Brandenburgische Technische Universität Cottbus
Lehrstuhl für Numerische und Angewandte Mathematik

3. Cottbuser Workshop
Aspekte Neuronalen Lernens

CoWAN'98

Programm

Die Veranstalter

L. Cromme **T. Kolb** **H. Koch**

danken

SUN Microsystems GmbH
Holiday Inn / Holiday Inn Express Hotel Cottbus

für die gewährte Unterstützung.

Inhaltsverzeichnis

Allgemeine Hinweise

Tagungsleitung

Brandenburgische Technische Universität Cottbus
Lehrstuhl für Numerische und Angewandte Mathematik
CoWAN'98
Postfach 101344
D-03013 Cottbus
Tel.: 035569-2409
Fax.: 035569-2776
email: `cowan@math.tu-cottbus.de`
Homepage: `http://www.math.tu-cottbus.de/cowan/`

Tagungsraum

Der Workshop findet im Hörsaal 1 im Lehrgebäude 1 (LG 1, Erdgeschoß) auf dem Campus der BTU statt (siehe auch Campusplan auf der vorderen inneren Umschlagseite).

Tagungsbüro

Das Tagungsbüro befindet sich im Foyer zum Hörsaal 1 des Lehrgebäudes 1 im Erdgeschoß.
Es ist zu folgenden Zeiten besetzt:

Montag,	05.20.1998:	16.00–21.00 Uhr
Dienstag,	06.10.1998:	08.30–17.30 Uhr
Mittwoch,	07.10.1998:	08.30–13.00 Uhr

Verpflegung

Möglichkeiten zum Mittagessen bestehen in der Mensa sowie in der Cafeteria im Mensagebäude der BTU. Möglichkeiten zum Abendessen bestehen von etwa 17.30 Uhr bis 19.00 Uhr in der Bierklause, welche sich ebenfalls im Mensagebäude befindet.

Anfahrtswege

- Züge der DB:
von Dresden, Leipzig oder Berlin
- Kfz-Anreise von Dresden bzw. Berlin:
A13 bis zum Autobahndreieck Spreewald, dort auf der A15 bis zur Autobahnabfahrt Cottbus-West fahren; im Stadtgebiet Cottbus den Hinweisschildern zur Technischen Universität folgen (siehe auch Skizze der PKW-Verbindungen nach Cottbus und Skizze der Anfahrwege zur BTU auf den hinteren Umschlagseiten).

Allgemeine Hinweise

Öffentlicher Verkehr in Cottbus

Hauptbahnhof – BTU Cottbus:
Haltestelle Hbf. mit Buslinie 16 Richtung Schillerplatz bis zur Haltestelle Mensa/BTU in der J.-Gagarin-Str.

Rahmenprogramm

Am Abend des 6. Oktobers besteht für die Teilnehmer des Workshops die Möglichkeit zu einem gemeinsamen Abendessen in einem Cottbuser Restaurant. Die Veranstaltung wird durch die Tagungsleitung für die Teilnehmer mit Anmeldung organisiert.

Tagungsunterlagen

Die Tagungsunterlagen einschließlich das vorliegende Programm liegen für die Teilnehmer im Tagungsbüro bereit.

Tagungsband

Der Tagungsband zu diesem Workshop wird beim Shaker Verlag erscheinen. Die Vortragenden werden gebeten, einen Artikel von etwa 5-20 Seiten Länge vorzubereiten und zum Workshop auf Diskette mitzubringen oder spätestens bis eine Woche nach dem Workshop auf elektronischen Wege an die Tagungsleitung zu senden. Der Tagungsband wird in LaTeX erstellt, sodaß wir um Einreichung der Beiträge im LaTeX-Format bitten. Wir bitten die Vortragenden, hierzu die Bemerkungen im Abschnitt „Hinweise für Autoren“ dieses Programmheftes bzw. die entsprechenden WWW-Seiten zu beachten. Unmittelbar nach seiner Fertigstellung wird der Tagungsband an alle Teilnehmer versandt.

Wissenschaftliches Programm

Montag, 05.10.1998, 19.00-21.00
Eröffnung (Hörsaal 1, LG 1)

	Eröffnungsvortrag:
19:00-21:00	*K. Obermayer* (Berlin): „Neuronale Netze für die Visualisierung und die statistische Datenanalyse“

Dienstag, 06.10.1998, 09.00-12.30
Lerntheorie (Hörsaal 1, LG 1)

	Hauptvortrag:
09:00-09:50	*P. Auer* (Graz): „Boosting-Algorithmen und Neuronale Netze“
	Einzelvorträge:
09:50-10:20	*F. Hamker* (Ilmenau): „Lebenslang lernfähige Zellstrukturen: Eine Lösung des Stabilitäts-Plastizitäts-Dilemmas“
10:40-11:10	*S. Behnke* (Berlin): „Wachsende Neuronale Netze mit Hyperbolischen Radialen Basisfunktionen“
11:10-11:40	*M.-H. Breitner* (Clausthal): „SQP-Training von Perzeptrons mit dem Neurorechner SYNAPSE 3 zur schnellen Berechnung heuristischer Optionpreismodelle“
11:40-12:10	*L. Cromme* (Cottbus): „Lernverfahren in Feedforward-Netzen“

Dienstag, 06.10.1998, 13.30-17.30
Intelligent Control (Hörsaal 1, LG 1)

	Hauptvortrag:
13:30-14:20	*A. Bühlmeier* (Dortmund, Bremen):„Intelligent Control: Von neuronalen Konzepten zur Anwendung“
	Einzelvorträge:
14:20-14:50	*F. Leisch* (Wien): „Zur Stationarität autoregressiver ANN Modelle“
14:50-15:20	*T. Kolb* (Cottbus): „Lernen in Rekurenten Neuronalen Netzen?“
	Schwerpunkt Industrielle Anwendungen:
15:40-16:10	*M. Borschbach* (Münster): „Aspekte Neuronalen Lernens für die medizinische Analyse von Gehirnaktivitäten anhand von MEG (Magnetenzephalographie) - Daten“
16:10-16:40	*B. Herm* (Leipzig): „Die Wertigkeit des Corticotropin-Releasing-Hormon (CRH)-Tests im Vergleich zum Insulinhypoglykämietest in Bezug auf die Beurteilung der Hypothalamus-Hypophysen-Nebennierenrinden-Achse bei wachstumsgestörten Patienten'“
16:40-17:10	*B. Beirow* (Cottbus): „Nutzung neuronaler Netze zur Lösung dynamischer Problemstellungen bei turmartigen Bauwerken“

Wissenschaftliches Programm

Mittwoch, 02.10.1998, 09.00-12.30
Hybride Systeme (Hörsaal 1, LG 1)

