bei der Ausführung dieser Alternativen wurde jedoch deutlich, daß ein immenser Aufwand an Zyklen erforderlich wird, um den Lernfehler noch wesentlich zu verbessern. Eine Erhöhung der Lernrate η führt auch nicht in jedem Fall zu einer Verringerung des Aufwandes, bzw. des Lerfehlers. Bei vielen Versuchen traten im Konvergenzprozeß Sprünge auf, die ein anfangs besseres Konvergenzverhalten wieder zunichte machten. Zudem wurde festgestellt, daß die Höhe des Lernfehlers nicht unbedingt ein Maß für die Güte des Prognoseergebnisses ist. Durch zusätzliche Lernzyklen verminderte sich zwar immer der Lernfehler, doch traten im Verhalten des Testfehlers und der Trefferrate die unterschiedlichsten Entwicklungen auf. In einem Teil der Versuche wurde tatsächlich eine Verbesserung des Prognoseergebnisses erreicht, in anderen Fällen verschlechterten sich der Testfehler und die Trefferrate. Durch Ausführung von weiteren Zyklen kehrte sich vereinzelt die anfangs beobachtete Entwicklung um, dies teilweise sogar mehrmals, sodaß keine einheitliche Aussage über das Verhalten getroffen werden kann. Aus diesen Gründen wurde letztendlich die Zyklenzahl und die Fehlergrenze beibehalten. In der Regel wurden die Versuche mit einer Lernrate von η=0.2 durchgeführt. Beispiele für oben beschriebenes Verhalten werden im folgenden noch angeführt.

Die in Tabelle 6.12 beschriebenen Versuchsergebnisse weisen, mit Ausnahme der Hoechst-Aktie, eine relativ zu den bisherigen Ergebnissen der anderen Titel hohe Genauigkeit auf. Da für diese Titel für 1991 keine Daten vorliegen, sind keine direkten Vergleichswerte vorhanden. Es treten allerdings Trefferraten von unter 50% auf. Es sei jedoch daran erinnert, daß eine dreiwertige Prognose durchgeführt wird. So beträgt bei den Kursen der BHF-Bank der Anteil der Kursentwicklung 'unverändert' in der Testmenge beispielsweise 10%. Ein gleichwahrscheinliches Raten der drei Fälle 'steigend','unverändert' und 'fallend' ergäbe nur eine Trefferqoute von 33.3%, selbst mit einer Einschränkung auf die Fälle 'steigend' und 'fallend' ist nur ein Ergebnis von 45% erreichbar. Die sehr genauen Prognoseergebnisse der Badenwerk-Aktie sind in Abbildung 6.3 dargestellt. Hier erkennt man auch, daß der Kurs dieser Aktie nur geringen Schwankungen unterliegt und in einem sehr engen Wertebereich verläuft. Dadurch läßt sich die hohe Genauigkeit der Prognosen teilweise erklären.

Eine Vertauschung von Lern- und Testmustermenge bei der Hoechst-Aktie führt zu einer Trefferrate von 54%, die Abweichung liegt jedoch mit 10.50 immer noch relativ hoch.

Betrachtet wird im folgenden ein Netz mit drei inneren Schichten zu je acht Zellen. Die Aktivierungsfunktion aller Zellen ist die logistische Funktion. Die Lernrate beträgt η=0.2. Für die Prognose der Absolutkurse der Degussa- und der Hoechst-Aktie erhält man folgende Resultate.

Netz:14→8→8→8→1 η:0.2 Prognose: Absolutkurs Muster:139a/100b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	70000	0.03575	0.07849	48%	5.11 DM	1.56%
Hoechst	70000	0.01103	0.85446	48%	10.05 DM	3.65%

Tabelle 6.13: Ergebnis Degussa u. Hoechst, Absolutkurs, 1988

Für die Degussa-Aktie wird, hinsichtlich der Genauigkeit, das bisher beste Ergebnis erzielt. Eine weitere Verbesserung der Ergebnisse läßt sich durch Verwendung der Sinus-Aktivierungsfunktion für die Zellen der dritten inneren Ebene erreichen (Tabelle 6.14).

Netz:14→8→8→8(s)→1 η:0.2 Prognose: Absolutkurs Muster:139a/100b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	70000	0.02497	0.03938	57%	3.55 DM	1.08%
Dresdner Bank	70000	0.01187	0.58221	46%	10.64 DM	4.01%
RWE	70000	0.00589	0.78948	57%	7.21 DM	3.34%
Badenwerk	70000	0.01004	0.36184	51%	1.96 DM	1.09%
BHF-Bank	70000	0.01364	0.05190	47%	5.19 DM	1.45%
Hoechst	70000	0.01440	0.68573	58%	8.81 DM	3.20%

Tabelle 6.14: Ergebnis Gesamttitel, Absolutkurs, 1988

Auffallend ist wiederum die hohe Genauigkeit der Prognosen der Degussa- und der Badenwerk-Aktie. Bei allen bisherigen Versuchen wurden nur höhere relative Abweichungen erreicht. Dabei ist das Ergebnis der Degussa-Aktie höher zu bewerten. In Tabelle 4.1 ist aufgeführt, daß die Degussakurse zwischen 268 und 386 DM verlaufen, das ist ein Intervall der Länge 118, während sich die Kurse der Badenwerk-Aktie nur in einem Intervall der Länge 22, zwischen 169 und 191, bewegen. In gleicher Weise verhält es sich mit den

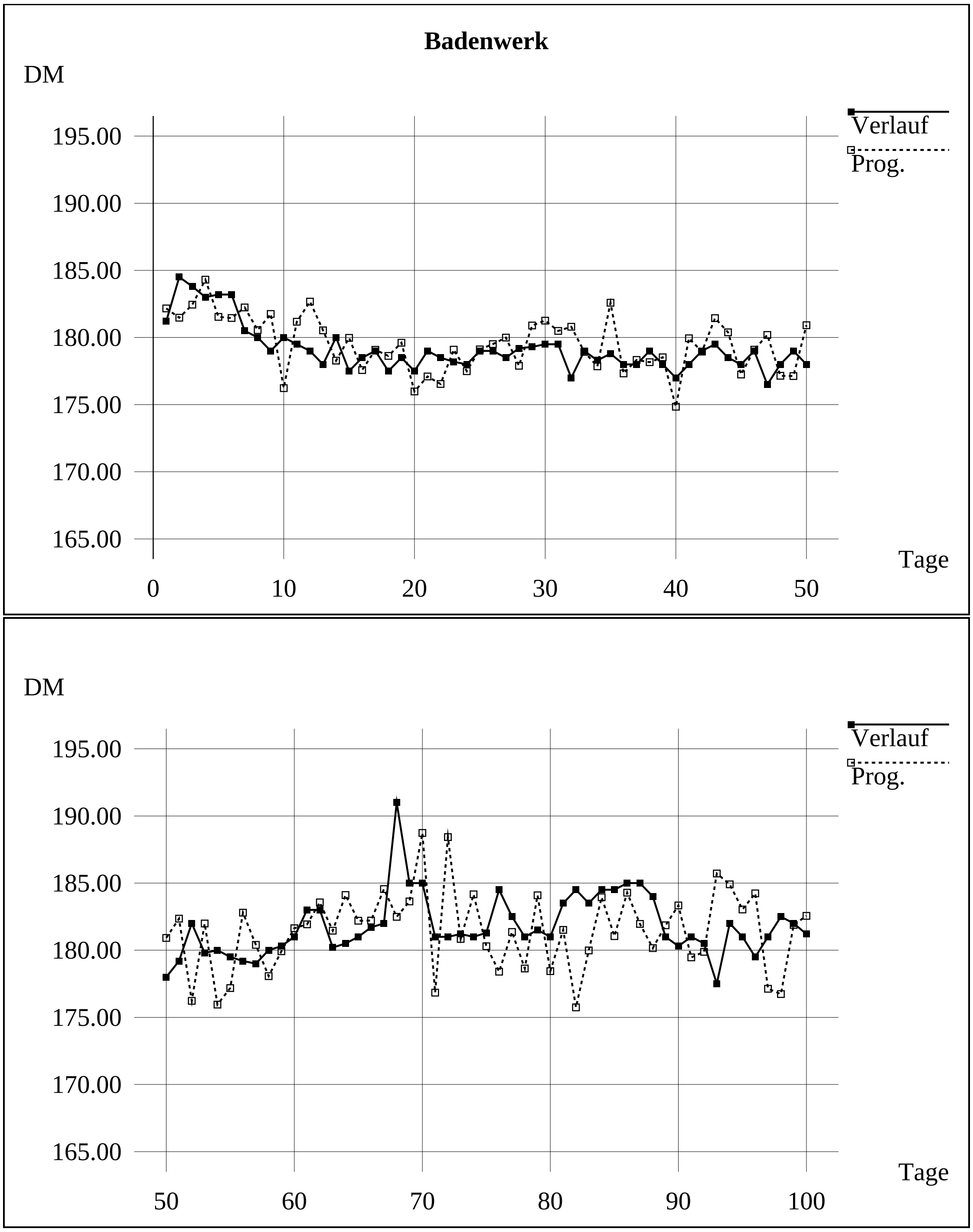


Abbildung 6.3: Prognoseergebnis Badenwerk aus Tabelle 6.12

Kursänderungen der beiden Titel. Die Degussa-Aktie weist die größeren Änderungen auf. Abbildung 6.4 im Vergleich zu Abbildung 6.3 verdeutlicht diesen Umstand. Die Prognosekurve paßt sich auch größeren Schwankungen des Kursverlaufs an.

