

Abverkaufsprognose im Supermarkt mit Neuronalen Netzen

Frank M. Thiesing und Oliver Vornberger

Universität Osnabrück

Fachbereich Mathematik/Informatik

Albrechtstr. 28, 49069 Osnabrück

Tel: 0541-969-2558; Fax: 0541-969-2770

E-mail: frank@informatik.uni-osnabrueck.de

Zusammenfassung

In diesem Beitrag werden Neuronale Netze zur Vorhersage der zu erwartenden Nachfrage nach Produkten eines Supermarktes eingesetzt. Feedforward-Multilayer-Perceptron Netzwerke werden mit Hilfe des Backpropagation-Algorithmus darauf trainiert, den wöchentlichen Absatz zu prognostizieren. Als Einflußfaktoren werden dazu Preise, Werbung und Feiertage berücksichtigt. Vorgestellt wird die Implementation eines Prototyps des entwickelten Neuronalen Prognosesystems, der in der Zentrale einer hessischen Supermarktkette im Einsatz ist. Die Prognosequalität des Neuronalen Netzes wird verglichen mit zwei statistischen Prognoseverfahren, die derzeit in dem Unternehmen eingesetzt werden.

Keywords: Neuronale Netze, Zeitreihen, Absatzprognose

1 Einleitung

Ein zentrales Problem in Wirtschaft und Wissenschaft ist die Prognose von Zeitreihen. Die Beispiele reichen von Bedarfsberechnungen für Fernwärme oder elektrischen Strom über die Bestimmung volkswirtschaftlicher Kenngrößen bis hin zur Börsenkursprognose [1, 2, 3].

In der Vergangenheit wurden zahlreiche statistische Verfahren entwickelt und zur Prognose eingesetzt [4]. Dabei muß für jedes Problem ein individuelles Modell unter Annahme des zugrundeliegenden Trends entwickelt werden. Die Spezifikation

dieses Modells erweist sich bei mehrdimensionalen Zeitreihen mit nicht-linearen Abhängigkeiten als äußerst schwierig.

Eine Alternative für die Prognose von mehrdimensionalen Zeitreihen stellen Neuronale Netze dar [5, 6]. Deren Einsatzfähigkeit für die Prognose wirtschaftlicher Vorgänge wurde anhand einer Reihe von Anwendungen unter Beweis gestellt [7, 8].

Die meisten Ansätze verwenden Feedforward-Multilayer-Netzwerke mit einem überwachten Trainingsalgorithmus, um die Abbildung zwischen Input- und Output-Daten zu generalisieren und die impliziten Regeln in der Bewegung der Zeitreihe zu entdecken sowie deren Verhalten in die Zukunft fortzuschreiben.

Es gibt zwei Ansätze, wie die Daten dem Neuronalen Netz präsentiert werden. Zum einen werden die Werte mehrerer verschiedener Indikatoren zum Zeitpunkt t im Inputvektor und der Wert der zu prognostizierenden Zeitreihe zum Zeitpunkt $t + 1$ als Output benutzt. Dieser Ansatz basiert auf der Annahme, daß die Entwicklung der verschiedenen Indikatoren zur Erklärung des Verhaltens der gesuchten Zeitreihe notwendig ist.