	Hauptvortrag:
09:00-09:50	*W.-M. Lippe* (Münster): „Hybride Systeme“
	Einzelvorträge:
09:50-10:20	*P. Feucht* (Karlsruhe): „Interaktiv lernende Diagnosesysteme für Industrieanwendungen“
10:40-11:10	*T. Adam* (Cottbus): „Minimierung der Energiebezugskosten in Kläranlagen durch Optimierung der Energieversorgung mittels Kostenmodell und dem Einsatz künstlicher Intelligenz neuronaler Netze“
11:10-11:40	*A. Ultsch* (Marburg): „Hybrid-Neuronale Systeme und ihre Anwendung zum Data Mining“
11:40-12:10	*O. Schleider* (Cottbus) „Parameteradaption hydrodynamisch - numerischer Modelle mit neuronalen Methoden“

„Neuronale Netze für die Visualisierung und die statistische Datenanalyse“

K. Obermayer
Berlin

In den letzten Jahren haben sich Neuroinformatik und Statistik einander sehr stark angenähert, und es sind eine Reihe interessanter Arbeiten im Spannungsfeld zwischen Künstlichen Neuronalen Netzen und den Methoden der klassischen Statistik entstanden. Ich möchte in meinem Vortrag drei moderne Entwicklungen umreißen, die alle drei bereits zu Erfolgen in der Praxis geführt haben.

Im ersten Teil werde ich über neue Verfahren zur unüberwachten Separation von Signalgemischen sprechen und diese an Hand von Anwendungen in der Bild- (optische Messung von Hirnaktivität) und Signalverarbeitung (Trennung akustischer Signalgemische) illustrieren.

Im zweiten Teil möchte ich Zusammenhänge zwischen Autoencodern und Clustering-Verfahren (z.B. Kohonen Karten) aufzeigen und daraus neue, robustere Verfahren zur Gruppierung von Daten ableiten. Ich werde diese Verfahren u.a. an Hand zweier Beispiele näher erläutern:

(1) "source-channel coding" Probleme bei der Übermittlung von Daten und

(2) Clustering auf Distanzmatrizen am Beispiel anatomischer Daten von der Hirnrinde der Katze.

Im dritten Teil, schließlich, bespreche ich das Prinzip der "structural risk minimization" (siehe, z.B., Vapnik, 1995). Dieses Prinzip hat zu einem neuen Ansatz zur Lösung von Klassifikationsproblemen geführt, der unter dem Namen "support vector machines" in den letzten Jahren große Verbreitung gefunden hat. Nach einer kurzen Einführung in die Methode werde ich diese auf die Analyse paarweiser und relationaler Daten erweitern. Es zeigt sich, daß die neue Methode konsistent bessere Klassifikationsergebnisse liefert als bisher übliche Verfahren wie z.B. die Methode der K nächsten Nachbarn.

„Boosting-Algorithmen und Neuronale Netze“

P. Auer
Graz

Boosting-Algorithmen zählen zu den erfolgreichsten Verfahren des Überwachten Lernens für Klassifikationsprobleme. Zum Beispiel schnitt bei einem Vergleich von Lernalgorithmen zur Erkennung von handgeschriebenen Ziffern [1] ein Neuronales Netz kombiniert mit einem Boosting-Algorithmus am besten ab.

Boosting ist eine allgemeine Methode zur Verbesserung der Genauigkeit von Klassifikationsalgorithmen. Dabei werden durch unterschiedliche Gewichtung der Trainingsbeispiele mehrere Klassifikatoren berechnet, deren Linearkombination dann einen verbesserten Klassifikator liefert.

In meinem Vortrag werde ich das grundlegende Boosting-Verfahren [2] erläutern und auch eine neuere Variante [3] vorstellen, die sich insbesonders zum Trainieren von Neuronalen Netzen eignet. Dann möchte ich Arbeiten besprechen, die versuchen die Effektivität von Boosting zu erklären [4, 5, 6, 7]. Daraus ergeben sich konkrete Fragen zur Kombination von Neuronalen Netzen und Boosting:

1. In mehreren Arbeiten wird Boosting als sehr unempfindlich gegen Overfitting beschrieben. Wie allgemein gültig ist diese Beobachtung?

2. Ist die Kombination von mehreren kleinen Neuronalen Netzen mittels Boosting einem komplexeren Neuronalen Netz überlegen?

3. Die Analyse zeigt, daß Boosting eine andere Fehlerfunktion als den quadratischen Fehler minimiert. Ist diese Fehlerfunktion dem quadratischen Fehler überlegen?

Den Versuch einer Antwort auf diese Fragen will ich an Hand von vorläufigen Ergebnissen einer empirischen Untersuchung unternehmen.

Literatur

[1] Y. LeCun, L. Jackel, L. Bottou, A. Brunot, C. Cortes, J. Denker, H. Drucker, I. Guyon, U. Müller, E. Säcklinger, P. Simard, and V. Vapnik. Comparison of learning algorithms for handwritten digit recognition. In *Proceedings of ICANN'95*, pages 53–60, 1995.

[2] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory: Eurocolt '95*, pages 23–37. Springer-Verlag, 1995.

[3] R. E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, 1998.

[4] Robert E. Schapire, Yoav Freund, Peter Bartlett, and Wee Sun Lee. Boosting the margin: a new explanation for the effectiveness of voting methods. In *Proc. 14th International Conference on Machine Learning*, pages 322–330. Morgan Kaufmann, 1997.

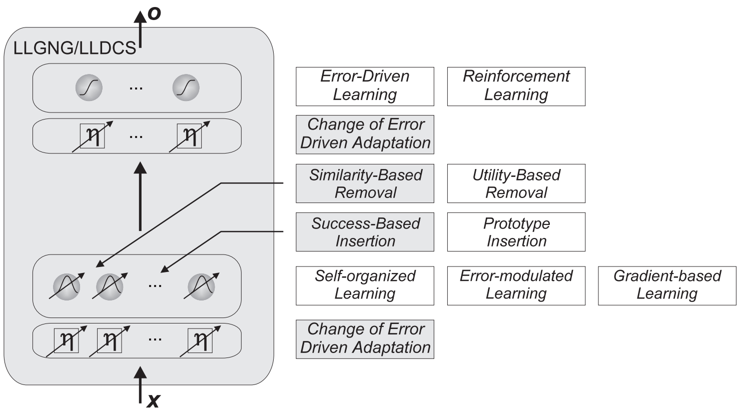
[5] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. Preprint, `http://www-stat.stanford.edu/~trevor/Papers/`, 1998.

[6] A. J. Grove and D. Schuurmans. Boosting in the limit: Maximizing the margin of learned ensembles. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, 1998.

[7] L. Mason, P. Bartlett, and J. Baxter. Direct optimization of margins improves generalization in combined classifiers. Technical report, Deparment of Systems Engineering, Australian National University, 1998.