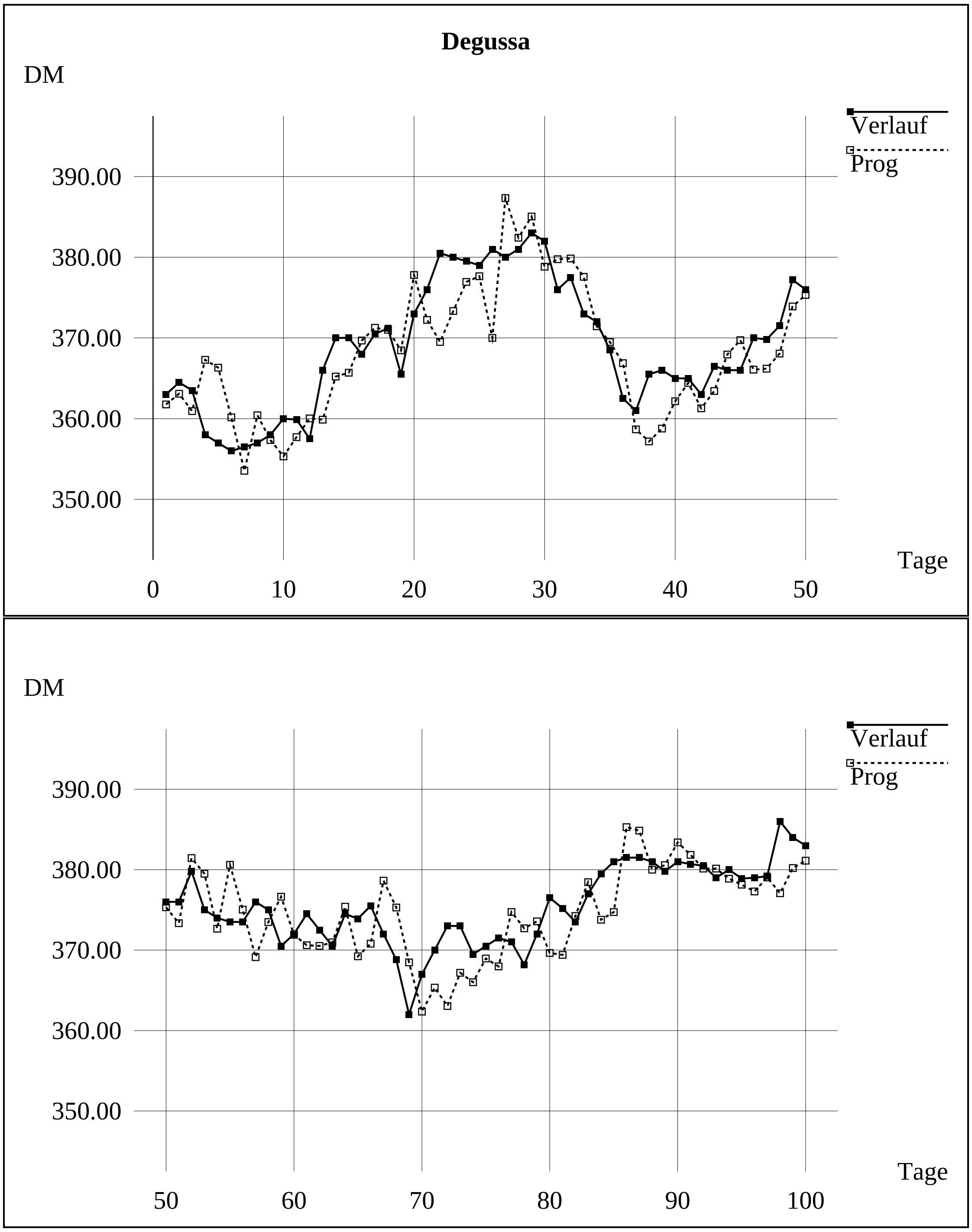


Abbildung 6.4: Prognoseergebnis Degussa aus Tabelle 6.14

Die Abbildung erweckt den Eindruck, daß die Prognosekurve dem tatsächlichen Kursverlauf immer um einige Tage verschoben folgt. Zur Überprüfung werden die Prognoseergebnisse gegenüber den tatsächlichen Ergebnissen jeweils um n Tage verschoben, d.h. die Prognose von Tag x wäre schon am Tag x-n gemacht worden. Eine Verschiebung um einen Tag ergibt keine Verbesserung, es würden nur 56 von 99 Mustern richtig prognostiziert werden. Durch die Verschiebung fehlt die letzte Prognose. Bei zwei Tagen wären 62 von 98 Prognosen = 63.3%, bei drei Tagen 67 von 97 Prognosen = 69% richtig. Danach nimmt die Trefferquote wieder ab. Würde die Prognosekurve tatsächlich nur dem wahren Kursverlauf hinterher hinken, so müßten sich hier höhere Trefferraten ergeben.

Aus den Ergebnisseen der Hoechst-Aktie in den Tabellen 6.12, 6.13 und 6.14 ist abzulesen, daß in diesem Fall ein höherer Lernfehler jeweils zu einem geringeren Testfehler führt, während bei den Ergebnissen der Degussa-Aktie in Tabelle 6.13 und 6.14 die umgekehrte Beziehung besteht. Diese Beobachtungen gelten jedoch nur für diesen Fall und sind nicht etwa für diese Titel charakteristisch.

Der Rückschauzeitraum von 7 Tagen in Schema B wird nun auf 5 Tage verkürzt. Es ergeben sich so nur noch 10 Eingabewerte. Das Netz Tabelle 6.14, mit entsprechend angepaßter Eingabeebene, liefert für die Degussa-Aktie eine Trefferquote von 53%. Der Lernfehler beträgt 0.04170, der Testfehler 0.07404. Die durchschnittliche Abweichung ist 4.86 DM, was einer relativen Abweichung von 1.49% entspricht. Für die Dresdner Bank-Aktie wird eine Trefferrate von 49% erreicht. Der Lernfehler beträgt hier 0.01878. Mit einem Testfehler von 0.50022 wird eine Abweichung von 9.85 oder 3.72% erreicht. In beiden Fällen ist der Lernfehler also höher als bei den Versuchen mit längerem Rückschauzeitraum. Verwendet man ein Netz mit nur zwei inneren Ebenen mit je zehn Zellen, so sind ähnliche Beobachtungen zu machen. Für die Degussa-Aktie wird ein Lernfehler von 0.05324 erreicht. Der Testfehler beträgt 0.05783, was zu einer Trefferrate von 54% führt. Die durchschnittliche Abweichung ist 4.42 DM = 1.35%. Verkürzt man den Rückschauzeitraum weiter, so nimmt die Trefferquote für die Degussa-Aktie weiter ab. Der Lernfehler nimmt zu. So ergibt ein Rückschauzeitraum von 3 Tagen eine Trefferrate von 52%. Der Lernfehler beträgt 0.06186, der Testfehler ist 0.05908. Bei einer durchschnittlichen Abweichung von 4.48 DM = 1.38%, ist die Genauigkeit jedoch nahezu unverändert.

Für die folgenden Versuche werden die Mustermengen weiter unterteilt. Die Menge mit 139 Mustern wird in Mengen von 70 und 69, die Menge von 100 Mustern wird in zwei mal 50 Muster unterteilt. Die Reihenfolge der Mengen wird mit dem Zusatz a,b,c und d angegeben. Eingabeschema ist Schema B. Die Fehlergrenze ist jetzt wieder 0.00010, wegen der verringerten Lernmustermengen. Die genaueren Prognoseergebnisse werden erreicht, wenn die Testmenge eine der 50-elementigen Mengen ist. So ergeben sich für ein Netz mit zwei inneren Ebenen, wobei die Zellen der zweiten Ebene die Sinus-Aktivierungsfunktion besitzen, die Resultate in Tabelle 6.15.

Mit den anderen Lern-/Testmusterkombinationen erhält man nur vereinzelt gute Resultate. So führt die Kombination 70a/69b bei der Degussa-Aktie mit demselben Netz zu einer Trefferrate von 59.4%. Für die Dresdner Bank ergibt die Aufteilung 50c/69b eine Quote von 53.6%. Das Netz aus Tabelle 6.14 liefert für RWE mit der Kombination 69b/50c eine Trefferquote von 62%. Die Abweichung dieser Prognose liegt bei 6.10 DM = 2.82%. Bei der Prognose der Degussa-Aktie soll hier die Möglichkeit der Verbesserung der Ergebnisses durch Ausführung zusätzlicher Lernzyklen geprüft werden. Für die Kombination 50c/50d ergibt sich nach 70000 Lernzyklen ein Lernfehler von 0.00475. Die Trefferrate beträgt 52%, der Testfehler ist 0.00863. Die Abweichung ist 2.28 DM = 0.70%. Nach zusätzlichen 20000

Netz:14→10→10(s)→1 η:0.2 Prognose: Absolutkurs Muster:50c/50d						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	70000	0.00197	0.02366	60%	3.20 DM	0.98%
Dresdner Bank	70000	0.00017	0.13420	60%	6.38 DM	2.41%
RWE	70000	0.00168	0.09104	50%	3.70DM	1.71%