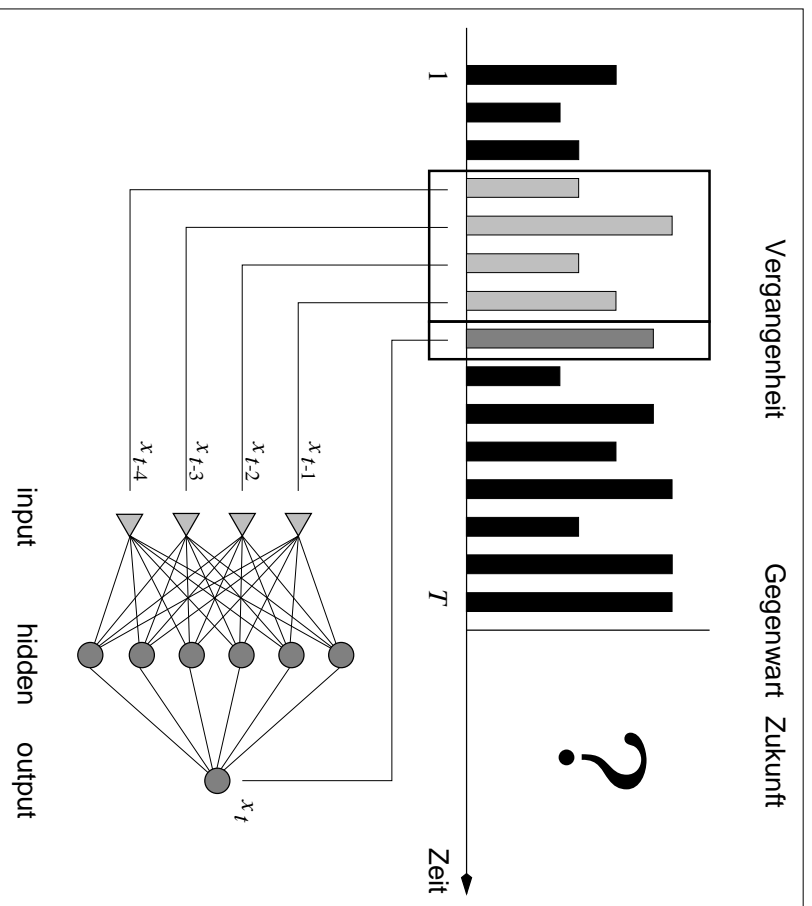


Abbildung 1: Univariante Zeitreihenprognose mit Neuronalen Netzen

Der zweite Ansatz zur Zeitreihenprognose wird realisiert, indem n aufeinander-

folgende Vergangenheitswerte bis zum Zeitpunkt t der zu prognostizierenden Zeitreihe als Inputvektor und der Wert zum Zeitpunkt $t + 1$ als gewünschter Output benutzt werden. Dieses Verfahren generiert die Zeitreihe in einem gleichenden Zeitfenster der Breite n (siehe Abbildung 1). Die Voraussetzung für diese Vorgehensweise ist, daß die gesamte zur Prognose benötigte Information (z.B. periodische Schwankungen) in der Zeitreihe selbst versteckt ist.

In diesem Beitrag werden Backpropagation-Netze darauf trainiert, die wöchentlichen Abverkaufszahlen von Produkten in einem Supermarkt vorherzusagen. Dazu werden mehrerer externe Einflußfaktoren benutzt. Die Prognosequalität der Neuralen Netze wird mit zwei Vorhersagemethoden verglichen, die im Supermarkt im Einsatz sind.

2 Die betrachteten Zeitreihen

Bei den hier betrachteten Zeitreihen handelt es sich um die Abverkaufszahlen von 20 Artikeln einer Produktgruppe in einem Supermarkt. Die Information über die Anzahl der verkauften Artikel und den damit erzielten Umsatz liegt wöchentlich vor ab September 1994. Bei der Analyse der Daten zeigen sich signifikante Einflüsse auf den Abverkauf: Werbekampagnen verbunden mit zeitlich begrenzten Preissenkungen fördern den Absatz, Feiertage schränken die Öffnungszeiten ein; außerdem zeigen sich saisonale Schwankungen.

Für die Prognose werden die Abverkaufszahlen der jüngeren Vergangenheit, die Preise, die Werbekampagnen und die Feiertage herangezogen. Als Saisonindikator wird der akkumulierte Gesamtumsatz in der Produktgruppe benutzt. Dessen Verhalten über den Zeitraum von mehr als 16 Monaten ist in Abbildung 2 schematisch dargestellt.