„Lebenslang lernfähige Zellstrukturen: Eine Lösung des Stabilitäts-Plastizitäts-Dilemmas“

F. Hamker
Ilmenau, Frankfurt



Aufbau der „lebenslang lernfähigen Zellstrukturen“ bzw. des LLGNG/LLDCs Algorithmus. Die neu entwickelten Verfahren sind grau hinterlegt

Diese Arbeit beschreibt eine Weiterentwicklung der Zellstrukturen endlich begründet durch die Beiträge von Fritzke [1992; 1994; 1995], Bruske und Sommer [1995] bzw. Bruske [1998] zur Lösung des Stabilitäts-Plastizitäts-Dilemmas. Das als "lebenslang lernfähige Zellstrukturen" bezeichnete neuronale Netz verwendet die in der Abbildung hervorgehobenen Module, um das Einfügen von Knoten bei Datenströmen mit einer stationären Wahrscheinlichkeitsdichteverteilung, sogar in überlappenden Entscheidungsregionen, selbständig begrenzen können, mit der gleichzeitigen Fähigkeit, bei einer relevanten Änderung der Wahrscheinlichkeitsdichte erneut Knoten einfügen zu können. Weiterhin wird demonstriert, daß die Bewegung der Knoten bei einer stationären Wahrscheinlichkeitsdichteverteilung bis auf Null reduziert werden kann, ohne das Netz fest einzufrieren, sondern, daß sich bei einer relevanten Änderung der Wahrscheinlichkeitsdichteverteilung die betreffenden Knoten erneut bewegen können. Weiterhin bleibt die Fähigkeit zum Einfügen neuer Knoten gewahrt, um dem Netz zu ermöglichen, selbständig eine angemessene Anzahl von Knoten zur Lösung der Aufgabe zu bestimmen. Die dafür notwendige Stabilität sowie die Plastizität wirkt nur lokal, d.h., während das Netz an einigen Stellen hochgradig plastisch ist, kann es an anderen Stellen in einem vollkommen stabilen Zustand verbleiben. Der entwickelte Algorithmus bietet keine Insellösung, sondern eine Erweiterung der bestehenden Zellstrukturen mit ihren Varianten. Das bisher zuungunsten der Plastizität ausgelegte Dilemma wird im Kontext des lebenslangen Lernens neu interpretiert und eine leistungsfähigere beide Komponenten berücksichtigende Lösung entwickelt.

Konkret beruhen die Erweiterungen auf zwei Säulen: Erstens, durch die nachträgliche Bewertung des Einfügens von neuen Knoten lernt das Netz, ob sich ein weiteres Einfügen an dieser entsprechenden Stelle noch lohnt. Zweitens, durch Beobachtung der Leistung des Netzes mit zwei Fehlerzählern unterschiedlicher Zeitkonstante können relevante, lokale Änderungen in der Wahrscheinlichkeitsdichteverteilung entdeckt und die Lernparameter entsprechend adaptiert werden.

Die empirischen Untersuchungen mit künstlichen Daten und der Benchmarks zeigen am Beispiel des *Error-driven* Lernens, daß die entwickelte Lösung des Stabilitäts-Plastizitäts-Dilemmas einen optimalen Kompromiß zwischen Stabilität und Plastizität darstellt. Dieser optimale Kompromiß ist dadurch charakterisiert, daß

- die Stabilität sowie die Plastizität lokal im Netz, d.h. für jede einzelne Stützstelle definiert ist
- die Stabilität sowie die Plastizität die Anzahl der Stützstellen des Approximators, die Lernrate zur Positionierung der Stützstellen sowie die Lernrate des Gradientenabstiegs im Fehlergebirge betrifft
- die Anzahl der Stützstellen im Netz keiner Vorgabe bedarf
- bei einmal erreichter Stabilität, eine lokale Plastizität erst über eine relevante Änderung der Wahrscheinlichkeitsdichteverteilung, d.h. eine Änderung der Fehlerdichteverteilung eintritt.

„Wachsende Neuronale Netze mit Hyperbolischen Radialen Basisfunktionen“

S. Behnke
Berlin

Zweischichtige neuronale Netze mit radialen Basisfunktionen (RBF) in den Units der verdeckten Schicht erfreuen sich großer Beliebtheit für Probleme aus den Bereichen Klassifikation und Funktionsapproximation, da die Effekte der radialen Units gut lokalisiert sind. In der Regel werden beide Schichten mit unterschiedlichen Verfahren trainiert. Zunächst wird die verdeckte Schicht unüberwacht initialisiert, z.B. durch Clusterung der Eingabevektoren. Die Gewichte der Ausgabeunits werden anschließend durch ein überwachtes Verfahren trainiert. Diese Vorgehensweise hat den Nachteil, daß die Positionen der radialen Units unabhängig von der gewünschten Ausgabe sind. Man kann diesen Nachteil vermeiden, wenn die Gewichte beider Schichten durch einen fehlerminimierenden Lernalgorithmus, wie Backpropagation, angepaßt werden. Wie kürzlich von N.B. Karayiannis gezeigt wurde, eignen sich die üblicherweise verwendeten Gaußschen Basisfunktionen nur bedingt für den Gradientenabstieg, da ihre Ableitung an der Stelle der höchsten Antwort verschwindet. Besser geeignet sind hyperbolische Basisfunktionen, bei denen sowohl Antwort, als auch Ableitung zum Zentrum hin zunehmen.

Im vorliegenden Beitrag untersuchen wir, wie mittels Wachstumsverfahren eine geeignete Architektur für die Lösung von Klassifikationsproblemen mit hyperbolischen RBF-Netzen gefunden werden kann. Dabei wird das Netz zunächst mit einer radialen Unit initialisiert. Die Gewichte beider Schichten werden durch Gradientenabstieg angepaßt. Neue radiale Units werden an Stellen eingefügt, an denen das Netz noch nicht perfekt arbeitet. Der Radius der hinzugefügten Units wird immer kleiner, so daß die zu approximierende Funktion in einer Grob-zu-Fein-Strategie angenähert wird.

Die vorgeschlagenen Wachstumsverfahren wurden an einer synthetischen Datenmenge, dem Problem der verschlungenen Spiralen, und an einer Datenmenge aus dem Bereich der Spracherkennung getestet. Dabei wurden sehr zufriedenstellende Ergebnisse erzielt.

„SQP-Training von Perzeptrons mit dem Neurorechner SYNAPSE 3 zur schnellen Berechnung heuristischer Optionpreismodelle“

M.-H. Breitner
Clausthal

Das überwachte Backpropagation- und Momentum-Term-Lernen von Perzeptrons ist in der Regel relativ robust, aber sehr langsam. Verfahren der restringierten, nichtlinearen Optimierung – z. B. SQP- und GGN-Verfahren – zur Minimierung des Trainingsdatenfehlers haben entscheidende Vorteile:

(1) Suchrichtung und Lernrate werden für jeden Gewichtsupdate optimiert und die Trainingszeiten drastisch reduziert;

(2) Beschränkungen, z. B. der Gewichte oder der Gesamtkrümmung, werden exakt erfüllt.