Tabelle 6.15: Ergebnis Standardtitel, Absolutkurs, 1988

Zyklen beträgt der Lernfehler noch 0.00280. Der Testfehler hat sich jedoch auf 0.01042 erhöht. Die Trefferrate beträgt weiterhin 52%, es werden jedoch nicht mehr genau dieselben Muster richtig erkannt. Drei neue Muster werden richtig prognostiziert, während drei der bisher richtigen Prognosen nicht mehr erkannt werden. Nach weiteren 30000 Lernzyklen sinkt der Lernfehler nur noch auf0.00209. Der Testfehler hat sich weiter erhöht und liegt jetzt bei 0.01341. Die Trefferquote ist jetzt 56%, wobei drei neue Muster erkannt werden. Weitere 40000 Lernzyklen sind nötig, um den Lernfehler auf 0.00179 zu reduzieren. Der Testfehler beträgt jetzt 0.01711, der Anteil der richtig prognostizierten Kursentwicklungen ist jetzt wieder 52%. Die durchschnittliche Abweichung ist inzwischen auf 3.10 DM gestiegen. Es sind also 90000 Zyklen nötig, um den Lernfehler noch von 0.00475 auf 0.00179 zu reduzieren, wobei die Lernmustermenge hier nur 50 Muster umfaßt. Bei größeren Mustermengen ist der Aufwand noch höher. Eine Verbesserung des Prognoseergebnisses ist dadurch aber nicht garantiert. Hingewiesen sie noch auf die geringen Abweichungen, die hier für die Degussa-Aktie erreicht wurden. Eine Verkürzung der Zeiträume wird auch bei den nächsten Versuchen vorgenomme. Die Lernmustermenge beinhaltet 50 Muster. Es werden jeweils 20 Muster prognostiziert, dann wird das Lernfenster um diese 20 Muste weitergeschoben, und das Netz wird neu trainiert. So werden 5 Lernvorgänge nötig, um die letzten 100 Muster von 1988 abzuarbeiten. In der Summe ergibt sich für die Degussa-Aktie eine Trefferquote von 57%. Das liegt im Bereich der bisherigen Ergebnisse mit Schema B. In einzelnen Intervallen waren 14 von 20 Prognosen richtig. Diese guten Resultate wurden jedoch durch schlechtere Trefferquoten in anderen Intervallen wieder aufgehoben. Da diese Versuche aus Zeitgründen nur am Rande betrachtet wurden, ist hier noch ein großes Potential an weiteren Varianten hinsichtlich Länge der Zeiträume, Wahl der Parameter und Zyklenzahl vorhanden. Abschließend sei noch der Versuch betrachtet, mit den Daten von 1988 das Schema A, das mit entsprechend umfangreichen Parametern von 1991 prozentual bessere Ergebnisse liefert, in groben Zügen nachzuvollziehen. Dazu wird der Tageswert des Prognosetitels und eines Kokurrenzwertes sowie jeweils 5 Δ-Werte in ein Eingabeschema zusammengefaßt. Ergänzt wird dieses noch durch den Tageswert und die Änderung zum Vortag(Δ-Wert) des FAZ-Index. Dieser Versuch liefert keine befriedigenden Ergebnisse. Die Trefferquoten liegen bei 50%. In einen Fall wird bei der Degussa-Aktie eine durchschnittliche Abweichung von über 19 DM festgestellt.

Mit dem Backpropagation-Verfahren werden Trefferraten von bis zu 70% erzielt. Diese werden für die Prognose der Kursänderungen der Degussa- und der Dresdner Bank-Aktie mit den Daten von 1991 erreicht. Verwendet wird in beiden Fällen ein Netz der Topologie 17→10→10→1. Im zweiten Fall wird für die zweite inneren Ebene die Sinus-Aktivierungsfunktion verwendet. Die RWE-Aktie liefert ebenfalls mit diesem Netz ihr bestes Ergebnis. Dieses ist eine Trefferrate von 60%. Die relativen Abweichungen der besten Resultate liegen zwischen 1.5% und 2.8%. Mit den Daten von 1988 werden keine so hohen Trefferraten erzielt. Dies liegt vorallem an dem anderen Eingabeschema der Daten. Die besten Ergebnisse liegen hier bei 62%. Diese Trefferquoten werden jedoch nur erzielt, wenn mit kleinen Lern- und Testmustermengen gearbeitet wird, die 50 Muster umfassen. Es hat sich ein Netz der Form 14→8→8→8(s)→1 als geeignet erwiesen. Die Genauigkeit ist jedoch in Einzelfällen höher, als bei den Versuchen mit Schema A. Abschließend sei noch darauf hingewiesen, daß das Verhalten nicht nur zwischen den einzelnen Aktientiteln recht unterschiedlich ist, sondern auch bei derselben Aktie in verschiedenen Zeiträumen Unterschiede im Lern- und Prognoseverhalten auftreten.

6.2 Versuche mit Quickprop

Beim Quickprop-Verfahren müssen die Werte von drei Lernparametern gewählt werden. Für den Parameter ν, ein Verzögerungsfaktor der verhindern soll, daß zu große Gewichte errechnet werden, genügt ein kleiner Wert. Er wird bei allen Versuchen konstant auf 0.0001 gesetzt. Des weiteren muß die Lernrate η und der Faktor μ, der die maximale Gewichtsänderung bestimmt gewählt werden. Zunächst wird wieder ein Netz mit zwei inneren Ebenen zu je 10 Zellen untersucht. Es wird Schema A verwendet, die Ausprägung für die einzelnen Titel ist dieselbe wie bei den Backpropagation-Versuchen. Für die Degussa-Aktie ergibt sich mit η=0.2 und einem Wachstumsfaktor μ=1.2 folgendes Ergebnis. Die Fehlergrenze wird nach 16100 Zyklen erreicht, der Testfehler beträgt dann 0.48930. Es werden 53.3% der Kursentwicklungen richtig vorhergesagt. Die durchschnittliche Abweichung ist mit 14.56 DM = 4.13% sehr hoch. Vertauscht man Lern- und Testmustermenge, so sind 19600 Lernzyklen nötig. Es ergibt sich ein Testfehler von 0.29100. Der Anteil der tendenziell richtigen Prognosen beträgt 60%, die Abweichung 12.14 DM. Auch hier werden trotz kleinerer Lernmustermenge mehr Zyklen benötigt. Mit den Kursen der Dresdner Bank werden nur 6300 Zyklen benötigt, der Testfehler ist dann 0.46791. Die Trefferquote ist 60%. Erhöht man η auf 0.3, so sind unterschiedliche Entwicklungen zu beobachten. Bei der Degussa-Aktie wird mit der ursprünglichen Reihenfolge der Muster 40a/30b eine leichte Verminderung der Zyklenzahl erreicht. Sie beträgt nun 14000. Eine Trefferquote von 56.6% resultiert aus einem Testfehler von 0.18028. Bei vertauschten Mustermengen ist jetzt ein höherer Lernaufwand von 21300 Zyklen nötig, um die Fehlergrenze von 0.00010 zu erreichen. Der Testfehler hat dann eine Höhe von 0.24419, bei einer Trefferquote von 62.5%. Für die anderen Standardtitel ergeben sich die folgenden Resultate, die Degussa-Werte sind wegen der Übersichtlichkeit nochmals aufgeführt.

Ein Vergleich mit den Backpropagation-Ergebnissen aus Tabelle 6.3 ergibt Unterschiedliches. Bei der Dresdner Bank-Aktie verringert sich die Zahl der Lernzyklen um ein Drittel, bei der Degussa-Aktie ist der Rückang immerhin noch ca. 1/5. Der Lernaufwand bei der RWE-Aktie steigt von 16200 Zyklen auf 25000 Zyklen an. Auch eine Erhöhung von η auf 0.4 bringt keine Verbesserung gegenüber dem Backpropagation-Versuch. Der Wachstumsfaktor μ hat, wie später noch gezeigt wird, meist geringeren Einfluß auf den Lernaufwand. Aus dem Verlauf des Fehlers in Abhängigkeit von der Zyklenzahl bei RWE erkennt man, daß die Konvergenz beim Quickprop-Verfahren anfangs schneller verläuft, als beim Backpropagation-Verfahren. So ist beispielsweise bei dem Quickprop-Versuch, ausgehend von einem Fehler

Netz:17→10→10→1 η:0.3 μ:1.2 Prognose:Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	14000	0.00010	0.18028	56.6%	8.22 DM	2.33%
Dresdner Bank	5700	0.00010	0.31423	63.3%	7.00 DM	1.88%
RWE	25000	0.00010	1.33519	56.6%	17.20 DM	4.36%

Tabelle 6.16: Ergebnis Absolutkurs, Standardtitel 1991

nach der Initialisierung der Verbindungsgewichte von 1.20477, bereits nach 5000 Zyklen ein Fehler von 0.00169 erreicht. Nach weiteren 5000 Zyklen beträgt der Fehler 0.00032, aber dann werden noch 15000 Zyklen benötigt, um die Fehlergrenze von 0.00010 zu erreichen, wobei immer wieder kleine Sprünge nach oben auftreten. Beim Backpropagation-Versuch beträgt der Fehler nach 5000 Zyklen erst 0.01023, bei einem Ausgangsfehler von 1.45460. Bei 10000 Zyklen ist bereits ein Fehler von 0.00097 erreicht.

Die Prognose der Kursänderungen mit unveränderten Parametern erbringt die Ergebnisse aus Tabelle 6.17.

Netz:17→10→10→1 η:0.3 μ:1.2 Prognose:Kursänderung Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	11000	0.00010	0.59160	60.0%	8.18 DM	2.32%
Dresdner Bank	2700	0.00010	0.68802	56.6%	5.78 DM	1.55%
RWE	13800	0.00010	1.26237	56.6%	10.40 DM	2.63%

Tabelle 6.17: Ergebnis Kursänderung, Standardtitel 1991

Es werden wiederum weniger Zyklen gebraucht, als bei dem Lernvorgang zur Prognose von Absolutkursen, wie auch schon in Abschnitt 6.1 festgestellt wurde. Die Genauigkeit nimmt leicht zu.

Das Lernverhalten ist bei den einzelnen Aktientiteln unterschiedlich. So ergibt ein Versuch mit η=0.4 und μ=1.2 für die Prognose von Absolutkursen bei der Dresdner Bank noch eine Verringerung des Lernaufwandes auf 4700 Zyklen (vgl. Tabelle 6.16). Mit einem Testfehler von 0.29840 werden 60% der Kursentwicklungen richtig vorhergesagt. Bei der RWE-Aktie ist keine Verbesserung erreichbar.

Ein überraschendes Ergebnis erhält man bei der Prognose der Richtung der Kursänderung für die Aktie der Dresdner Bank. Mit einer Testmenge von 40 Mustern wird nach nur 127 Zyklen die Fehlergrenze von 0.00010 erreicht. Der Ausgangsfehler nach der Initialisierung betrug 6.8505. Mit einem Testfehler von 6.55117 werden 53.3% der Tendenzen richtig vorhergesagt. Im Vergleich dazu treten beim Backpropagation-Verfahren bei dieser Prognoseart oft lokale Minima auf, sodaß alle 40000 Zyklen ablaufen.