Für die Prognose werden Feedforward-Multilayer-Perceptron (MLP) Netzwerke eingesetzt, die mit dem Backpropagation-Verfahren trainiert werden [9, 10]. Zur Prognose des zukünftigen Abverkaufs werden als Input die Werte mehrerer betrachteter Zeitreihen angelegt. Aufgrund des Dispositionsablaufs liegt eine Lücke von einer Woche zwischen dem letzten verfügbaren Abverkaufswert und dem Prognosezeitpunkt (siehe Abbildung 3). Dadurch steht diese Woche für die Erstellung der Prognose und zur Disposition zur Verfügung. Die Information über Preis und Werbekampagnen für ein Produkt sowie über Feiertage steht bereits für die Prognosewoche fest, weshalb auch diese Information mit einfließt, wie in Abbildung 3 dargestellt.

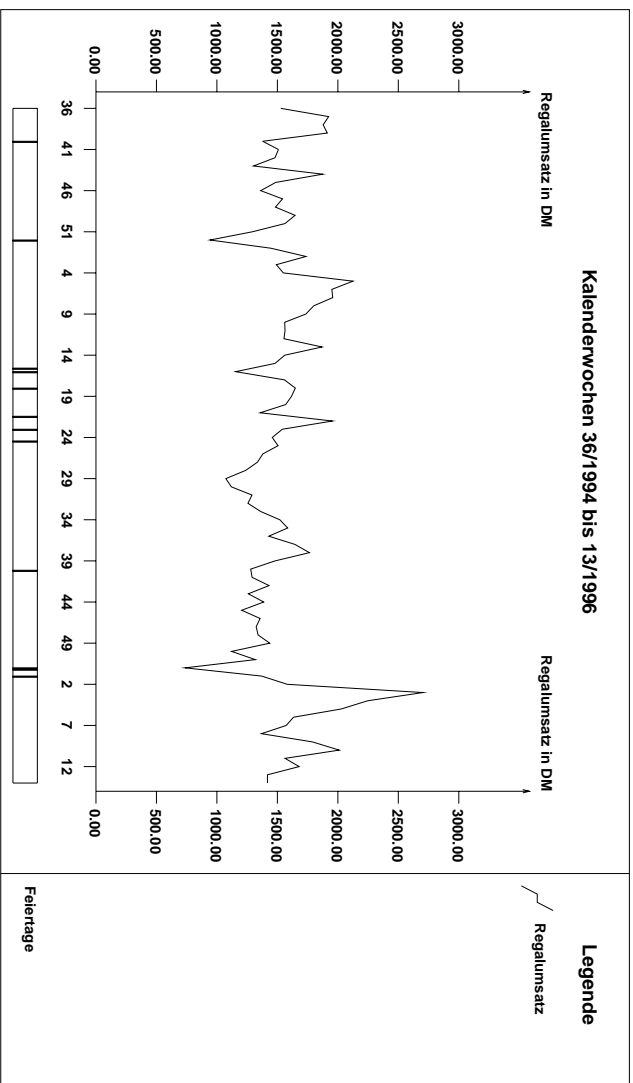


Abbildung 2: Akkumulierter Umsatz der Produktgruppe

3 Vorverarbeitung der Input-Daten

Die rohen Zeitreihendaten benötigen eine Vorverarbeitung, bevor sie in das Neuronale Netze eingespeist werden können. Im allgemeinen ist es besser, die rohen Daten in Indikatoren umzuwandeln, die die zugrundeliegende Information expliziter repräsentieren [11, 12, 13]. Dazu werden die Abverkaufszeitreihen normiert und wegen der sigmoiden Aktivierungsfunktion der Neuronen auf das Intervall $[0, 1]$ skaliert. Diese Skalierung unterstützt zum einen das Neuronale Netz bei seiner Lernaufgabe, zum anderen werden durch die Normierung die unterschiedlichen Indikatoren gleichgewichtet. Es werden verschiedene Skalierungen (z_t) für die Abverkäufe und die akkumulierten Umsätze (x_t) getestet:

$$z_t = \frac{x_t - \min(x)}{\max(x) - \min(x)} \cdot 0.8 + 0.1 \quad \text{bzw.}$$

$$z_t = \frac{x_t - \mu}{c \cdot \sigma} + 0.5$$

wobei *min* und *max* das Minimum bzw. Maximum der Zeitreihe sind, und μ und σ sind der Mittelwert bzw. die Standardabweichung. c ist ein Faktor, mit dem der Wertebereich der Skalierung kontrolliert werden kann.

Bei den Preisen ist die Preisveränderung der Indikator mit dem größten Einfluß, der gegensätzlich wirkt, abhängig davon, ob der Preis sinkt oder steigt. Deshalb

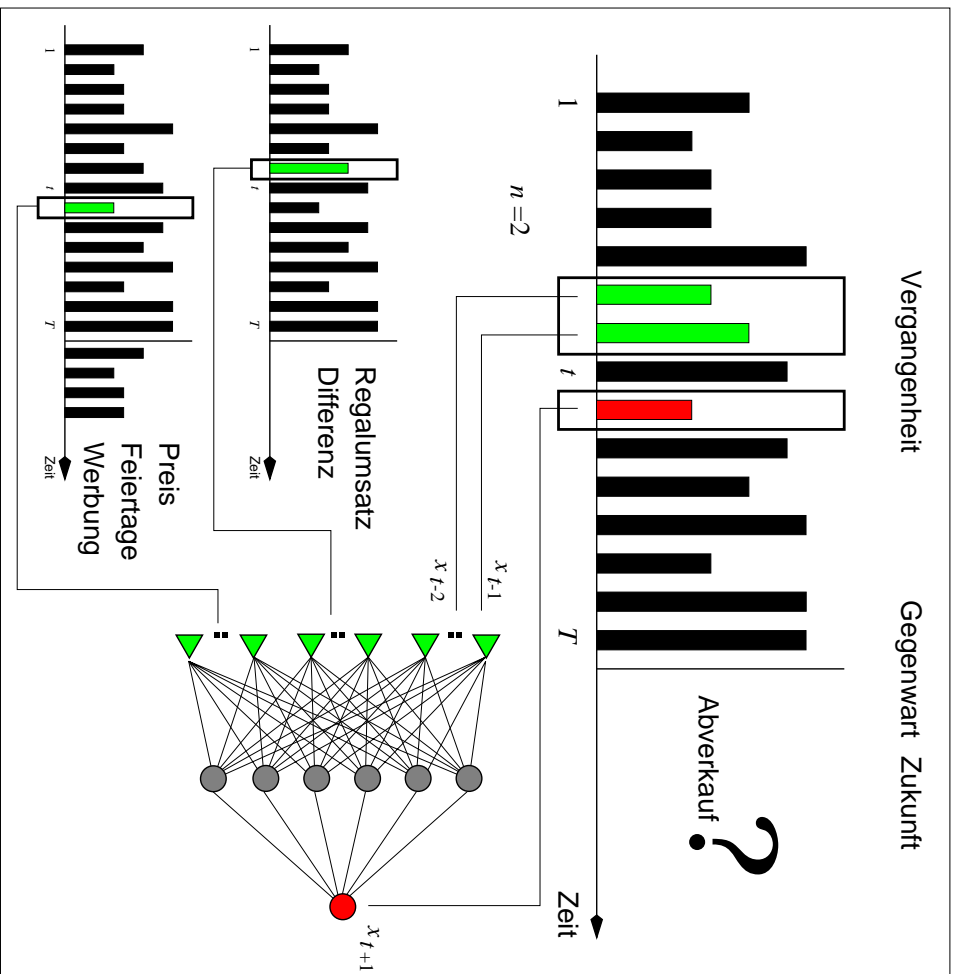


Abbildung 3: Ein- und Ausgabe des Neuronalen Netzes

werden die Preise wie folgt modelliert:

$$pri_t := \left. \begin{cases} 0.9 & : \text{ Preis steigt} \\ 0.0 & : \text{ Preis bleibt gleich} \\ -0.9 & : \text{ Preis fällt} \end{cases} \right\} \text{ in Woche } t$$

