

Berichte aus der Informatik

Andreas Rudolph

**Data Mining in action:
Statistische Verfahren der Klassifikation**

Shaker Verlag
Aachen 1999

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Rudolph, Andreas:

Data Mining in action: Statistische Verfahren der Klassifikation /
Andreas Rudolph. - Als Ms. gedr. - Aachen : Shaker, 1999
(Berichte aus der Informatik)

ISBN3-8265-6172-4

Copyright Shaker Verlag 1999

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

Als Manuskript gedruckt. Printed in Germany.

ISBN 3-8265-6172-4

ISSN 0945-0807

Shaker Verlag GmbH • Postfach 1290 • 52013 Aachen
Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9
Internet: www.shaker.de • eMail: info@shaker.de

Das vorliegende Buch wendet sich nicht an Hochschullehrer, die besonders elegante Beweisführungen suchen. Es wendet sich aber an diejenigen, die ein praktisches Problem des Data Mining mit einem adäquaten und geeigneten Verfahren angehen möchten und gegebenenfalls das entsprechende Verfahren auch auf den Rechner bringen wollen, um die Qualität ihrer Arbeit zu steigern. In diesem Sinne ist das Buch eine Kommentierung von Verfahren und Sätzen aus der Statistik im Hinblick auf ihre Anwendbarkeit in der Praxis und versucht deshalb, eher die mathematische Intuition des Lesers anzusprechen.

Von daher wendet sich das Buch an

- Studenten der Stochastik, die sich auf das spätere Berufsleben vorbereiten wollen
- Praktiker mit gewissem mathematischen Background, die sich informieren wollen
- Informatiker, die die mathematischen Verfahren kennenlernen wollen

Die Verfahren, die in diesem Buch angesprochen werden, umfassen

- Clusteranalysen, insbesondere partitionierende wie auch hierarchische Verfahren
- Verfahren der Klassifikation, insbesondere parametrische wie auch nicht-parametrische Diskriminanzanalysen, neuronale Netze, logistische Regressionsmodelle wie auch Entscheidungs- und Regressionsbäume

In jedem Fall wird versucht, eine Abgrenzung der Verfahren vorzunehmen und ihre Vor- und Nachteile gegeneinander abzuwägen.