Nun wird wieder für die Zellen der letzten inneren Schicht die Sinus-Aktivierungsfunktion verwendet. Bei unveränderten Parameterwerten wird zunächst der Absolutkurs prognostiziert. Mit 40 Lern- und 30 Testmustern werden die Ergebnisse aus Tabelle 6.18 erreicht.

Im Gegensatz zu den Backpropagation-Versuchen, wo deutliche Abnahmen des Lernauf-

Netz:17→10→10(s)→1 η:0.3 μ:1.2 Prognose:Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	15450	0.00010	0.17127	66.6%	8.54 DM	2.42%
Dresdner Bank	5200	0.00010	0.42286	60.0%	7.78 DM	2.09%
RWE	11800	0.00010	1.02782	56.6%	15.45 DM	3.91%

Tabelle 6.18: Ergebnis Absolutkurs, Standardtitel 1991

wandes festgestellt wurden, ist hier diese Entwicklung nicht in jedem Fall zu verzeichnen. Bei der Prognose der Degussa-Aktie steigt, gegenüber den Versuchen mit dem Netz ohne Sinus-Aktivierungsfunktion, die Zyklenzahl an. Bei der Dresdner Bank ist nur ein sehr geringer Rückgang von 5700 auf 5200 Zyklen vorhanden. Nur bei der RWE-Aktie ergibt sich eine deutliche Abnahme. Die Genauigkeit der Prognosen liegt im bisherigen Rahmen der jeweiligen Titel. Aus der Prognose der Kursänderungen resultieren die Werte in Tabelle 6.19.

Netz:17→10→10(s)→1 η:0.3 μ:1.2 Prognose:Kursänderung Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	17000	0.00010	0.67250	50.0%	8.98 DM	2.59%
Dresdner Bank	1800	0.00010	0.76872	50.0%	5.82 DM	1.56%
RWE	19730	0.00010	0.94589	56.6%	8.56 DM	2.16%

Tabelle 6.19: Ergebnis Kursänderung, Standardtitel 1991

Mit Ausnahme der Vorhersagen der Kurse der Dresdner Bank, werden jetzt, entgegen der Beobachtungen bei den bisherigen Versuchen, mehr Zyklen für die Prognose der Kursänderungen als für die Prognose der Absolutkurse benötigt. Ein Vergleich der Ergebnisse aus den Tabellen 6.18 und 6.19 mit den Backpropagation-Resultaten aus den Tabellen 6.7 und 6.8 zeigt, daß bei den Quickprop-Versuchen erheblich mehr Lernaufwand nötig ist, als bei den Backpropagation-Versuchen mit demselben Netz. Einzige Ausnahme hiervon ist wiederum die Dresdner Bank-Aktie. Bei dieser Aktie ist in beiden Fällen ein Rückgang der Zyklenzahl festzustellen. Bei der Prognose der anderen Titel ist das Verhalten des Quickprop-Lernvorgangs ähnlich, wie eingangs für die RWE-Aktie beschrieben.

Auch mit diesem Netz ist bei der Prognose der Tendenzen der Kursentwicklung der Dresdner Bank-Aktie ein äußerst geringer Lernaufwand zu beobachten. Die Fehlergrenze wird bereits nach 110 Lernzyklen erreicht. Die Trefferquote beträgt 53.3%. Für die Prognose des Absolutkurses unter Verwendung der Kassennotierungen und des WestLB-Branchenindex erhält man die Ergebnisse in Tabelle 6.20.

Beide Ergebnisse stellen hinsichtlich des Lernaufwandes eine Verbesserung gegenüber den entsprechenden Backpropagation-Versuchen dar.

Auch mit Quickprop wird die Eignung eines Netzes mit einer inneren Ebene untersucht. Es werden wieder die Absolutkurse und die Kursänderungen der RWE-Aktie vorhergesagt. Für ein Netz mit 14 Zellen in der inneren Schicht ergeben sich folgende Resultate:

Netz:17→10→10(s)→1 η:0.3 μ:1.2 Prognose:Absolutkurs Muster:40a/30b						
Kassennotierung						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Dresdner Bank	2700	0.00010	0.44747	60.0%	9.00 DM	2.41%
Kassennotierung + WestLB-Index						
Dresdner Bank	2500	0.00010	0.68642	60.0%	10.57 DM	2.83%

Tabelle 6.20: Ergebnis Dresdner Bank, 1991

Netz:17→14→1 η:0.3 μ:1.2 Prognose:Absolutkurs Muster:40a/30b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
RWE	40000	0.00022	0.83072	53.3%	12.90 DM	3.26%
Prognose: Kursänderung						
RWE	40000	0.00016	0.77857	56.6%	7.02 DM	1.78%

Tabelle 6.21: Ergebnis RWE, 1991

Mit der festgelegten Zyklenzahl kann die Fehlergrenze nicht mehr erreicht werden. Bei der Prognose des Absolutkurses wird nach 20000 Zyklen ein Fehler von 0.00070 erreicht, danach vollzieht sich die Abnahme des Fehlers nur noch in äußerst kleinen Schritten. Im zweiten Fall, bei der Vorhersage der Kursänderungen, ist dieses Verhalten noch extremer ausgeprägt. Bereits nach 10000 Lernzyklen beträgt die Höhe des Fehlers nur noch 0.00028. Die restlichen 30000 Zyklen ergeben nur noch eine Verringerung um 0.00012. Mit erheblich weniger Lernaufwand wäre also ein Ergebnis erreicht worden, das in einem ähnlichen Gütebereich liegen würde wie das jetzige. Der Vergleich mit den Backpropagation-Ergebnissen viele dann auch nicht so schlecht aus, wie das jetzt in vielen Versuchen der Fall ist. Ein Heraufsetzen der Fehlergrenze war jedoch nicht mehr möglich, da zum Zeitpunkt der Quickprop-Versuche die Tests mit Backpropagation schon zum größten Teil abgeschlossen waren. Somit wäre keine einheitliche Beurteilungsgrundlage mehr vorhanden gewesen, da die Backpropagation-Versuche mit einer höheren Fehlergrenze auch andere Ergebnisse erbracht hätten.

Ein Netz mit drei inneren Ebenen zu je acht Zellen, wobei die Zellen der letzten Ebene mit der Sinus-Aktivierungsfunktion arbeiten, wird im folgenden noch für die Prognose der Degussa-Aktie getestet. Mit vertauschten Mustermengen ergeben sich hier die besten Resultate. Für die Prognose der Absolutkurse sind 7600 Zyklen nötig, der Testfehler beträt dann 0.43925. Werden die Kursänderungen vorhergesagt, so ist der Lernaufwand 4500 Zyklen, bei einem Testfehler von 0.48542. In beiden Fällen erhält man eine Trefferrate von 62.5%. Die durchschnittliche Abweichung ist 11.66 DM bei den Absolutkursen und 6.24 DM bei den Kursänderungen.

Der höchste erreichte Anteil von tendenziell richtigen Prognosen liegt bei 70% mit Backpropagation, bzw. 66.6% mit Quickprop. In Tabelle 6.22 sind die richtig erkannten Muster von mehreren Versuchen mit unterschiedlichen Netzen, Prognosearten und Lernverfahren bei der Vorhersage der Werte der Degussa AG mit Schema A aufgeführt. Die Testmenge umfaßt die 30 letzten Muster aus der 70-elementigen Mustermenge von 1991.

Muster	Tab. 6.17	Tab. 6.18	Tab. 6.3	Tab. 6.7	Tab. 6.8
1	*	*	*	*	*
2	*	*		*	*
3	*		*		*
4	*	*	*	*	*
5	*	*	*	*	*
6	*	*	*	*	*
7		*	*	*	*
8					
9	*	*	*		*
10	*		*		
11		*	*		*
12	*	*	*	*	*
13	*	*	*	*	*
14		*			*
15	*	*	*	*	*
16		*			
17	*	*	*		*
18	*		*	*	*
19	*	*	*	*	*
20					
21					
22					
23					
24					
25	*	*	*	*	*
26	*	*	*	*	*
27					
28	*	*	*	*	*
29		*		*	*
30	*	*	*	*	*
	18	20	19	16	21
* = Tendenz richtig erkannt					

Tabelle 6.22: Tendenziell richtig prognostizierte Muster, 1991

Es fällt auf, daß bei den Mustern 8, 20-24 und 27 in keinem der Fälle die Tendenz richtig erkannt wird. Das Vorhersagepotential in Schema A scheint also mit der besten Lösung von 21 richtigen Prognosen fast ausgeschöpft zu sein.

