

Inhaltsverzeichnis

Vorwort	9
1 Einleitung	15
1.1 Machine Learning – ein Überblick	15
1.1.1 Anwendungsszenarien	15
1.1.2 Funktionen von SAP HANA	17
1.2 Zentrale Begriffe und Methoden	20
1.2.1 Daten für Machine Learning	21
1.2.2 Beaufsichtigtes Lernen	24
1.2.3 Unbeaufsichtigtes Lernen	29
1.3 Beispieldaten mit Eclipse importieren	29
1.4 Python-Tools verwenden	38
2 Predictive Analysis Library mit SQLScript	47
2.1 Entscheidungsbäume für Klassifikationsprobleme	49
2.1.1 Trainieren eines Entscheidungsbaums	51
2.1.2 Testen des trainierten Modells	63
2.2 Clusteranalysen mit K-Means	70
3 Automated Predictive Library mit SQLScript	85
3.1 Ad-hoc-Analysen mit Predictive Business Functions	86
3.1.1 Statistiken und Datenprofil berechnen	87
3.1.2 Einflussfaktoren ermitteln	94
3.2 Predictive Model Functions	102
3.2.1 Modelle trainieren	102
3.2.2 Modelle testen und anwenden	123
3.2.3 Modelle bewerten	131
3.2.4 Weiterführende Informationen	140
4 Der Python Machine Learning Client für SAP HANA	143
4.1 Datenexploration und -vorbereitung mit HANA-Techniken	143

4.2	Datenvisualisierung im Jupyter Notebook	158
4.2.1	<i>Visualisierungsfunktionen des Python ML Client</i>	158
4.2.2	<i>Verwendung von Pandas und Plotly</i>	172
4.3	PAL in Python aufrufen	181
4.3.1	<i>Klassifikationsmodelle</i>	181
4.3.2	<i>Clusteranalysen</i>	207
4.3.3	<i>Modelle speichern und wiederverwenden</i>	219
4.4	APL in Python aufrufen	226
5	Text Mining	231
5.1	<i>Import der Beispieldaten</i>	233
5.2	<i>Verwendung des Volltextindex</i>	237
5.2.1	<i>Erstellen des Volltextindex</i>	237
5.2.2	<i>Dokumentensuche und Schlagwörter</i>	243
5.3	<i>Themenextraktion</i>	246
5.3.1	<i>Normalisieren der Texte</i>	248
5.3.2	<i>Der Algorithmus Latent Dirichlet Allocation</i>	251
5.3.3	<i>Visualisierung der Themenscores in Python</i>	259
5.4	<i>Textklassifikation mit dem Python ML Client</i>	267
6	Integration in SAP-Anwendungen	275
6.1	<i>Vergleich der SAP-Produkte und -Techniken</i>	275
6.1.1	<i>Native SAP-HANA-Entwicklung</i>	276
6.1.2	<i>SAP S/4HANA</i>	276
6.1.3	<i>SAP BW/4HANA oder BW on HANA</i>	277
6.1.4	<i>SAP Predictive Analytics</i>	278
6.1.5	<i>SAP Data Intelligence</i>	278
6.1.6	<i>SAP Analytics Cloud</i>	279
6.1.7	<i>SAP AI Business Services</i>	280
6.1.8	<i>Python, SQLScript oder ABAP</i>	280
6.2	<i>PAL-Prognosemodelle in Calculation Views</i>	281
7	Aufbau einer Testumgebung	289
7.1	<i>Einrichten der HANA-Instanz</i>	289
7.2	<i>Einrichten der Python-Umgebung</i>	293

8 Fazit und Ausblick	297
A Der Autor	304
B Quellcodeverzeichnis	305
C Literaturverzeichnis	307
D Index	311
E Disclaimer	316