Sehr problematisch ist, daß der Trainingsdatenfehler für viele Trainingsmuster und/oder viele Neuronen sehr viele lokale Minima hat. Notwendig ist also, verschiedene Topologien mit jeweils vielen verschiedenen Gewichtsinitialisierungen zu trainieren, um mit hoher Wahrscheinlichkeit gute Perzeptrons zu finden. Expert-Council-Topologien verbessern dann die globalen Approximationseigenschaften. Die extrem langen Rechenzeiten für viele Anwendungen aus der Praxis erfordern den Einsatz von Höchstleistungsrechnern, d. h. teuren Parallel- oder preiswerteren Neurorechnern. Der Neuro-, besser Matrixrechner SYNAPSE 3 ist über den PCI-Bus in Standard-PCs einsetzbar und liefert bis zu 6 Giga-Festpunktoperationen pro Sekunde (16 Bit Addition, 48 Bit Multiplikation). Für Perzeptrons ist die Backpropagation, d. h. die Berechnung des Fehlergradienten, durch die verfügbaren Matrixoperationen des SYNAPSE 3 in nur 5 – 20% der PC-Rechenzeit möglich. In Zusammenarbeit mit der Dresdner Bank, Frankfurt, und der Market-Maker GmbH, Kaiserslautern, werden heuristische Optionspreismodelle entwickelt. Optionen und Optionsscheine verbriefen bedingte Termingeschäfte. Der Käufer einer Kauf- beziehungsweise Verkaufsoption (Call bzw. Put) kauft zum Optionspreis vom Stillhalter das Optionsrecht. Dann kann der Käufer, muß aber nicht, das Basisobjekt innerhalb bzw. am Ende der Laufzeit (amerikanische bzw. europäische Option) zum vereinbarten Basispreis kaufen bzw. verkaufen. Das Hedging der Emissionshäuser von Optionsscheinen, d. h. der Aufbau von Gegenpositionen zu verkauften Optionen zur Neutralisierung des Risikos, erfordert möglichst genaue Optionspreismodelle. Im volatilen Jahr 1997 haben viele große Emissionshäuser, z. B. die Citibank und die Schweizerische Bankgesellschaft, auf Grund zu ungenauer Optionspreismodelle (Black/Scholes, Cox/Ross/Rubinstein) Verluste in dreistelliger Millionenhöhe hinnehmen müssen. Durch Beobachtung der Marktpreise und Volumen von tausenden, zehntausenden, oder mehr Optionsgeschäften (viele Handelstage, Börsen aus aller Welt) können Muster zum Training empirischer, neuronaler Optionspreismodelle für verschiedene Basisobjekte erstellt werden.

„Lernverfahren in Feedforward-Netzen“

L. Cromme
Cottbus

Lernverfahren in neuronalen Netzen sind spezielle Optimierungsalgorithmen. Wir gehen auf Wechselwirkungen zwischen beiden Gebieten ein, insbesondere auf Konvergenzaussagen für Verfahren 1. Ordnung (Gradientenverfahren vs. Backpropagation) und auf die Struktur der HESSE-Matrix in Feedforward-Netzen, die für Verfahren 2. Ordnung von großer Wichtigkeit ist.

„Intelligent Control: Von neuronalen Konzepten zur Anwendung“

A. Bühlmeier
Dortmund, Bremen

Die besondere Herausforderung beim Einsatz neuronaler Netze für Steuerungsaufgaben liegt darin begründet, daß ein reales System mit der realen Umwelt interagiert, ohne daß ein Mensch die jeweiligen Ein- und Ausgaben des Systems überwacht. Es ist daher erforderlich, Vorwissen in das System zu integrieren, damit gefahrlose Adaption möglich wird. Ein Lernverfahren, das besonders für 'Intelligent Control' geeignet erscheint, ist das Reinforcement Lernen, da nicht immer erwartet werden kann, daß ein weiteres System (z.B. ein Mensch) zur Verfügung steht, das dem System Lösungen vorgibt. Neuronale Ansätze, die lediglich für Klassifikationsaufgaben entwickelt wurden, können somit nicht in jedem Fall verwendet werden.

Darüber hinaus möchte man dynamische Effekte erzielen, wie das Reagieren auf raumzeitliche Muster. Hierfür müssen dynamische neuronale Modelle verwendet werden, die neben der Langzeitdynamik (Lernen) eine Kurzzeitdynamik besitzen.

Wir geben zunächst einen Überblick über den Einsatz neuronaler Netze für 'Intelligent Control'. Wir zeigen dann an dem Beispiel eines adaptiven Servoassistenten für einen elektrisch angetriebenen und mit Sensorik ausgestatteten Rollstuhl, wie eine von der Biologie inspirierte neuronale Architektur eingesetzt wird. Basis bilden dabei Prinzipien der klassischen und operanten Konditionierung. Die von uns vorgenommene spezielle Umsetzung des Prinzips des Reinforcement Lernens stellt ein interessantes Modell für die Interaktion mit dem Benutzer da: Solange der Anwender keine Systemziele gefährdet, tritt kein Reinforcement auf, und das neuronale Netz beeinflut Benutzerkommandos nicht. Erst durch vermeidbare Fehler des Anwenders reagiert das Netzwerk und greift abhängig von der Wahrscheinlichkeit eines Fehlers in die Steuerung ein.

„Zur Stationarität autoregressiver ANN Modelle“

F. Leisch
Wien

Neurale Netze erfreuen sich großer Beliebtheit zur Modellierung von nichtlinearen, autoregressiven Zeitreihen, d.h., Modellen der Form

$$x_t = g(x_{t-1}, ..., x_{t-p}) + e_t \tag{1}$$

wobei $g()$ eine “beliebige” Funktion ist und e_t ein iid Rauschprozeß. Wenn g linear ist, erhält man ein klassisches AR(p) Modell.

Eine der zentralen Fragen der Zeitreihentheorie ist die Stationarität des Modells, d.h., ob die probabilistischen Strukturen konstant oder zumindest asymptotisch konstant über die Zeit sind. Überraschenderweise wurde diese Frage in der NN Literatur bisher fast nicht behandelt, sondern AR-NN Modelle wie (1) “einfach verwendet”, ohne die Eigenschaften des Modells genauer zu untersuchen.

Wir verwenden Resultate aus der Theorie dynamischer Systeme und Markoffkettentheorie, um zu zeigen, unter welchen Bedingungen AR-NN (asymptotisch) stationär sind. Insbesonders zeigt sich z.B., daß ein klassisches Multi Layer Perceptron ohne Shortcuts nur stationäre Prozesse lernen kann. Sind lineare Shortcuts zwischen Inputs und Outputs des Netzes vorhanden, so bestimmen nur die Gewichte der Shortcuts, ob das Gesamtsystem stationär ist oder nicht.