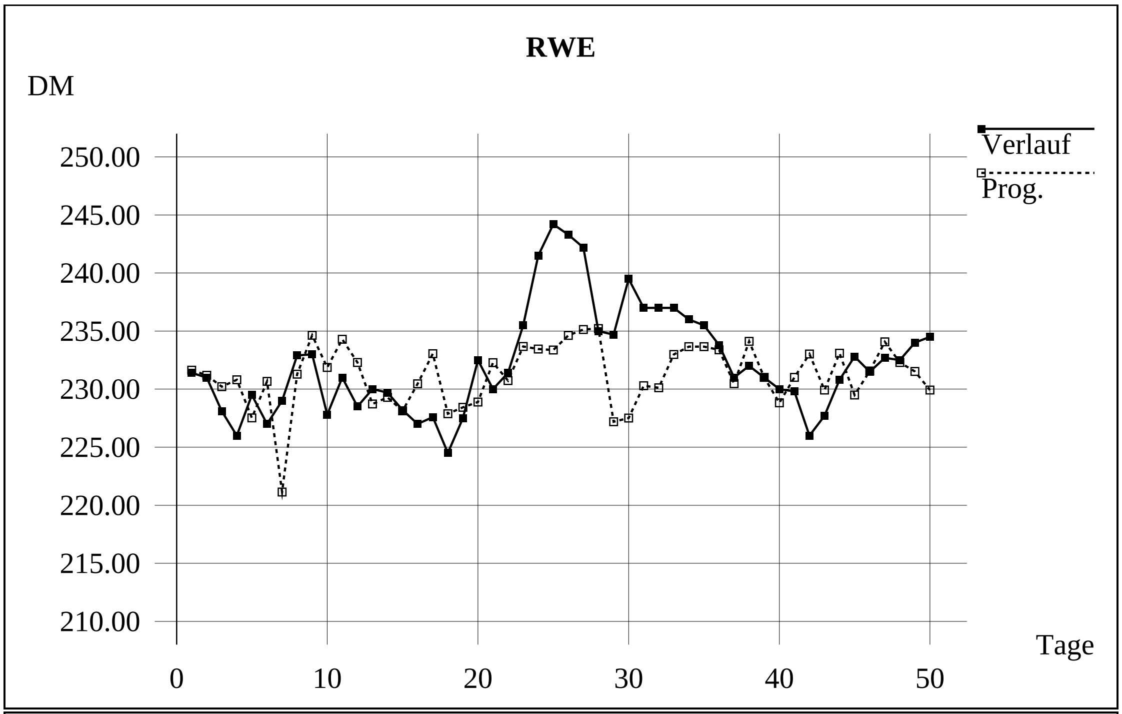
Bei den Versuchen mit den Daten von 1988 gestaltet sich die Wahl der Parameter zunehmend schwieriger. Die einzelnen Aktientitel verhalten sich unterschiedlich. Zunächst wird das Netz mit drei inneren Ebenen zu je acht Zellen betrachtet. Die Aktivierungsfunktion aller Zellen ist die logistische Funktion. Die Eingabedaten sind nach Schema B organisiert. Für die Prognose der Absolutkurse bei der Degussa-Aktie wird mit η=0.3 und μ=1.4 das beste Ergebnis erreicht. Nach 70000 Lernzyklen ergibt sich ein Lernfehler von 0.02505, der Testfehler beträgt 0.12593. Bei einer durchschnittlichen Abweichung von 6.73 DM werden 51% der Kursentwicklungen richtig prognostiziert. Bei der Vorhersage der Kursänderungen führt die Parameterkombination η=0.2/μ=1.3 zwar zu den besten Resultaten was die Trefferrate betrifft, mit einem Wachstumsbegrenzungsfaktor von nur μ=1.1 wird jedoch ein gleich gutes Ergebnis hinsichtlich des Lernaufwandes erzielt. Bei der RWE-Aktie wird bei der Absolutprognose mit η=0.2 und μ=1.1 ein Lernfehler von 0.03345 erreicht. Die Trefferrate beträgt 57%, der Testfehler 0.45481. Mit η=0.2 und μ=1.7 wird ein Lernfehler gleicher Höhe erzielt. Er beträgt dann 0.03275, der Testfehler ist 0.35613 und die Trefferquote 61%. Dieses Ergebnis ist in Abbildung 6.5 dargestellt. Die durchschnittliche Abweichung beträgt 4.75 DM = 2.20%.

Für die Dresdner Bank-Aktie führt auch die Parameterwahl η=0.2/μ=1.7 zum besten Ergebnis. Die Trefferqoute ist dann 50%, der erreichte Lernfehler 0.01101. Eine Erhöhung von η auf 0.3 verschlechtert die Konvergenz des Lernvorgangs. Es treten größere Sprünge auf, der erreichte Lernfehler ist nur noch 0.05462. Arbeiten die Zellen der letzten Ebene mit Sinus-Aktivierungsfunktion, so erhält man die Ergebnisse in Tabelle 6.23.

Netz:14→8→8→8(s)→1 η:0.2 μ:1.5 Prognose:Absolutkurs Muster:139a/100b						
Titel	Zyklen	Lernfehler	Testfehler	Quote	d. Abw.	rel. Abw.
Degussa	70000	0.01787	0.10825	52%	5.94 DM	1.82%
Dresdner Bank	70000	0.04022	2.85295	43%	26.01DM	9.82%
RWE	70000	0.01012	0.68520	52%	6.29 DM	2.91%

Tabelle 6.23: Ergebnis RWE, 1991

Der Vergleich mit den Backpropagation-Ergebnissen für dasselbe Netz fällt unterschiedlich aus. Bei der Degussa-Aktie wird ein kleinerer Lernfehler erreicht, bei den Aktien Dresdner Bank und RWE ist der Lernfehler höher. Was die Prognoseergebnisse betrifft, so werden hier keine so guten Resultate erzielt wie mit Backpropagation. Lediglich bei der Genauigkeit der RWE-Vorhersagen ist eine Verbesserung zu verzeichnen. Eine Verkürzung des Rückschauzeitraumes auf 5 Tage ergibt mit diesem Netz auch keine Verbesserung gegenüber den Backpropagation-Versuchen. Für die Prognose der Degussa-Aktie wird mit der Paramaterausprägung η=0.2, μ=1.3 ein Lernfehler von 0.05499 erreicht. Beim Backpropagation-Versuch betrug der Fehler 0.04170. Der Testfehler beträgt 0.24149. Mit einer durchschnittlichen Abweichung von 9.82 DM = 3.0% werden 52% der Muster richtig prognostiziert. Bessere Resultate erhält man für ein Netz mit zwei inneren Ebenen. Jede Ebene enthält 10 Zellen, die alle mit der logistischen Aktivierungsfunktion arbeiten. Nach 70000 Zyklen ist ein Lernfehler von 0.04157, gegenüber 0.05324 bei Backpropagation, erreicht. Der Testfehler beträgt 0.06446. Die Trefferquote ist 55%. Abschließend wird auch hier der FAZ-Index in die Eingabedaten einbezogen. Schema B wird um den FAZ-Index-Tageswert und die Änderung zum Vortag erweitert. Verwendet wird wieder das Netz mit zwei inneren Ebenen. Die Eingabeebene umfaßt jetzt 16 Zellen. Es ergeben sich 188 Muster, da der FAZ-Index nicht lückenlos vorhanden ist. Die Trainingsmenge beinhaltet 100 Muster. Prognostiziert wer-



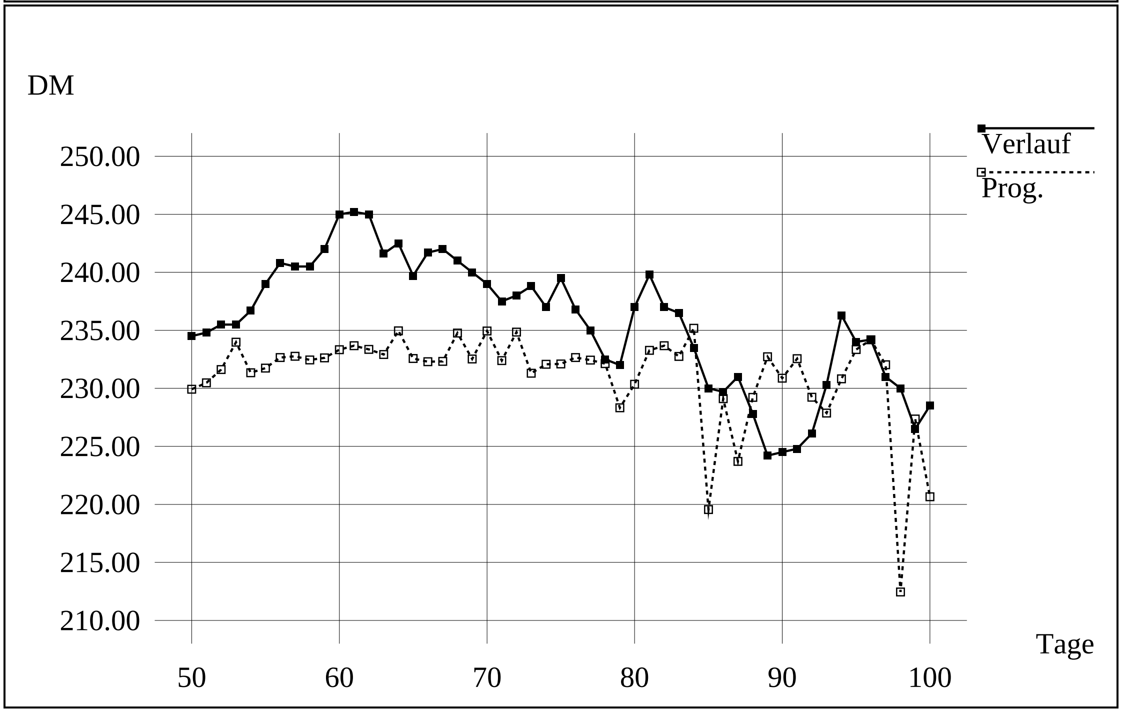


Abbildung 6.5: Prognoseergebnis RWE 1988

den die Absolutkurse der Degussa-Aktie, die Parameterwerte sind η=0.2 und μ=1.2. Der erreichte Lernfehler beträgt 0.00627, der Testfehler ist 0.20559. Der Anteil der richtigen

Prognosen ist 53.4%, die Abweichung 8.49 DM = 2.60%. Ein Netz mit nur einer inneren Ebene liefert schlechtere Ergebnisse. Es wird nur noch ein Lernfehler von 0.00938 erreicht, die Trefferrate sinkt auf 47.7%.

