

Inhalt

Vorworte	11
Einleitung	15
1 Einführung	19
<hr/>	
1.1 Themenabgrenzung	20
1.1.1 Business Intelligence	21
1.1.2 Statistik	21
1.1.3 Künstliche Intelligenz	21
1.1.4 Data Science	24
1.1.5 CRISP-DM	25
1.2 Data Science in der SAP-Welt	27
1.2.1 Standardsoftware	28
1.2.2 Wiederverwendbare Dienste	29
1.2.3 SAP Business Technology Platform	31
1.2.4 AI-Ethik	35
2 SAP HANA als Data-Science-Umgebung	37
<hr/>	
2.1 SAP HANA	38
2.1.1 SAP HANA, Express Edition	40
2.1.2 SAP HANA Cloud	53
2.1.3 SAP Data Warehouse Cloud	60
2.2 SAP HANA Client	64
2.3 Python-Entwicklungsumgebung	64
3 Erste Schritte	71
<hr/>	
3.1 Python	71
3.1.1 Erste Schritte mit Python	72
3.1.2 Verbindung zu SAP HANA	76

- 3.1.3 Daten laden 81
- 3.1.4 Datenverarbeitung 82
- 3.2 Die R-Programmiersprache 84**
- 3.3 Die SQL-Sprache für SAP HANA 89**
 - 3.3.1 SAP HANA Database Explorer 89
 - 3.3.2 DBeaver 95

4 Explorative Datenanalyse und Datenvorbereitung 99

- 4.1 Analyse einer Tabelle 100**
- 4.2 Analyse einzelner Variablen 105**
 - 4.2.1 Numerische Variablen 105
 - 4.2.2 Kategorische Variablen 109
- 4.3 Analyse mehrerer Variablen 111**
- 4.4 Datenvorbereitung 114**
 - 4.4.1 Datenbeschreibung 115
 - 4.4.2 Daten laden 115
 - 4.4.3 Datenmodifikation 117

5 Automated Predictive Library 129

- 5.1 Einführung in die APL 131**
- 5.2 Klassifizierung mit der APL 133**
 - 5.2.1 Binäre Klassifizierung mit der APL 133
 - 5.2.2 Multiclass-Klassifizierung mit der APL 158
- 5.3 Regression mit der APL 173**
- 5.4 Zeitreihen mit der APL 186**
 - 5.4.1 Einzelne Zeitreihe 186
 - 5.4.2 Einzelne Zeitreihe mit Prädiktoren 196
 - 5.4.3 Mehrere Zeitreihen 204
- 5.5 Weitere Informationen 211**

6	Predictive Analysis Library	213
6.1	Einführung in die PAL	214
6.2	Klassifizierung mit der PAL	216
6.2.1	Binäre Klassifizierung mit der PAL	216
6.2.2	Multiclass-Klassifizierung mit der PAL	222
6.3	Regression mit der PAL	229
6.4	Zeitreihen mit der PAL	236
6.4.1	Zeitreihenzerlegung	237
6.4.2	Vorhersagemethoden in der Zeitreihenanalyse	238
6.4.3	Dynamische Zeitverzerrung	247
6.4.4	Spektralanalyse	251
6.5	Cluster-Analyse	253
6.6	Survival Analysis	260
6.7	Ausreißeranalyse	282
6.7.1	Statistische Methoden	284
6.7.2	Clustering-basierte Methoden	287
6.7.3	Regressions- und Klassifizierungsmethoden	288
6.8	Automated Machine Learning	289
6.9	State-enabled Deployment	299
7	Spezialisierte Analyse-Engines	301
7.1	Geodatenanalyse	302
7.1.1	Einführung eines praktischen Beispiels	303
7.1.2	Installieren von zusätzlichen Packages	304
7.1.3	Datenbeschreibung und Daten laden	304
7.1.4	Datenanalyse	308
7.2	Graphanalyse	317
7.2.1	Einführung eines praktischen Beispiels	318
7.2.2	Datenbeschreibung	319
7.2.3	Daten laden	320
7.2.4	Datenanalyse	322
7.2.5	Graphanalyse durchführen	323

7.3	Textanalyse	328
7.3.1	Beispiel und Datenbeschreibung	329
7.3.2	Datenvorverarbeitung und -analyse	330
7.3.3	Weiterführende Analysen	336

8 Deployment-Optionen 341

8.1	SAP Data Intelligence	342
8.1.1	JupyterLab Notebooks	344
8.1.2	Template Pipelines	346
8.1.3	Python-3-Operator	349
8.2	Cloud Foundry	357
8.3	Kyma	363
8.3.1	Scheduling von Code	366
8.3.2	Triggern der Ausführung über einen API-Endpunkt	369
8.3.3	Vorhersagen über REST APIs	372

9 Tipps und Tricks 377

Anhang 397

A	Checkliste	397
----------	-------------------------	-----

Die Autoren	401
Index	403

Diese Leseprobe haben Sie beim
 edv-buchversand.de heruntergeladen.
Das Buch können Sie online in unserem
Shop bestellen.
[Hier zum Shop](#)