Es ist daher z.B. nicht möglich, mit einem “normalen” MLP (ohne Shortcuts) eine nichtstationäre Zeitreihe zu lernen, wenn auch über endliche Zeitintervalle zu approximieren. Weiters kann durch klassischen Bedingungen an die Wurzeln des Shortcut-Polynoms getestet werden, ob eine stationäre oder nicht-stationäre Zeitreihe gelernt wurde (die restlichen Gewichte sind für diese Frage irrelevant).

Wenn man als Beispiel Zeitreihen aus der Finanzwirtschaft wie Aktienkurse hernimmt, so heißt dies, daß ein MLP ohne Shortcuts nur die Returns, nicht aber den Kurs selber lernen kann, da derartige Zeitreihen typischerweise integriert und damit nicht stationär sind.

„Lernen in Rekurrenten Neuronalen Netzen?“

T. Kolb
Cottbus

Diese Arbeit beschäftigt sich mit Fragen und Problemen, die durch Einführung von Rückkopplungen in die Netztopologie verschiedener neuronaler Architekturen entstehen. Nach einer Diskussion verschiedener unproblematischer Fälle wird gezeigt, daß in der Praxis verwendete Lernansätze für große Klassen von Architekturen und Anwendungsproblemen zumindest fragwürdig sind. Für den Fall der Folgengenerierung werden Rahmenbedingungen formuliert und Verfahren beschrieben, die es ermöglichen, rekursive Neuronale Netze trotz der obengenannten Schwierigkeiten einzusetzen.

„Aspekte Neuronalen Lernens für die medizinische Analyse von Gehirnaktivitäten anhand von MEG (Magnetenzephalographie) - Daten“

M. Borschbach
Münster

MEG ist ein in der Gehirnforschung etabliertes Verfahren mit dem die Aktivität der Grosshirnrinde zeitlich und räumlich untersucht werden kann. Es wird anhand von synthetischen generierten MEG Daten gezeigt, daß das zugrundeliegende Modell den Voraussetzungen entspricht, die zur Blinden Separation notwendig sind. Mehrere Neuronale Lernalgorithmen mit einer 2-Schichtnetzstruktur werden dabei auf ihre Leistungsfähigkeit untersucht. Die Resultate werden erörtert und ein Ausblick auf weitere Forschung gegeben.

„Die Wertigkeit des Corticotropin-Releasing-Hormon (CRH)-Tests im Vergleich zum Insulinhypoglykämietest in Bezug auf die Beurteilung der Hypothalamus-Hypophysen-Nebennierenrinden-Achse bei wachstumsgestörten Patienten“

B. Herm
Leipzig

Das System der Hypothalamus-Hypophysen-Nebennierenrinden-Achse dient der Regulation der Ausschüttung der Nebennierenrindenhormone in Anpassung an den körperlichen Bedarf. Für eine ausreichende Streßadaption des Körpers ist die Funktionsfähigkeit der Hypothalamus-Hypophysen-Nebennierenrinden-Achse von großer Bedeutung. Die Beurteilung der Funktionsfähigkeit ist jedoch dadurch erschwert, daß eine normale Alltagsfunktion eine Minderfunktion in körperlichen Streßsituationen (Erkrankungen, Traumen, Operationen etc.) nicht ausschließt. Diese kann sich in einem akuten Kreislaufzusammenbruch manifestieren und über Koma bis hin zum Tod führen. Da bei wachstumsgestörten Kindern eine zusätzliche Störung der Hypothalamus-Hypophysen-Nebennierenrinden-Achse auftreten kann, empfiehlt sich eine unbedingte Abklärung der Funktionsfähigkeit derselben. Zur Diagnostik und Differenzierung von Störungen der Hypothalamus-Hypophysen-Nebennierenrinden-Achse werden derzeit mehrere Funktionstests eingesetzt, die geeignet sind, die korrekte Funktion von verschiedenen Elementen oder Systemteilen oder auch des gesamten Systems zu prüfen. Dazu gehören auch der Insulinhypoglykämietest und der CRH-Test. Die Insulin-induzierte Hypoglykämie ist trotz ihrer Gefährlichkeit und Nachteile bisher das Standardverfahren zur Beurteilung der Funktionsfähigkeit der Hypothalamus-Hypophysen-Nebennierenrindenachse. Der Corticotropin-Releasing-Hormon-Test jedoch wird derzeit hauptsächlich zur Differentialdiagnose des Cushing-Syndromes eingesetzt.

Ziel dieser Arbeit ist es, die Aussagen des Insulin-Hypoglykämie-Testes und des Corticotropin-Releasing-Hormon-Testes hinsichtlich der Hypothalamus-Hypophysen-Nebennierenrinden-Achse zu vergleichen, um herauszufinden, ob eine Einstufung wachstumsgestörter Kinder im Hinblick auf das Risiko einer mangelnden Streßadaption ausschließlich mit dem CRH-Test möglich ist. Die Folge wäre der wünschenswerte Verzicht auf den Insulinhypoglykämietest als diagnostisches Hilfsmittel. Es wurden die Daten von wachstumsgestörten Kindern erfaßt, bei denen in den vergangenen 10 Jahren aus diagnostischen Gründen mindestens einer der o.g. Tests durchgeführt wurde. Unter Verwendung künstlicher Neuronaler Netze soll der Korrelationsgrad zwischen den Testdaten ermittelt werden.

Nach bisherigen Trainingsläufen wurde feststellt, daß beide Testdatenreihen einen unterschiedlichen Informationsgehalt besitzen. Jedoch erscheint es wahrscheinlich, daß sich aus dem Kurvenverlauf des CRH-Testes und dem zugehörigem Beurteilungsschema der Befund des Insulintestes vorhersagen läßt. Dies könnte aus medizinischer Sicht ggf. ausreichen, um auf den Insulintest zur Beurteilung der o.g. Achse zu verzichten.