Die wichtigste Erkenntnis der Quickprop-Versuche ist, daß die erhoffte drastische Reduzierung des Lernaufwandes nur vereinzelt eintritt. Es ist von Titel zu Titel verschieden, ob die Fehlerkurve durch eine Parabel mit genügender Genauigkeit approximiert werden kann. Die Genauigkeit der liegt im Rahmen der Bachpropagation-Versuche. In einigen Fällen wird eine höhere Trefferquote erreicht. Als Beispiel sei das Ergebnis der RWE-Aktie genannt. Hier wird eine Trefferrate von 61% erreicht, was das beste Ergebnis mit vollständigen Mustermengen von 1988 darstellt. Es erweisen sich dieselben Netze als geeignet wie bei den Backpropagation-Versuchen.

6.3 Versuche mit Counterpropagation

Ein Counterpropagation-Netzwerk besitzt, wie in Abschnitt 2.2.3 beschrieben, nur eine innere Ebene. Diese Ebene arbeitet nach dem 'winner-takes-all'-Prinzip, d.h. nur eine Zelle erzeugt als Ausgabe eine Eins, alle anderen erzeugen eine Null. Somit bestimmen die Gewichte der Verbindungen von der einen aktivierten Kohonenzelle zur Ausgabeebene die Ausgabe des Netzes. Das bedeutet wiederum, daß nur soviele unterschiedliche Ausgabevektoren erzeugt werden können, wie in der Kohonenebene Zellen vorhanden sind. Die Ausgabe eines kontinuierlichen Wertebereichs, wie es bei den bisherigen Prognosen von Absolutkurs und Kursänderung mit Backpropagation und Quickprop der Fall war, ist somit nicht möglich. Deshalb wird hier nur die Tendenz der Kursänderung prognostiziert. Hierfür muß nur die Ausgabe 0.0, 0.5 oder 1.0 erzeugt werden. Es sind drei Parameter zu wählen. Für den Parameter γ, der die Auswahl eines neuen Kohonenneurons steuert, ist in der Regel ein Standardwert von 4.0 ausreichend. Der Parameter α steuert die Anpassung der Gewichte der Kohonenebene und somit die Stärke, mit der der Gewichtsvektor an den betrachteten Eingabevektor angeglichen wird. Der Parameter β findet bei der Änderung der Gewichte zwischen Kohonen- und Grossbergebene Verwendung. Er bestimmt somit wie stark die Ausgabe einer Grossbergzelle der gewünschten Ausgabe des aktuellen Eingabevektors angepasst wird und dabei evtl. die Ausgabe eines früheren Eingabevektors verändert wird. Es werden erst die Daten von 1991 betrachtet. Das verwendete Eingabeschema ist bei allen Versuchen Schema A. Die Muster werden wiederum in zwei Mengen zu 40 und 30 Mustern aufgeteilt. Es wird ein Netz mit 40 Kohonenzellen untersucht (Tabelle 6.24).

Netz:17→40→1 α:0.9 β:0.9 Prognose: Tendenz Muster:40a/30b				
Titel	Zyklen	Lernfehler	Testfehler	Quote
Degussa	550	0.00010	4.87326	53.3%
Dresdner Bank	450	0.00010	5.47912	53.3%
RWE	450	0.00010	4.36491	50.0%

Tabelle 6.24: Ergebnis Tendenz, Standardtitel 1991

Da genauso viele Kohonenzellen wie Lernmuster vorhanden sind, wird jedes Muster einer neuen Kohonenzelle und somit neuen Klasse zugeordnet. Die Parameter α und β

können so groß wie möglich gewählt werden. Dabei spielt nur β eine Rolle. Setzt man α auf 0.05, so erhält man mit demselben Aufwand das gleiche Ergebnis. Da jede Klasse nur ein Muster enthält, spielt es keine Rolle, ob der Gewichstvektor nun in kleinen oder großen Schritten angepaßt wird. Setzt man jedoch β auf 0.5 und läßt die anderen Parameter unverändert, so werden 1000 Zyklen benötigt, um die Fehlergrenze zu erreichen. Das Erlernen der gewünschten Ausgabe geht nun langsamer voran. Der Fall daß gelerntes Wissen durch zu große Ausprägung von Lernparametern zerstört wird, tritt aus dem oben genannten Grund nicht auf.

Die Anwendung des Counterprogagation-Verfahrens hat zum Ziel, die Eingabe in verschiedene Klassen zu unterteilen. Bisher hat jede Klasse nur ein Element. Eine wünschenswerte Entwicklung ist die, daß mehrere Muster einer Klasse zugeordnet werden und so die Klassifizierung allgemeinere Gültigkeit besitzt. Die Zahl der Kohonenzellen wird deshalb zunächst auf 30 verringert. Vertauscht man Lern- und Testmustermenge, so tritt derselbe Fall ein wie oben. Die Ergebnisse dieser Versuche sind aus Tabelle 6.25 ersichtlich.

Netz:17→30→1 α:0.9 β:0.9 Prognose: Tendenz Muster:30b/40a				
Titel	Zyklen	Lernfehler	Testfehler	Quote
Degussa	430	0.00010	3.51942	67.5%
Dresdner Bank	400	0.00010	4.19654	65.0%
RWE	400	0.00010	4.32417	62.5%

Tabelle 6.25: Ergebnis Tendenz, Standardtitel 1991

Bei der Durchführung der Versuche mit der ursprünglichen Reihenfolge der Muster kommt kein kontinuierlicher Konvergenzprozeß mehr zustande. Während der ersten Zyklen nimmt der Lernfehler noch ab, dann treten große Sprünge auf und nur gelegentlich ergibt sich noch ein geringerer Lernfehler. Dieses Verhalten ist auch durch Variation der Parameter nicht zu ändern. Die Versuche wurden deshalb mit Hilfe eines Progamms durchgeführt, das das Netz beim Auftreten eines neuen niedrigsten Fehlers abspeichert. Nach einer gewissen Zahl von Lernzyklen wurde der Lernvorgang dann abgebrochen. Die Ergebnisse der Versuche sind in Tabelle 6.26 dargestellt. Es wurden jeweils 4000 Zyklen durchgeführt. In der Spalte *kl. Lernfehler* ist der kleinste erreichte Lernfehler angegeben.

Netz:17→30→1 α:0.05 β:0.2 Prognose: Tendenz Muster:40a/30b				
Titel	Zyklen	kl. Lernfehler	Testfehler	Quote
Degussa	4000	1.20572	6.07540	60.0%
Dresdner Bank	4000	1.14362	5.24750	56.6%
RWE	4000	1.60640	3.73608	56.6%

Tabelle 6.26: Ergebnis Tendenz, Standardtitel 1991

Die Ursache für oben beschriebenes Verhalten ist die, daß die Zuordnung der Muster zu den Klassen durch selbstständiges Lernen geschieht. Es werden die Eingabevektoren in einer Klasse zusammengefasst, die im normierten Vektorraum dicht beisammen liegen. Die

Ausgabe des Musters wird dabei nicht berücksichtigt. So können Muster in dieselbe Klasse eingeordnet werden, die eine unterschiedliche Ausgabe haben. Der resultierende Ausgabewert, den die Grossbergebene zu dieser Klasse erzeugt, ist dann ein Wert, der in Richtung eines Mittelwertes der einzelnen Ausgaben geht. Während des Lernvorgangs ändert sich die Klasseneinteilung, was aber nicht heißen muß, daß zwingend eine bessere Lösung gefunden wird. Die resultiernden Ausgaben sind nicht mehr so eindeutig, wie bei den Versuchen in Tabelle 6.24. Im obigen Versuch werden bei der Prognose der Degussa-Aktie die Klassen so gebildet, daß zu vier Lernmustern nicht die korrekte Ausgabe erzeugt wird. Durch weitere Verringerung der Zahl der Kohonenzellen wird nun versucht eine Klassifizierung zu finden, die einen geringeren Fehler erzeugt. Die Anzahl der Kohonenneuronen wird auf 20 reduziert. Es ergeben sich die Resultate in Tabelle 6.27.

Netz:17→20→1 α:0.05 β:0.4 Prognose: Tendenz Muster:40a/30b				
Titel	Zyklen	kl. Lernfehler	Testfehler	Quote
Degussa	4000	2.36378	6.64759	60.0%
Dresdner Bank	4000	1.93508	4.68773	50.0%
RWE	4000	1.64270	5.54533	60.0%

Tabelle 6.27: Ergebnis Tendenz, Standardtitel 1991

Hier werden für die Degussa-Aktie schon 8 von 40 Lernmuster nicht mehr korrekt gelernt. Etwas bessere Resultate erhält man bei Vertauschung der Mustermengen. Jetzt werden nur 3 von 30 Lernmustern falsch klassifiziert. Die Trefferquote bei Degussa und RWE beträgt 62.5%, bei der Dressdner Bank-Aktie werden 55% richtig vorhergesagt. Eine weitere Reduzierung der Zahl der Zellen in der Kohonenebene bringt ähnliche Ergebnisse. Besonders erwähnt sei davon nur das Ergebnis der Degussa-Aktie für ein Netz mit 15 Kohonenneuronen. Es werden 20 von 30 Testmustern = 66.6% richtig prognostiziert. Nocheinmal komplett aufgeführt werden die Ergebnisse, die mit einen Netz mit 5 Kohonenzellen erreicht wurden (Tabelle 6.28).

Netz:17→5→1 α:0.05 β:0.2 Prognose: Tendenz Muster:40a/30b				
Titel	Zyklen	kl. Lernfehler	Testfehler	Quote
Degussa	4000	3.69920	3.71376	56.6%
Dresdner Bank	4000	3.34229	3.58391	66.6%
RWE	4000	3.46066	3.44085	50.0%

Tabelle 6.28: Ergebnis Tendenz, Standardtitel 1991

Für die Mustermengenaufteilung 40a/30b ergibt sich für die Aktie der Dresdner Bank das beste Ergebnis. Generell läßt sich sagen, daß das Verhalten der einzelnen Titel auch hier nicht einheitlich ist.