Sowohl für die Zeitreihe der Feiertage als auch für die Werbekampagnen werden binäre und lineare Modellierungen getestet, die die Zeitreihen mit Tageswerten wöchentlich aggregieren. Das binäre Modell unterscheidet lediglich, ob eine Werbeaktion in einer Woche stattfindet oder nicht, während die lineare Skalierung über die Anzahl der Aktionstage in einer Woche eine quantitative Aussage macht. Diese Indikatoren sind:

$$y_t := \left. \begin{cases} 0.9 & : \text{ falls Feiertag- bzw. Werbetag in Woche } t \\ 0.0 & : \text{ sonst} \end{cases} \right\} \text{ bzw. } \left. \begin{cases} \text{Anzahl der Feiertag- bzw. Werbetage in Woche } t & \text{(normalisierte Anzahl der Tage in Woche } t \text{)} \\ 6 & \end{cases} \right\}$$

4 Ergebnisse

Zur Ermittlung der günstigsten Parameterkonstellation, wurden in zahlreichen Tests unterschiedliche Einstellungen des Neuronalen Netzes und der Daten untersucht:

- Modellierung der Input-Zeitreihen
- Breite der Zeitfenster für jede Zeitreihe
- Anzahl der verdeckten Neuronen
- Intervall der initialen zufälligen Netzgewichte
- Lernrate und Momentumterm
- Anzahl und Wahl der Validierungsmuster
- Trainingsabbruch zur Verhinderung des Übertrainings

Zur Bewertung der Prognosequalität des Neuronalen Netzes (“Neuro”) werden die Ergebnisse in Tabelle 1 mit einem naiven Prognoseansatz (“Naiv”) und einer statistischen Prognose (“MovAvg”) für den Zeitraum der Kalenderwochen 44 des Jahres 1995 bis 13 des Jahres 1996 verglichen. In diesen Zeitraum fallen sowohl zahlreiche Aktionen als auch das Weihnachtsfest, das in der Vergangenheit den Umsatz dieser Produktgruppe negativ beeinflusst hat.

Gute Ergebnisse mit den Neuronalen Netzen wurden erzielt für eine Zeitfensterbreite von $n = 2$ für die Abverkaufszeitreihe. Außerdem erwies sich die Eingabe der Differenz zweier aufeinanderfolgender Abverkaufswerte ($x'_t = x_t - x_{t-1}$) als hilfreich für die Lernaufgabe der Neuronalen Netze. Zusätzlich fließen noch der Regalmassatz der jüngsten bekannten Woche ein sowie Feiertage, Werbeaktionen und Preisveränderungen in der Prognosewoche.

Daraus ergeben sich 7 Input-Neuronen. Für jeden Artikel wird ein solches Netz mit 4 verdeckten Neuronen darauf trainiert, die Prognose für die übernächste Woche vorzunehmen. Bei den Skalierungen zeigen sich die binäre Skalierung für die Tagesinformation und die lineare σ - μ -Skalierung für die Abverkaufszahlen überlegen. Als Lernrate wird 0.3 und als Impuls 0.1 eingestellt. Die initialen Gewichte werden zufällig aus dem Intervall $[-0.5, 0.5]$ gewählt. Der Trainingsabbruch erfolgt anhand des Netzfehlers auf einer Menge von zwölf Validierungsmustern.