„Nutzung neuronaler Netze zur Lösung dynamischer Problemstellungen bei turmartigen Bauwerken"

B. Beirow
Cottbus

Turmartige Bauwerke sind im besonderen Maße Wind- und Eiseinwirkungen ausgesetzt, sei es durch Extrem- oder Dauereinwirkungssituationen, und damit aufgrund ihrer Schlankheit schwingungsanfällig. Die Kenntnis des Schwingverhaltens, d.h. der Zusammenhang zwischen beliebiger Einwirkung und Systemantwort, ist hinsichtlich der Beurteilung der Nutzungsfähigkeit von besonderem Interesse. Bei bekanntem mechanischen System kann mittels herkömmlicher Verfahren die Systemantwort im Falle bekannter Erregung berechnet werden. In der Regel sind die Einwirkungen jedoch stochastischer Natur, Erregermechanismen und -arten mischen sich. Der Einsatz dieser Methoden ist daher u. U. mit erheblichen Schwierigkeiten verbunden. Da zur Beurteilung der Güte der Rechnungen ohnehin eine Vergleichsbasis in Form von Wind- (Einwirkung) und Schwingungsmessungen (Antwort) geschaffen werden muß, stellt deren Nutzung zum Training eines neuronalen Netzes eine zweckmäßige Alternative zur Ableitung des angesprochenen Zusammenhanges dar. Dieses hybride Verfahren gewinnt vor allem aus den Tatsachen an Bedeutung, daß die Kenntnis des mechanischen Systems nicht notwendig ist und sich neuronale Netze zur Abbildung unscharfer Zusammenhänge besonders eignen. Im April 1997 wurde durch den Lehrstuhl für Statik und Dynamik der BTU Cottbus ein den spezifischen Bedingungen von Fernseh- und Fernmeldetürmen angepaßtes Monitoringsystem im Fernmeldeturm (FMT) Cottbus installiert, welches permanent Schwinggeschwindigkeiten aufgezeichnet. Parallel erfolgt durch den Lehrstuhl für angewandte Physik die meßtechnische Erfassung des Windes. Die Meßdaten liegen jeweils in Form von 10-Minutenmittelwerten vor. Durch Vorklassierungen der Meßdaten werden die jeweiligen Erregermechanismen berücksichtigt. Die klassierten Daten werden dann zum Training eines feed forward-Netzes (im einfachsten Fall 2 Eingänge mit Windmeßdaten, 2 Ausgänge mit Schwingungsmeßdaten) nach dem Lernverfahren Backpropagation mit Momentumterm verwendet. Die Anwendung des Verfahrens wurde bisher bei Auftreten wirbelinduzierter Schwingungen erfolgreich getestet. Ein weiteres Anwendungsgebiet neuronaler Netze in der Baudynamik ist im Bereich der Identifikation globaler Bauwerksänderungen angesiedelt. Als praktisches Beispiel dient auch wieder der FMT Cottbus. Hier wird allerdings die Kenntnis des mechanischen Systems vorausgesetzt. Als Indikatoren für die angesprochenen Änderungen können - meßtechnisch erfaßbare - Änderungen modaler Parameter (z. B. Eigenfrequenzen) fungieren. Durch Simulationsrechnungen mit Hilfe eines um Schadensmodelle erweiterten Rechenmodells muß ein Musterkatalog erstellt werden, der möglichen Änderungen nach Ort und Intensität (lokale Steifigkeitsänderungen) die entsprechenden geänderten modalen Parameter zuordnet. Die Indentifikationsprozedur besteht nun darin, rückwärts aus gemessenen Änderungen modaler Parameter die zugehörigen Bauwerksänderungen im Rahmen des verwendeten Modells zu diagnostizieren. Diese Vorgehensweise würde allerdings einen sehr großen Musterkatalog erfordern, dessen Erstellung durch Simulationsrechnungen allein zu aufwendig wäre. Durch Verwendung einer geringen Anzahl ausgewählter Simulationsrechnungen für das Training eines neuronalen Netzes (feed forward, Backpropagation mit Momentum term) gelingt die Ableitung des Zusammenhangs für beliebige Eingangsgrößen mit erheblich reduziertem Aufwand.

„Hybride Systeme“

W.-M. Lippe
Münster

In den letzten dreißig Jahren wurde eine Reihe von Konzepten zur Informationsverarbeitung entwickelt, die sich an biologischen Vorbildern orientieren. Hierzu gehören vor allem

– Künstliche Neuronale Netze

  Aufbauend auf den grundlegenden Arbeiten von McCulloch, Pitts, Rosenblatt und anderen, sind hier inzwischen verschiedene, teilweise sehr unterschiedliche, Modelle entstanden. Allen Modellen gemeinsam ist ein Netz einfacher Knoten, den Neuronen, zusammen mit einer Lernregel, die es dem Netz ermöglichen, sich selbständig – in Abhängigkeit von präsentierten Trainingsdaten – zu modifizieren. Dieses Training mit Beispieldaten wird so lange durchgeführt, bis das Netz über gewünschte Eigenschaften verfügt. Neben der Modellbildung für biologische und medizinische Untersuchungen finden Künstliche Neuronale Netze inzwischen eine breite kommerzielle Anwendung, vor allem in Bereichen, wie Mustererkennung, Klassifikation oder Prozeßsteuerung.

– Genetische Algorithmen

  Genetische Algorithmen sind stochastische Such- und Optimierungsverfahren, die ihre Motivation aus ihrer Ähnlichkeit zu den Prozessen der biologischen Evolution beziehen. Wie beim biologischen Vorbild gibt es zu jedem Zeitpunkt eine Population von "Individuen", wobei dem Individuum über eine Güte- oder Fitneß-Funktion ein Güte- oder Fitneßgrad in bezug auf die zu optimierende Zielfunktion zugeordnet wird. Ziel der genetischen Algorithmen ist es, im Verlauf des Verfahrens immer bessere Individuen (Individuen mit besserer Fitneß) zu erzeugen. Die Erzeugung neuer Individuen geschieht dadurch, daß die besten Individuen über einen stochastischen Auswahlprozess proportional zu ihrer relativen Fitneß zur Bildung von Nachkommen selektiert werden. Da das Konzept der genetischen Algorithmen relativ neu ist – nach ersten Ansätzen von Vogel, Owens und Walsh 1996, erfolgte eine systematischere Untersuchung erst ab Ende der siebziger Jahre – sind hier kommerzielle Einsätze erst in neuester Zeit zu beobachten.

– Fuzzy-Logik

  Fuzzy-Logik wurde ab Mitte der sechziger Jahre von L. Zadeh entwickelt. Es kann als eine Verallgemeinerung der klassischen Logik aufgefaßt werden. Sie dient zur Modellierung unscharfen Wissens, wie sie auch in der menschlichen Ausdrucksweise verwendet wird. Man kann mit ihr linguistische Terme, wie "fast", "klein", "mittel", "groß" oder "ziemlich" modellieren. Die praktische Anwendung findet sich vor allem in der Prozeßsteuerung in Form von Fuzzy-Controllern. Hierbei handelt es sich um eine Weiterentwicklung von klassischen Controllern unter Verwendung von unscharfen Regeln.

In neuerer Zeit sind diese zunächst unabhängigen Konzepte zunehmend zusammengewachsen. Aber auch z. B. bei Künstlichen Neuronalen Netzen findet man immer mehr Systeme, die aus einer Kombination unterschiedlicher Modelle bestehen. Derartige Kombinationen bezeichnet man als "Hybride Systeme".

