

Ulrich Kathöfer

# **Data Mining mit Genetischen Algorithmen**

Logos Verlag Berlin

**i**

λογος

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Motivation . . . . .	3
1.2	Überblick . . . . .	5
<b>2</b>	<b>Codierung und Operatoren für Genetische Algorithmen</b>	<b>7</b>
2.1	Genetische Algorithmen . . . . .	8
2.1.1	Das natürliche Vorbild . . . . .	8
2.1.2	Der Grundalgorithmus . . . . .	9
2.1.3	Codierung/Evaluierung . . . . .	11
2.1.4	Genetische Operatoren . . . . .	14
2.1.5	Warum funktioniert ein genetischer Algorithmus? . . . . .	18
2.1.6	Anwendungen . . . . .	22
2.2	Grundgerüst/Java-Programmierung . . . . .	22
2.3	Erweiterungen bei binärer Codierung . . . . .	36
2.3.1	Übersicht . . . . .	36

---

2.3.2	Einsatz der Selektion . . . . .	37
2.3.3	Diskussion der klassischen Crossover-Techniken . . . . .	39
2.3.4	Mutation . . . . .	41
2.3.5	Verbesserungsoperatoren . . . . .	43
2.4	Genetische Algorithmen mit „reellen Genen“. . . . .	47
2.5	Genetische Algorithmen und kombinatorische Probleme . . . . .	53
2.6	Zusammenfassung . . . . .	58
<b>3</b>	<b>Anwendung von Genetischen Algorithmen im Data Mining</b>	<b>61</b>
3.1	Data Mining . . . . .	61
3.2	Clusteranalyse . . . . .	65
3.2.1	Fragestellung und übliche Verfahren. . . . .	65
3.2.2	Anwendung von Genetischen Algorithmen. . . . .	75
3.2.3	Bewertung . . . . .	108
3.3	Klassifizierung/Diskriminanzanalyse . . . . .	110
3.3.1	Klassische Verfahren. . . . .	110
3.3.2	Anwendung von Genetischen Algorithmen. . . . .	120
3.3.3	Bewertung. . . . .	135
3.4	Regelableitung . . . . .	137
3.4.1	Problemstellung und herkömmliche Lösungsmethoden . . . . .	137
3.4.2	Anwendung von Genetischen Algorithmen. . . . .	141

---

3.4.3	Bewertung . . . . .	158
<b>4</b>	<b>Zusammenfassung und Ausblick</b>	<b>159</b>
<b>A</b>	<b>Programmcode</b>	<b>163</b>
A.1	Cluster-Algorithmus 1. . . . .	164
A.2	Cluster-Algorithmus 2. . . . .	176
A.3	Diskriminanz-Algorithmus binär. . . . .	186
A.4	Diskriminanz-Algorithmus reell. . . . .	195
A.5	Regelableitungs-Algorithmus. . . . .	202
A.6	Weitere Programmteile . . . . .	210
<b>B</b>	<b>Testdaten</b>	<b>213</b>
B.1	Kredit-Scoring-Daten. . . . .	213
B.2	Generierte Beispieldaten. . . . .	231
<b>C</b>	<b>Testläufe</b>	<b>235</b>
C.1	Cluster-Algorithmus 1. . . . .	236
C.2	Cluster-Algorithmus 2. . . . .	244
C.3	Diskriminanz-Algorithmus binär. . . . .	251
C.4	Diskriminanz-Algorithmus reell. . . . .	256
C.5	Regelableitungs-Algorithmus. . . . .	261