Tabelle 1: Prognosefehler: RMSE/Mean

Artikel	Mean	Neuro	MovAvg	Naiv
036252	7.79	1.037	1.217	1.181
078924	8.97	1.084	1.409	1.551
180689	9.92	0.973	1.243	1.231
215718	12.73	0.468	0.601	0.612
215732	12.97	0.325	0.402	0.395
215749	8.63	0.485	0.471	0.535
228558	17.40	0.481	0.709	0.923
229104	63.88	0.431	0.591	0.649
289573	8.29	0.992	1.154	1.325
304962	15.35	0.971	1.267	1.523
341110	6.79	0.672	0.815	0.755
341127	5.79	0.953	1.114	1.293
362238	10.88	1.221	1.573	2.138
372206	9.69	3.119	3.253	4.513
392785	19.79	0.513	0.624	0.534
399883	11.47	0.407	0.530	0.444
468978	23.99	0.263	0.318	0.387
468985	17.19	0.942	1.135	1.420
567411	3.95	0.586	0.569	0.623
852234	7.75	0.891	1.125	1.211
Durchschnitt	—	0.841	1.006	1.162

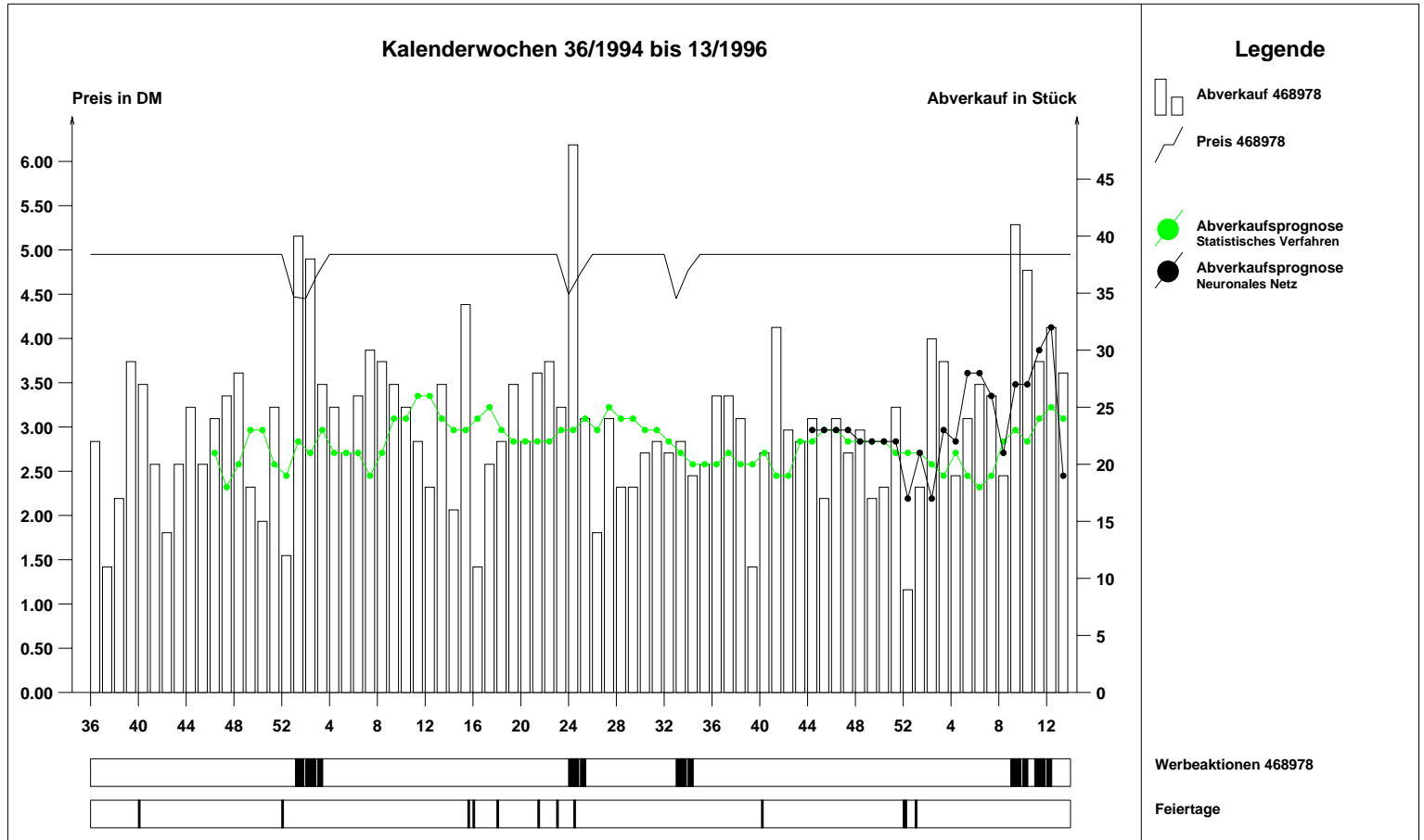
4.1 Naive Prognose

Die naive Prognosemethode benutzt den letzten bekannten Wert der Zeitreihe als Prognosewert für die Zukunft: $\hat{x}_{t+1} := x_{t-1}$. Diese Prognosemethode wird häufig von den Angestellten des Supermarktes angewandt.

4.2 Statistische Prognose

Die statistische Methode ist zur Zeit als Bestandsempfehlung im Einsatz. Es handelt sich dabei um die Berechnung des gleitenden Durchschnitts von maximal 9 aufeinanderfolgenden Abverkaufswerten der zuvor durch Ausreißerbehandlung geglätteten Zeitreihe.

Abbildung 4: Statistische und neuronale Abverkaufsprognose des Artikels 468978



4.3 Vergleich der Prognosemethoden

Zur Messung des Fehlers wird die Wurzel des mittleren quadratischen Fehlers (RMSE) dividiert durch den Mittelwert (“Mean”) der Zeitreihe. Die Ergebnisse in Tabelle 1 wurden durch die sukzessive Prognose der genannten 22 Wochen zum Jahreswechsel 1995/96 ermittelt.

