

Berichte aus der Physik

Christoph Bunzmann

Effizientes Online-Training neuronaler Netzwerke

Shaker Verlag
Aachen 2003

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Bunzmann, Christoph:

Effizientes Online-Training neuronaler Netzwerke/Christoph Bunzmann.

Aachen : Shaker, 2003

(Berichte aus der Physik)

Zugl.: Würzburg, Univ., Diss., 2002

ISBN3-8322-1502-6

Copyright Shaker Verlag 2003

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 3-8322-1502-6

ISSN 0945-0963

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: www.shaker.de • eMail: info@shaker.de

Effizientes Online-Training neuronaler Netzwerke

Christoph Bunzmann

Das Lernen aus Beispielen ist eine Fähigkeit, die man bei Menschen und Tieren beobachtet. Es erfolgt durch die Anpassung von Denk- und Verhaltensweisen an Gesetze, die aus Beispielen in Form von konkreten Erlebnissen gewonnen werden. Eine mathematische Beschäftigung mit diesem Phänomen zeigt Modellsituationen auf, in denen die wesentlichen Eigenschaften dieses komplexen Vorgangs durch relativ einfache mathematische Modelle beschrieben werden können. Während die anfangs biologisch inspirierten Studien bald in technische Anwendungen und mathematische Analysen mündeten, blieben diese doch der informationsverarbeitenden Einheit treu, mit der alles angefangen hatte: dem Neuron – nun als mathematisches Modell einer Informationsverarbeitung durch einfache Einheiten.

Inzwischen existiert eine Vielzahl von Algorithmen, um komplizierte neuronale Netzwerke an Beispieldaten anzupassen, d.h. zu „trainieren“. Der begrenzende Faktor ist dabei oft, dass nicht genügend Beispiele zur Verfügung stehen, um gute Lernergebnisse zu erzielen. Diese Arbeit präsentiert – nach einer kurzen Einführung in die Problemstellung – zwei Ansätze, um mit einer begrenzten Anzahl von Beispielen ein neuronales Netzwerk an die Regel hinter den Daten anzupassen.

Besondere Aufmerksamkeit wird Problemen beim Training von Netzwerken mit zwei informationsverarbeitenden Schichten gewidmet. Als Modell solcher Netzwerke wird in dieser Arbeit die sogenannte Komitee-Maschine gewählt. Betrachtet man die Anzahl der Beispiele, die pro Freiheitsgrad des Netzwerkes benötigt werden, um ein gutes Lernergebnis zu erzielen, so wächst diese bei herkömmlichen Lernverfahren mit der Dimension der Eingabeschicht. Die Effizienz des Trainings sinkt also mit der steigenden Anzahl der Eingabeparameter.

Die Arbeit präsentiert einen Algorithmus zum Training der Komitee-Maschine, der dieses Problem löst. Dabei wird die schrittweise Optimierung des Netzwerkes durch ein Lernen mit zwei grundsätzlich verschiedenen Abschnitten ersetzt: Vor der Anwendung der schrittweisen Optimierung wird die Dimension des Lernproblems reduziert. Die Lösung des Optimierungsproblems kann damit in einem niedrig dimensionalen Raum gesucht werden. Nachdem die technische Umsetzung des Algorithmus demonstriert wurde, wird im Limes einer unendlich dimensional Eingabeschicht dessen Effizienz theoretisch analysiert.

Ein Kapitel widmet sich der Analyse des Lernens mit Zurücklegen, d.h. unter mehrmaliger Verwendung derselben Beispiele. In diesem Fall besteht das Problem darin, den Endzustand des Lernens bezüglich der relevanten Größen zu charakterisieren. Diese werden entscheidend von der Größe der Optimierungsschritte beeinflusst. Während Methoden existieren, die den Einfluss der Schrittgröße auf das Lernergebnis wiedergeben, sind diese angesichts der recht einfachen Problemstellung überraschend aufwändig. Hier wird ein Ansatz vorgestellt, mit dem ohne den Rückgriff auf aufwändige mathematische Methoden der Zusammenhang zwischen Schrittgröße und der Verallgemeinerungsfähigkeit des Netzwerkes erfasst wird.