Die Daten von 1988 werden nach Schema B organisiert. Die 239 Muster werden in Mengen von 139 bzw. 100 Mustern aufgeteilt. Die erste Menge wird als Lernmenge benutzt. Untersucht wird deshalb zunächst ein Netz mit 139 Zellen in der Kohonenebene. Jedes

Muster wird wieder einem neuen Kohonenneuron zugeordnet. Die Prognoseergebnisse sind Tabelle 6.29 zu entnehmen.

Netz:14→139→1 α:0.9 β:0.9 Prognose: Tendenz Muster:139a/100b				
Titel	Zyklen	kl. Lernfehler	Testfehler	Quote
Degussa	700	0.00010	18.23612	52.0%
Dresdner Bank	740	0.00010	22.81733	49.0%
RWE	720	0.00010	13.47041	52.0%

Tabelle 6.29: Ergebnis Tendenz, Standardtitel 1988

Bei einer Reduzierung der Zellen in der Kohonenebene kommt es zum gleichen Verhalten wie bei den Versuchen mit den Daten von 1991. Die Versuche werden mit 20000 Lernzyklen durchgeführt. Eine Reduzierung auf 100 Kohonenzellen erbringt keine Verbesserung der Ergebnisse, sondern eine Verschlechterung. Es treten Trefferquoten zwischen 45% und 48% auf. Der kleinste erreichte Fehler liegt zwischen 6 und 9. Eine weitere Reduzierung auf 80 Zellen führt zum besten Ergebnis für die Dresdner Bank-Aktie. Mit einer Parameterausprägung von α=0.2 und β=0.2 wird ein Lernfehler von 10.63642 erreicht. Der Testfehler beträgt 12.71452. Es werden 52% der Muster richtig erkannt. Mit weiteren Verringerungen der Zahl der Kohonenzellen werden Ergebnisse erreicht, die sich nicht wesentlich von den bisherigen unterscheiden. Die Trefferraten liegen meist etwas niedriger. Tabelle 6.30 zeigt die Ergebnisse für ein Netz mit 12 Kohonenzellen.

Netz:14→12→1 α:0.05 β:0.2 Prognose: Tendenz Muster:139a/100b				
Titel	Zyklen	kl. Lernfehler	Testfehler	Quote
Degussa	20000	14.34815	11.97279	53.0%
Dresdner Bank	20000	14.57838	12.41723	49.0%
RWE	20000	13.76529	12.32585	52.0%

Tabelle 6.30: Ergebnis Tendenz, Standardtitel 1988

Es werden fast dieselben Trefferraten erzielt wie in Tabelle 6.29.

Als Fazit der Coutnerpropagation-Versuche läßt sich sagen, daß die Trefferraten, zumindest für die Daten von 1991, im Bereich der bisherigen Resultate liegen. Die Ergebnisse der Versuche mit Daten von 1988 liegen nur leicht unter denen, der Versuche, die mit anderen Lernverfahren durchgeführt wurden. Ein Nachteil sind natürlich die eingeschränkten Möglichkeiten bei der Wahl der Prognoseart.

Kapitel 7

Vergleich mit konventionellen Verfahren

In diesem Kapitel werden die besten erzielten Prognoseergebnisse mit Ergebnissen konventioneller Verfahren zur Bestimmung des Kurswertes verglichen. Es wird dabei nicht nach Lernverfahren unterschieden, sondern das beste Ergebnis aller Versuche ausgewählt. Die konventionellen Verfahren sind die Bildung von gleitenden Durchschnitten und die Extrapolation nach der Methode von Neville. Die Durchschnittsbildung ist kein echtes Vorhersageverfahren, hier wird als neuer Kurswert nur ein Mittelwert aus einer bestimmten Anzahl alter Kurswerte gebildet. Wenn der tatsächliche Kurs sehr beständig verläuft, wird der vorhergesagte Kurs wird also immer dem tatsächlichen Kurs hinterher laufen. Nur wenn sich die Tendenz des Kurses häufig ändert, können richtige Prognosen erzielt werden. Beim Extrpolationsverfahren wird mit Hilfe von n Stützstellen der Wert der Funktionswert an der Stelle n+1 ermittelt. Die Ergebnisse für die Kurse von 1991 sind Tabelle 7.1 zu entnehmen. Sie beziehen sich auf die Prognose der letzten 30 Muster(30b).

1991	Extrapolation			Gl. Durchschnitt		
	n	Quote	Abweich.	n	Quote	Abweich.
Degussa	2	50%	19.84 DM	3	50%	5.43 DM
Dresdner Bank	3	50%	19.60 DM	2	53.3%	4.61 DM
RWE	4	46.6%	15.78 DM	2	53.3%	3.04 DM

Tabelle 7.1: Ergebnisse mit konventionellen Verfahren, 1991

Für die Degussa-Aktie wird mit neuronalen Netzen eine Trefferquote von 70% erreicht. Die durchschnittliche Abweichung beträgt 8.36 DM. Die höchste Trefferrate der Dresdner Bank-Aktie ist 63.3%, mit einer Abweichung von 7.00 DM. Das beste Ergebnis der RWE-Aktie ist eine Trefferquote von 60%. Die Abweichung beträgt hierbei 9.35 DM. Alle Trefferraten der herkömmlichen Verfahren liegen also unter denen, der Prognosen mit neuronalen Netzen. Beim Extrapolationsverfahren sind auch die Abweichungen erheblich größer. Die geringen Abweichungen bei der Durchschnittsbildung sind dadurch zu erklären, daß nur selten größere Kurssprünge auftreten, somit liegen die Mittelwerte immer in der

Nähe des tatsächlichen Kurses.

Die Ergebnisse für die Kurse von 1988 sind in Tabelle 7.2 dargestellt. Es werden die letzten 100 Kurse (100b) vorhergesagt.

1988	Extrapolation			Gl. Durchschnitt		
	n	Quote	Abweich.	n	Quote	Abweich.
Degussa	3	52%	18.59 DM	3	53%	2.67 DM
Dresdner Bank	4	56%	33.66 DM	3	51%	3.22 DM
RWE	3	54%	6.96 DM	4	47%	2.59 DM

Tabelle 7.2: Ergebnisse mit konventionellen Verfahren, 1988

Es folgen wieder die besten Ergebnisse der Prognosen mit neuronalen Netzen. Die höchste Trefferrate für die Degussa-Aktie liegt 1988 bei 57%. Die Abweichung beträgt 3.55 DM. Mit verkleinerten Mustermengen (50) werden sogar 60% erreicht, die Abweichung beträgt dann nur noch 3.20 DM. Für die Dresdner Bank-Aktie ist nur eine geringere Trefferraten zu verzeichnen. Sie liegt bei 50%. Mit verkleinerten Mustermengen ist eine Trefferrate von 60% erreichbar. Vergleichswerte der konventionellen Verfahren liegen hierfür allerdings nicht vor. Für die RWE-Aktie liefern die neuronalen Netze wieder die höhere Trefferrate. Sie beträgt 61%. Die Abweichung ist dann 4.75 DM. Auch hier ergibt sich eine höhere Trefferrate, wenn die Mustermengen verkleinert werden. Sie liegt dann bei 62%. Die Trefferraten der Prognosen mit neuronalen Netzen liegen auch hier ,abgesehen von der Dresdner Bank-Aktie, immer über denen der konventionellen Verfahren. Allerdings sind die Differenzen nicht mehr so groß, wie bei den Versuchen mit Daten von 1991. Die Genauigkeit der Durchschnittsbildung ist, aus dem oben genannten Grund, auch hier etwas höher.

Es läßt sich also sagen, daß mit neuronalen Netzen höhere Trefferraten erzielt werden, als mit diesen konventionellen Verfahren. Man muß allerdings hinzufügen, daß die betrachteten Verfahren zu den einfacheren Verfahren gehören. Im Bereich der Extrapolation könnten beispielsweise andere, leistungsfähigere Verfahren angewendet werden.

Kapitel 8

Zusammenfassung

Es wurden die Lernverfahren *Backpropagation, Quickprop* und *Counterpropagation* untersucht. Mit den beiden erstgenannten Verfahren kann der genaue Kurswert prognostiziert werden, mit dem letztgenannten Verfahren ist nur die Prognose der Tendenz der Kursentwicklung möglich. Für die Versuche standen zwei Datenmengen zur Verfügung, die unterschiedliche Kurs- und Indexwerte enthielten. Es wurden unterschiedliche Netztopologien und Eingabeschemata verwendet und unterschiedliche Prognosearten durchgeführt. Die Versuche wurden mit den Aktientiteln Degussa, Dresdner Bank und RWE durchgeführt, die vereinzelt durch zusätzliche Titel ergänzt wurden. Für die Degussa-Aktie ergab sich die beste Trefferrate mit den Daten von 1991. Für die Prognose der Absolutkurse wurde mit dem Backpropagation-Verfahren eine Quote von 70% erreicht. Dieser Wert, der auch noch für die Prognose der Kursänderungen der Dresdner Bank erreicht wurde, stellt gleichzeitig die höchste Trefferrate aller Versuche dar. Mit der RWE-Aktie wurden 1991 nur Trefferraten von 60% erzielt. Die relative Abweichung der Prognosen liegt zwischen 1.5% und 2.5%. Diese Ergebnisse wurde alle mit dem Backpropagation-Verfahren erzielt. Bei den Versuchen mit den Daten von 1988 wurden meist geringere Trefferraten erzielt, dafür lag die Genauigkeit der Prognosen bei mehreren Versuchen über der, der Versuche mit Daten von 1991. Es wurden relative Abweichungen von 1% und weniger erzielt. Mit dem Quickprop-Verfahren ergaben sich nur in Einzelfällen erhebliche Verbesserungen des Lernaufwandes. Oft lag der Aufwand, wegen Sprüngen im Konvergenzprozeß, über dem Aufwand, der bei Backpropagation-Versuchen mit demselben Netz benötigt wurde. Oft trat auch keine Konvergenz im Lernprozeß ein. Die Approximation der Fehlerkurve durch eine Parabel scheint also nicht in jedem Fall ausreichend genau möglich zu sein. Betrachtet man die Gesamtheit der Versuche, so hat sich das Backpropagation-Verfahren als das stabilere Verfahren erwiesen. Welches Verfahren angewandt werden soll, muß von Fall zu Fall getestet werden, da das Verhalten mit verschiedenen Aktientiteln recht unterschiedlich ist. Aus diesem Grund läßt sich auch die Frage nach dem 'besten' Netz nicht eindeutig beantworten. Gute Resultate wurden mit verschiedenen Netztopologien erreicht, wobei keine Topologie für alle Aktientitel das beste Ergebnis lieferte. Zu nennen sind hier Netze mit zwei und drei inneren Ebenen, sowohl mit der logistischen, als auch mit der Sinus-Aktivierungsfunktion in der letzten inneren Ebene. Für größere Mustermengen wurden mit dem Netz mit drei inneren Ebenen bessere Resultate erzielt. Dieses ist in Abbildung 8.1 dargestellt.