Der Tabelle ist zu entnehmen, daß der naive Ansatz von den beiden anderen Prognosemethoden deutlich in den Schatten gestellt wird. Für 18 der 20 Artikel können mit den Neuronalen Netzen bessere Prognoseergebnisse als mit der zur Zeit angewendeten Methode erzielt werden.

Eine Betrachtung der beiden Zeitreihen, die von der Prognosemethode des gleichenden Durchschnitts favorisiert werden, zeigt, daß es sich hierbei um Zeitreihen handelt, die durch unbekannte zufällige Einflüsse keine ausreichenden Gesetzmäßigkeiten erkennen lassen, die das Neuronale Netz erlernen könnte. Insbesondere beim Artikel 567411 ist dieses auch an seinem geringen durchschnittlichen Wochenabsatz (“Mean”) von weniger als 4 Stück zu erkennen.

Abbildung 4 zeigt exemplarisch die Ergebnisse für einen Artikel gegenübergestellt sowohl für die neuronale als auch die statistische Methode im Vergleich zu den wirklichen Werten. Auch der Preis, die Dauer der Werbekampagnen und die Feiertage sind in der Abbildung 4 dargestellt, die mit Hilfe des entwickelten Prognose- und Analysewerkzeugs erstellt wurde.

5 Zusammenfassung und Ausblick

Für eine Auswahl von 20 Artikeln einer Produktgruppe werden Neuronale Netze eingesetzt, um den Abverkauf in einem Supermarkt zu prognostizieren. Auf Basis der Abverkaufsinformation der Vergangenheit zusammen mit äußeren Einflüssen wie Preisen, Werbekampagnen und Feiertagen werden die Netze trainiert und anschließend für die Prognose abgefragt. Die Testergebnisse zeigen, daß die Neuronalen Netze besser in der Lage sind, die zukünftige Entwicklung zu prognostizieren, als die zum Vergleich herangezogenen Prognoseverfahren, die zur Zeit im Supermarkt zum Einsatz kommen.