In dem Vortrag werden zunächst die unterschiedlichen Konzepte vorgestellt und insbesondere ihre typischen Eigenschaften (Vor- und Nachteile) analysiert. Danach werden die prinzipiellen Kombinationsmöglichkeiten und der Stand der aktuellen Forschungen auf diesem Gebiet präsentiert. Beispielhaft wird anschließend etwas detaillierter auf die Optimierung von Parametern bei einem speziellen Backpropagation-Verfahren (Kombination von Momentum-Variante und $\bar{\delta}$-δ-Regel) mit Hilfe eines Fuzzy-Controllers bzw. auf die Optimierung von Fuzzy-Controllern mittels Neuronaler Netze eingegangen.

„Interaktiv lernende Diagnosesysteme für Industrieanwendungen“

P. Feucht
Karlsruhe

Im Forschungszentrum Informatik in der Gruppe IDS werden seit Jahren Diagnosesysteme für industrielle Anwendungen entwickelt. Hierbei werden vor allem Systeme betrachtet, welche von einer großen Menge von verrauschten und fehlerbehafteten Massendaten ausgehen. Das für den Diagnoseprozeß wichtige Bereichswissen liegt meist nur unstrukturiert und unvollstdndig und in unterschiedlichen Wissenrepräsentationen vor. Regelwissen und analytisches Wissen ist oft nur partiell vorhanden, die vom menschlichen Experten erworbenen Diagnosefähigkeiten können von diesem oft nur in Form von Beispielen akquiriert werden.

In der Präsentation soll ausführlich auf ein Konzept für eine hybride Architektur eingegangen werden, die die Integration und Verarbeitung unterschiedlicher Wissenrepräsentationen erlaubt. Die Basis der Architektur bildet ein interaktiv eingelernten Zustandsautomaten, dessen Zustdnden dann spezielle Klassifikatoren zugeordnet werden können. Mit Hilfe des Automaten werden die Massendaten so aufbereitet werden, daß anschließend Klassifikationsalgorithmen sinnvoll eingesetzt werden können.

Anhand von einem Beispiel aus dem industriellen Umfeld – Lackierdiagnose – wird gezeigt, wie mit Hilfe des Verfahrens ein leistungsfähiges Diagnosesystem aufgebaut werden kann.

„Minimierung der Energiebezugskosten in Kläranlagen durch Optimierung der Energieversorgung mittels Kostenmodell und dem Einsatz künstlicher Intelligenz neuronaler Netze“

T. Adam
Cottbus

Zentrale Kläranlagen reinigen durch eine Vielzahl von Verfahrensstufen industrielle und gewerbliche Abwässer einer, vom Gesetzgeber vorgeschriebenen, hohen Qualität. Die dafür notwendigen Aggregate wie Pumpen, Belüftungsanlagen und elektrische Antriebe zeichnen sich durch einen hohen Energiebedarf aus. Dies verursacht Energiekosten, die 1/5 der Betriebsaufwendungen einer Abwasserreinigungsanlage ausmacht.

Eine sinnvolle Möglichkeit der Betreiber von Klärwerken, auf die Betriebskosten unmittelbar einzuwirken, bietet die Energieversorgung. Durch den innerbetrieblichen Anfall von Faulgas bei anaerober Schlammbehandlung und dessen Einsatz in der Heizzentrale, bestehend aus Blockheizkraftwerken (BHK) und Heizkesseln (HK), leistet es einen Beitrag zur thermischen und elektrischen Bedarfsdeckung. Die daraus resultierende Optimierungsaufgabe, aus den vielfältigen Möglichkeiten der autarken und netzbezogenen Energieversorgung die Energiebezugskosten (Arbeit) zu minimieren, wird anhand dieses Kostenmodells dargestellt.

Daneben lassen sich durch künstliche neuronale Netze elektrische Lastspitzen, welche vor allem in Koppelung äußerer (Wetter) und innerer (Betrieb) Ursachen stochastisch auftreten, prognostizieren. Diese Kenntnis ermöglicht dem Betreiber durch geeignete Maßnahmen im Rahmen des Lastmanagements den Netzbezug zu vergleichmäßigen und somit die Leistungskosten zu minimieren.

Die Synthese beider Maßnahmen sind Ziel einer innerbetrieblichen Optimierung der Energieversorgung in Klärwerken und sollen im Vortrag methodisch und exemplarisch dargestellt werden.

„Hybrid-Neuronale Systeme und ihre Anwendung zum Data Mining“

A. Ultsch
Marburg

Als ein Hauptunterschied zwischen Konnektionistischen Modellen und Technologien der symbolischen künstlichen Intelligenz kann die Repräsentation von Wissen in subsymbolischer vs. symbolischer Form angesehen werden. Als eigentlich hybride Systeme werden von uns solche angesehen, die in irgendeiner Form einen Übergang zwischen den unterschiedlichen Repräsentationsformen realisieren. Die Arbeit stellt eine Einordung der verschiedenen hybriden Ansätze vor.

Weiterhin wird die Möglichkeit des Findens bislang unbekannter Strukturen in Datensammlungen (Data Mining /Knowqledge Discovery) mit solchen Systemen aufgezeigt. Von eigenen Erfahrungen mit praktischen (industriellen) Anwendungen solcher Systeme bei Produktionsprozessen, Banken und in der medizinischen Forschung wird berichtet.

„Parameteradaption hydrodynamisch-numerischer Modelle mit neuronalen Methoden'"

O. Schleider
Cottbus

Wichtig für den Betrieb hydrodynamisch-numerischer Modelle ist die Wahl geeigneter Parameter abhängig von der Problemstellung und von Faktoren wie der gewählten Orts- und Zeitdiskretisierung. Zu unterscheiden sind Parameter des numerischen Verfahrens und Parameter der dem hydrodynamischen Modell zugrunde liegenden Differentialgleichungen. Im Beitrag wird sich auf letztere konzentriert.

Das Einstellen von Simulationsparametern (Kalibrierung) bedeutet meist die Durchführung einer Reihe von Experimenten mit geschätzten oder empirischen Parameterwerten. In vielen Fällen resultiert ein Ingenieuransatz in solchen Versuchsreihen. Bei komplexeren Simulationsaufgaben etwa mit mehrdimensionalen Parametervektoren und deshalb aufwendiger manueller Parametereinstellung bietet sich jedoch die Verwendung von automatischen (wie z.B. neuronalen) Verfahren an.

Der Beitrag untersucht zunächst zwei Möglichkeiten der Anwendung neuronaler Methoden zur Parameteradaption hydrodynamisch-numerischer Modelle. Dann wird ein kombinierter Ansatz zur Parameteradaption durch ein neuronales Netz während der Laufzeit des numerischen Modells entwickelt. An einem konkreten Beispiel wird die Anwendung des kombinierten Ansatzes vorgestellt.