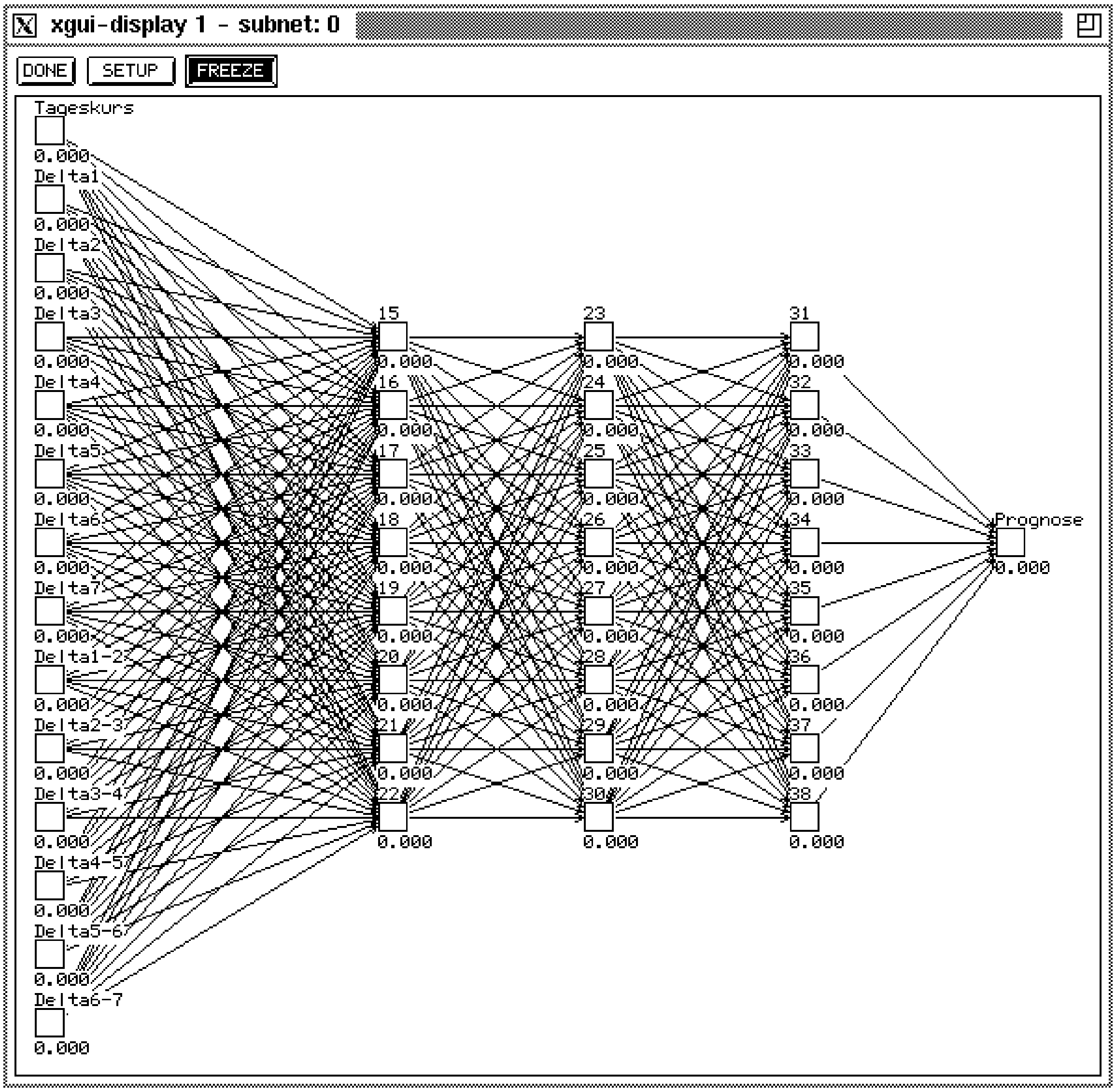


Abbildung 8.1: Neuronales Netz mit drei inneren Ebenen

Kapitel 9

Ausblick

In der vorliegenden Arbeit konnte nur ein kleiner Teil der vorhanden Variationsmöglichkeiten ausgenutzt werden. Neuronale Netze bieten noch ein großes Potential für weitere Versuche. Es stehen noch eine Vielzahl von Lernverfahren zu Verfügung, die nicht getestet wurden. Hier ist vor allem das von Jurik [Juri 91] entwickelte Backpercolation-Verfahren zu nennen. Hierbei werden in Feed-Forward-Netzen lokale Fehler eingeführt. Für verschiedene Prognose-Problemstellungen ,wie z.B. auch Vorhersagen der Sieger eines Rennens, wurden hiermit schon vielversprechende Ergebnisse erzielt. Aber auch hinsichtlich der Netztopologie können weitere Varianten getestet werden. Alternativen sind die Zahl der Ebenen, die Zahl der Zellen pro Ebene und die Verbindung der Ebenen untereinander.

Die unterschiedlichen Ergebnisse von Netzen mit logistischer und Sinus-Aktivierungsfunktion legen die Untersuchung von weiteren Aktivierungsfunktionen nahe. Interessant ist auch die Kombination von verschiedenen Aktivierungsfunktionen in einem Netz.

Die hohen Trefferraten bei den Versuchen mit der vielfältigen Datenmenge von 1991 geben Anlaß für weitere Untersuchungen mit noch umfangreicheren Eingabedaten. So könnten die Umsatzzahlen des Prognosetitels, oder die Zusätze, die das Verhältnis von Angebot und Nachfrage beschreiben (bezahlt, Brief, Geld, usw.) in kodierter Form einbezogen werden. Außerdem könnte darüber nachgedacht werden, in welcher Form Informationen über politische und wirtschaftliche Ereignisse und Nachrichten als Eingabedaten verwendet werden können.

Literaturverzeichnis

[Alek 89] I. Aleksander(Ed.): *Neural Computing Architectures: The Design of Brain-Like Machines.* 1989.

[Ande 88] D. Z. Anderson(Ed.): *Neural Information Processing Systems.* New York 1988.

[Eckm 90] R. Eckmiller(Ed.): *Parallel Processing in Neural Systems and Computers.* Participants Edition North-Holland, 1990.

[Fahl 88] S. E. Fahlman: *Faster-Learning Variations on Back-Propagation: An Empirical Study.* In [Tour 88] Part 1 S.38-51, 1988.

[Hech 90] R. Hecht-Nielsen: *Neurocomputing.* Addison-Wesley, 1990.

[Hint 89] G. E. Hinton: *Connectionist Learning Procedures.* In Artificial Intelligence, 40 S.185-234, 1989.

[Judd 90] J. S. Judd: *Neural Network Design and the Complexity of Learning.* MIT Press. London,England 1990.

[Juri 91] M. Jurik: *Backpercolation: assigning local error in feedforward perceptron networks.* Jurik Research & Consulting. 1991.

[Khan 90] T. Khanna: *Foundations of neural networks.* Addison Wesley, 1990.

[Krat 90] K. P. Kratzer: *Neuronale Netze: Grundlagen und Anwendungen.* Carl Hanser Verlag, 1990.

[Lape 88] A. Lapedes, R. Farber: *How Neural Nets Work.* In [Ande 88] S.442-456.

[Mill 90] R. K. Miller: *Neural Networks.* The Fairmont Press, Inc., 1989.

[Ritt 90] H. Ritter, T. Martinez, K. Schulten: *Neuronale Netze: eine Einführung in die Neuroinformatik selbstorganisierender Netze.* Bonn, München, Reading(Mass.). Addison-Wesley, 1990.

[Rumm 86]] D. E. Rummelhart, J. L. McClelland, PDP Research Group: *Parallel Distributed Processing. Explorations in the Microstructure of Cognition.* Volume I, II. MIT Press, 1986.

[Schö 90a] E. Schöneburg, N. Hansen, A. Gawelczyk: *Neuronale Netzwerke.* Markt & Technik-Verlag, München 1990.

[Schö 90b] E. Schöneburg: *Stock price prediction using neural networks: A project report.* In International Journal of Neurocomputing Vol. II, 1 S.17-27, 1990.

[Simp 90] P. K. Simpson: *Artificial neural systems: foundations, paradigms, applications and implementations.* Pergamon Press, 1990.

[Tour 88] D. S. Touretzky, G. Hinton, T. Sejnowski(Eds): *Proceedings of the 1988 Connectionist Models Summer School.* Morgan Kaufmann, San Mateo 1988.

[Wass 89] P. D. Wasserman: *Neural computing: theory and practice.* Van Nostrand Reinhold, 1989.

[Zell 91] A. Zell, T. Korb, N. Mache, T. Sommer: *SNNS (Stuttgarter Neuronale Netze Simulator) Benutzerhandbuch.* Bericht Nr. 1/91. Institut für Parallele und Verteilte Höchstleistungsrechner, Universität Stuttgart, 1991.