

---

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b> .....	<b>1</b>
1.1	Daten .....	1
1.2	Data Mining, Data Analytics und Knowledge Discovery .....	2
	Literatur .....	3
<b>2</b>	<b>Daten und Relationen</b> .....	<b>5</b>
2.1	Der Iris-Datensatz .....	5
2.2	Maßskalen .....	6
2.3	Mengen- und Matrixdarstellung .....	10
2.4	Relationen .....	11
2.5	Unähnlichkeitsmaße .....	12
2.6	Ähnlichkeitsmaße .....	14
2.7	Sequenzrelationen .....	16
2.8	Abtastung und Quantisierung .....	18
	Literatur .....	22
<b>3</b>	<b>Datenvorverarbeitung</b> .....	<b>23</b>
3.1	Fehlerarten .....	23
3.2	Behandlung fehlerhafter Daten .....	26
3.3	Filterung .....	27
3.4	Datentransformationen .....	32
3.5	Datenintegration .....	35
	Literatur .....	36
<b>4</b>	<b>Datenvisualisierung</b> .....	<b>37</b>
4.1	Diagramme .....	37
4.2	Hauptkomponentenanalyse .....	39
4.3	Mehrdimensionale Skalierung .....	43
4.4	Sammon-Abbildung .....	47
4.5	Auto-Assoziator .....	51
4.6	Histogramme .....	51

4.7	Spektralanalyse .....	54
	Literatur .....	58
<b>5</b>	<b>Korrelation .....</b>	<b>59</b>
5.1	Lineare Korrelation .....	59
5.2	Korrelation und Kausalität .....	61
5.3	Chi-Quadrat-Test auf Unabhängigkeit .....	63
	Literatur .....	65
<b>6</b>	<b>Regression .....</b>	<b>67</b>
6.1	Lineare Regression .....	67
6.2	Lineare Regression mit nichtlinearer Substitution .....	71
6.3	Robuste Regression .....	72
6.4	Neuronale Netze .....	72
6.5	Radiale Basisfunktionen .....	76
6.6	Kreuzvalidierung .....	78
6.7	Merkmalsselektion .....	80
	Literatur .....	82
<b>7</b>	<b>Prognose .....</b>	<b>83</b>
7.1	Endliche Zustandsautomaten .....	83
7.2	Rekurrente Modelle .....	85
7.3	Autoregressive Modelle .....	86
	Literatur .....	87
<b>8</b>	<b>Klassifikation .....</b>	<b>89</b>
8.1	Klassifikationskriterien .....	89
8.2	Naiver Bayes Klassifikator .....	93
8.3	Lineare Diskriminanzanalyse .....	96
8.4	Supportvektormaschine .....	98
8.5	Nächster-Nachbar-Klassifikator .....	100
8.6	Lernende Vektorquantisierung .....	101
8.7	Entscheidungsbäume .....	102
	Literatur .....	106
<b>9</b>	<b>Clustering .....</b>	<b>109</b>
9.1	Clusterpartitionen .....	109
9.2	Sequenzielles Clustering .....	111
9.3	Prototypbasiertes Clustering .....	114
9.4	Fuzzy-Clustering .....	116
9.5	Relationales Clustering .....	121
9.6	Clustertendenz .....	125
9.7	Clustervalidität .....	126

---

9.8 Selbstorganisierende Karte .....	127
Literatur .....	129
<b>Anhang: Optimierungsverfahren .....</b>	<b>131</b>
<b>Lösungen der Übungsaufgaben .....</b>	<b>135</b>
<b>Sachverzeichnis .....</b>	<b>139</b>