Die Parameteradaptionsaufgabe kann als Zuordnung bestimmter Parameterwerte zu gegebenen Charakteristika der Simulationsaufgabe angesehen werden. Ein entsprechendes neuronales Netz muß die unbekannte Funktion, die die Zuordnung beschreibt, erlernen. Dazu werden backpropagation-Netze eingesetzt. Problematisch ist hier die entsprechende Identifizierung charakteristischer Werte der Simulationsaufgabe.

Weiterhin kann die Parametereinstellung on-line erfolgen, dann steht dem Netz als weitere Information die Historie früher gewählter Parameterwerte und die entsprechende Reaktion des Simulationssystems durch einen entsprechend zu wählenden Fehlerwert zur Verfügung. Dafür werden rückgekoppelte neuronale Netze verwendet. Im entwickelten Verfahren für online Adaption wird ein rückgekoppeltes neuronales Netz simultan mit den Rechenschritten des numerischen Modells eingesetzt. Der hydrodynamische Prozess wird durch die Lösung eines Systems partieller Differentialgleichungen simuliert, deren Parameter mit dem neuronalen Netz approximiert werden. Anstatt die Simulationsläufe mit angepaßten Parametern zu wiederholen, berechnet ein für zeitabhängige Berechnungen geeignetes neuronales Netz während der Laufzeit die aktuellen Parameter.

Die beispielhafte Anwendung verwendet ein rückgekoppeltes neuronales Netz und ein numerisches Verfahren zur Lösung der vertikal integrierten 2D-Flachwassergleichungen. Der zu identifizierende Parameter ist der Bodenreibungskoeffizient. Die Meßwertewerden durch eine Referenzrechnung ersetzt. Das kombinierte System benutzt die Randbedingungen der Referenzrechnung für den numerischen Teil und muß den gewählten Parameter identifizieren.

Hinweise für Autoren

Da der Verlag von uns den Tagungsband als Postscript-Datei erhält und diese unverändert in Druck geht, wollen wir den Band komplett in LaTeX setzen. Dazu haben wir je zwei Style-Files für die Verwendung unter LaTeX2.09 bzw. unter LaTeX2e entwickelt.

LaTeX2.09: Der Style `cowanbook.sty` ist als Dokumentstil, das Style-File `cowan.sty` als letzte Dokumentstiloption zu verwenden.

LaTeX 2_ε: Der Style `cowan-book.cls` ist als Dokumentstil, das Style-File `cowan.sty` als letzte Dokumentstiloption zu verwenden.

Die beiden Varianten sollten zu identischen Druckresultaten führen. Der Tagungsband selbst wird mit LaTeX 2_ε erzeugt. Da der Tagungsband im Format A5 erscheinen wird und die Druckseiten deshalb um etwa ein Drittel verkleinert werden, sollte eine Standardschriftgröße von 12pt verwendet werden. Sie erleichtern uns die Arbeiten bei der Erstellung des Tagungsbandes, wenn Sie für Ihren Beitrag zudem die folgenden Hinweise beachten:

Die folgenden Felder sollten nach `\begin{document}` belegt werden:

`\title{...}`	Titel des Beitrages
`\author{...}`	Autoren incl. Einrichtung, Adresse, email-Adresse
`\maketitle`	Erzeugung des Titels
`\shorttitle{...}`	Kurzform des Titels incl. Autoren für Seiten-Header
`\contentstitle{...}`	Titel incl. Autoren fuer das Inhaltsverzeichnis

`\begin{abstract}`...Abstract des Beitrages...`\end{abstract}`

Ansonsten können alle notwendigen Dokumentstiloptionen (`epsf` etc.) sowie gewohnten LaTeX-Befehle und -Umgebungen für den Textteil verwendet werden. Sollten Sie nicht über die Möglichkeiten zur Erstellung mittels LaTeXoder zum Einbinden der Style-Files verfügen, setzen Sie sich bitte mit der Tagungsleitung in Verbindung.

Hinweise für Autoren

Es folgt die Struktur eines Beispielbeitrages:

```
\documentclass[12pt]{cowan-book}
\usepackage{german,cowan}

\begin{document}

  \title{Optimierte Lernverfahren f"ur Feedforward-Netze}
  \author{J.~Wille, T.~Kolb\\
          Brandenburgische Technische Universit"at Cottbus\\
          Lehrstuhl f"ur Numerische und Angewandte Mathematik\\
          Karl-Marx-Stra"se 17, D-03046 Cottbus\\
          email: wille@math.tu-cottbus.de, kolb@math.tu-cottbus.de}
  \maketitle

  \shorttitle{J.~Wille, T.~Kolb: Optimierte Lernverfahren}
  \contentstitle{J.~Wille, T.~Kolb: Optimierte Lernverfahren
                f"ur Feedforward-Netze}

  \begin{abstract}
    Auf dem Feld ... des Standardverfahrens gegeben.
  \end{abstract}

  \section{Mathematische Beschreibung}
  \subsection{Problemstellung}
   W"ahrend der Lernphase ... basierenden Gewichts"anderungen.

  \begin{thebibliography}{[00]}
    \bibitem[1]{rum86} McClelland, J. \& Rumelhart ...
  \end{thebibliography}

\end{document}
```

Die (verkleinerte) LaTeX-Ausgabe dieses Beispiels zeigt die nachfolgende Seite.

Optimierte Lernverfahren für Feedforward-Netze

J. Wille, T. Kolb
Brandenburgische Technische Universität Cottbus
Lehrstuhl für Numerische und Angewandte Mathematik
Karl-Marx-Straße 17, D-03046 Cottbus
email: wille@math.tu-cottbus.de, kolb@math.tu-cottbus.de

Zusammenfassung

Auf dem Feld der Beschäftigung mit den verschiedensten Typen Neuronaler Netze nehmen die Feedforward-Netze eine exponierte Stellung ein. Seit der Propagierung des Backpropagation-Algorithmus durch Hinton, Rumelhart et.al. [1] hat es eine Flut von Verbesserungen und Modifikationen des Standardverfahrens gegeben.

...

1 Mathematische Beschreibung

1.1 Problemstellung

Während der Lernphase werden dem Netz Eingabevektoren präsentiert, welche intern in einem mehrstufigen Vorgehen schichtweise verarbeitet werden und es wird eine Netzausgabe berechnet. Das Lernen, d.h. die zielgerichtete Veränderung der Netzstruktur, basiert auf der Rückvermittlung des Ausgabefehlers, der durch Abgleich der Netzausgabe mit der zugehörigen Sollausgabe bestimmt wird. Die Adaption erfolgt mittels auf diesem Fehler basierenden Gewichtsänderungen.

...

Literaturverzeichnis

[1] McClelland, J. & Rumelhart, D.E.: Parallel Distributed Processing – Explorations in the Microstructure of Cognition – Volume 1: Foundations. The MIT Press, Cambridge Massachusetts, 1986