Der implementierte Ansatz verwendet für alle Artikel die gleiche Netzarchitektur und die gleiche Vorverarbeitung der rohen Zeitreihen. Die Optimierung der vielen einstellbaren Parameter basiert auf den Ergebnissen für alle betrachteten Artikel und wurde nicht speziell für einen bestimmten angepaßt.

Der entwickelte Prototyp ist Teil eines Prognosesystems, das die Rohdaten auf-

bereitet, die notwendigen Vorverarbeitungen vornimmt, die Neuronalen Netze trainiert und die prognostizierten Werte liefert. Bestandteil dieses Systems ist auch eine interaktive Benutzeroberfläche, mit der die betrachteten Zeitreihen und ihre Abhängigkeiten untereinander graphisch analysiert werden können und die die Optimierung der vielen Parameter in Testläufen unterstützt.

Die weiteren Schritte auf dem Weg zur Verbesserung der Einsatzfähigkeit auf großen Datenmenge werden sein, zusätzliche adaptive Datentransformationen zu entwickeln, die Signifikanz der Input-Daten in Hinblick auf eine Reduzierung der Netzkomplexität zu prüfen und andere Trainingsverfahren Neuronaler Netze auf ihre Eignung hin zu untersuchen. Dies wird mit Hilfe des Stuttgarter Neuronale Netze Simulators SNNS geschehen [14].

Literatur

- [1] Schöneburg, E., "Stock Price Prediction Using Neural Networks: An Empirical Test," *Neurocomputing*, 2, 1, 1991.
- [2] Refenes A.N., Azema-Barac M., Chen L., Karousos, S.A., "Currency Exchange Rate Prediction and Neural Network Design Strategies," *Neural Computing & Applications*, 1(1) pp. 46–58, 1993.
- [3] Preisleben, B., Ripper, K., "Economic Forecasting Using Neural Networks," *Proceedings of the 1995 IEEE International Conference on Neural Networks*, Vol. 2, pp. 833–838, Perth, WA, 1995.
- [4] Weigend A.S., Gershenfeld, N.A., *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, 1994.
- [5] Rojas, R., *Theorie der neuronalen Netze*, Springer, 1993.
- [6] Zell, A., *Simulation Neuronaler Netze*, Addison-Wesley, 1994.
- [7] Chakraborty, K., Mehrotra, K., Mohan, C.K., Ranka, S., "Forecasting the Behaviour of Multivariate Time Series Using Neural Networks," *Neural Networks*, Vol. 5, pp. 961–970, 1992.
- [8] Vennuri, V.R., Rogers, R.D., *Artificial Neural Networks – Forecasting Time Series*, IEEE Computer Society Press, 1994.
- [9] Rumelhart, D.E., Hinton, G.E., Williams, R.J., "Learning internal representations by error propagation", in Rumelhart, D.E., McClelland (Eds.) *Parallel*

Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1, MIT Press, Cambridge, MA, 1986, pp. 318–362.

[10] Rumelhart, D.E., Hinton, G.E., Williams, R.J.: “Learning representations by back-propagating errors”, *Nature*, 323: 533-536.

[11] Rehkugler, H., Zimmermann, H.G., *Neuronale Netze in der Ökonomie*, Verlag Vahlen, München 1994.

[12] Zimmermann, H.-J. (Hrsg.), *Datenanalyse*, VDI Verlag, 1995.

[13] Joereßen, A., “Systematisierung der Datenverarbeitung für die Prognose”, Sitzung der Arbeitsgruppe “Prognoseverfahren” und “Fuzzy Sets und Neuronale Netze” in der DGOR, Stralsund, 12./13.9.1996.

[14] *SNNS Stuttgart Neural Network Simulator User Manual, Version 4.1*, Universität Stuttgart, Report No. 6/1995.