Bernd Heesen

Künstliche Intelligenz und Machine Learning mit R

Anwendungen im Bereich Business Analytics

Springer Gabler

Inhaltsverzeichnis

1 Nutzen von Machine Learning und Künstlicher Intelligenz 1
1.1 Künstliche Intelligenz 3
1.2 Menschliche Intelligenz 4
1.3 Rationale Entscheidungen 5
1.4 Entscheidungsautomation 7
1.5 Entscheidungsprozess bei Mensch und Maschine 9
1.6 Geschwindigkeit der Entscheidungsfindung 10
1.7 Datenbasis für Entscheidungen 12
1.8 Business Intelligence 14
1.9 Innovation 15
Literatur 16

2 Machine Learning 19
2.1 Überwachtes Lernen 21
2.2 Unüberwachtes Lernen 22
2.3 Bestärkendes Lernen 23
Literatur 24

3 Best Practices 25
3.1 Datenqualität 27
3.2 Visualisierung 29
3.3 Entwicklungsumgebung 31
3.4 Programmierung 32
Literatur 33

4 Einrichtung der Entwicklungsumgebung von R 35
4.1 R installieren 35
4.2 Rtools installieren 42
4.3 RStudio installieren 45
4.4 Benutzeroberfläche 49
4.5 Erweiterte Funktionen 53

4.6 Das Paket machinelearning 55
Literatur. 56

5 Grundlagen der Programmiersprache R 57
5.1 Datentypen 59
5.2 Datenstrukturen 63
5.2.1 Vektor. 63
5.2.2 Faktorvariable 66
5.2.3 Liste 70
5.2.4 Matrix und Array 71
5.2.5 Dataframe und Tibble 75
5.2.6 Übersichtstabelle 82
5.3 Operationen 82
5.3.1 Arithmetische Operationen 82
5.3.2 Logische Operationen 85
5.3.3 Selektive Operationen 87
5.3.4 Textoperationen 89
5.3.5 Datumsoperationen 91
5.4 Funktionen 97
5.4.1 Funktionen aufrufen. 97
5.4.2 Funktionen erstellen. 100
5.5 Kontrollstrukturen 105

6 Grundlagen des Machine Learnings mit R 111
6.1 Import und Export von Daten 111
6.1.1 Import von Daten 113
6.1.2 Export von Daten 114
6.2 Vorverarbeitung 118
6.2.1 Selektion 118
6.2.2 Validierung. 127
6.2.3 Fusion. 132
6.2.4 Umformung 139
6.2.5 Ergänzung 145
6.2.6 Zusammenfassung 151
6.2.7 Best Practices Fallstudie 157
6.3 Explorative Datenanalyse 166
6.3.1 Analysemethoden für Basistabellen. 168
6.3.2 Statistische Analysemethoden 179
6.3.3 Datenvisualisierung 208
6.4 Modellierung und Evaluation. 256
6.4.1 Modellierung 258
6.4.2 Stichproben 282
6.4.3 Trainings- und Testdaten 291
6.4.4 Evaluation 293

6.5 Das Paket Tidymodels 300
6.5.1 Modellwahl, Modelltraining und Modellanwendung 300
6.5.2 Resampling 318
6.5.3 Vorverarbeitung 334
6.5.4 Vergleich von Modellen 360
6.5.5 Parametertuning 373
6.5.6 Fazit 397
Literatur 397

7 Anwendung von Machine Learning mit R 399
7.1 Überwachtes Lernen 399
7.1.1 Regression 399
7.1.2 Klassifikation 446
7.2 Unüberwachtes Lernen 457
7.2.1 Clustering 457
7.2.2 Assoziationsregeln 473
Literatur 490

8 Ausblick 491
Literatur 494

Literatur 495