

Inhaltsverzeichnis

Über die Autoren	7
Widmung	8
Danksagung	9
Einführung	21
Über dieses Buch	21
Einige törichte Annahmen	22
In diesem Buch verwendete Symbole	22
Über dieses Buch hinaus	23
Wie es weitergeht	23
Teil I	
Erste Bekanntschaft mit Predictive Analytics	25
Kapitel 1	
Die Arena betreten	27
Predictive Analytics erklären	27
Data-Mining	27
Das Modell in den Mittelpunkt stellen	28
Den Geschäftswert erhöhen	30
Unendliche Möglichkeiten	30
Ihr Unternehmen stärken	31
Ein Predictive Analytics-Projekt starten	32
Unternehmenskenntnisse	33
Data-Science-Team und -Technologie	34
Die Daten	35
Den Markt begutachten	35
Auf Big Data reagieren	36
Mit Big Data arbeiten	36
Kapitel 2	
Predictive Analytics in der freien Wildbahn	39
Online-Marketing und Verkauf	41
Empfehlungsdienste	41
Realisierung eines Empfehlungsdienstes	43
Kollaboratives Filtern	43
Inhaltbasiertes Filtern	51
Hybrid-Empfehlungsdienste	55
Zielgruppenansprache mithilfe der Modellierung von Verhaltensweisen	57
Uplift-Modell	58

Zielgruppengerichtetes Marketing	56
Predictive Analytics gegen Betrug und Verbrechen	60
Inhalte und Texte analysieren	61

Kapitel 3

Datentypen und die damit verbundenen Techniken **63**

Ihre Datentypen erkennen	63
Strukturierte und unstrukturierte Daten	64
Statische und fließende Daten	66
Datenkategorien erkennen	67
Einstellungsdaten	68
Verhaltensdaten	69
Demografische Daten	70
Predictive Analytics einsetzen	70
Die datengesteuerte Analyse	71
Nutzergesteuerte Analyse	72
Die Verbindung zu benachbarten Disziplinen	73
Statistik	73
Data-Mining	74
Maschinelles Lernen	74

Kapitel 4

Die Komplexität von Daten **77**

Nutzen aus Ihren Daten ziehen	77
In Ihren Daten graben	78
Validität der Daten	78
Vielfalt der Daten	79
Ständige Änderung der Daten	79
Geschwindigkeit der Daten	80
Volumen der Daten	80
Schwierigkeiten beim Durchsuchen Ihrer Daten	81
Die Suche nach Schlüsselwörtern	81
Semantische Suche	81
Die Unterschiede zwischen Business Intelligence und der Analyse von Big Data	83
Die Visualisierung von Rohdaten	84
Die Attribute der Daten bestimmen	84
Die Visualisierung von Daten erkunden	85
Tabellarische Visualisierungen	85
Balkendiagramme	86
Tortendiagramme	86

Graphen	88
Schlagwortwolken	89
Kurvendiagramm	90
Darstellung durch Vogelschwärme	90

Teil II

Algorithmen bei der Modellbildung berücksichtigen **93**

Kapitel 5

Modelle verwenden **95**

Daten modellieren	95
Modelle und Simulationen	96
Modelle kategorisieren	98
Daten beschreiben und zusammenfassen	100
Bessere Geschäftsentscheidungen treffen	100
Fallstudien aus dem Gesundheitswesen	101
Suchanfragen bei Google als Anzeichen für eine Ansteckung	101
Die Chance, einen Krebs zu überleben	102
Fallstudien zum Marketing in sozialen Medien	103
Tweets als Anzeichen für den Aktienhandel	103
Eine Kaufhauskette sagt Schwangerschaften von Frauen vorher	104
Auf Twitter beruhende Anzeichen für Erdbeben	105
Mit Twitter die Ergebnisse politischer Kampagnen vorsagen	106

Kapitel 6

Ähnlichkeiten bei Daten erkennen **109**

Die Clusteranalyse erklären	109
Motivation	111
Rohdaten in eine Matrix verwandeln	113
Eine Matrix aus Ausdrücken erstellen	113
Auswahl der Terme	114
K-Gruppen in Ihren Daten bestimmen	115
k-Means-Algorithmus	115
Clustering mit den nächsten Nachbarn	119
Korrelationen zwischen Datenelementen entdecken	121
Biologisch inspirierte Clustermethoden verwenden	123
Vogelschwärme	123
Ameisenkolonien	128

Kapitel 7

Mithilfe der Klassifizierung von Daten die Zukunft vorhersagen **131**

Die Klassifizierung von Daten	132
Kreditvergabe	133
Marketing	133
Gesundheit	134
Was kommt als Nächstes?	135
Die Klassifizierung von Daten in Ihrem Unternehmen einführen	135
Den Prozess der Klassifizierung von Daten erklären	137
Mithilfe der Klassifizierung die Zukunft vorhersagen	139
Entscheidungsbäume	139
Support Vector Machine	141
Einfache Bayes-Klassifizierungsalgorithmen	143
Neuronale Netze	148
Das Markov-Modell	150
Lineare Regression	155
Die Genauigkeit durch die Verwendung von Ensemble-Methoden verbessern	156

Teil III

Einen Fahrplan entwickeln **159**

Kapitel 8

Ihr Management von der Einführung von Predictive Analytics überzeugen **161**

Für das Geschäftsmodell werben	162
Vorteile für das Unternehmen	163
Unterstützung von Projektbeteiligten bekommen	169
Mit Ihren Auftraggebern arbeiten	170
Fachkräfte beauftragen	172
Auf Informationstechnologie setzen	173
Schnelle Entwicklung eines Prototypen	177
Ihren Plan präsentieren	178

Kapitel 9

Vorbereitung der Daten **181**

Die Geschäftsziele auflisten	182
Verwandte Ziele erkennen	182
Die Voraussetzungen festlegen	183
Ihre Daten verarbeiten	183
Ihre Daten bestimmen	184
Die Daten reinigen	185
Abgeleitete Daten erzeugen	186
Die Dimensionalität der Daten reduzieren	187

Ihre Daten strukturieren	188
Ihre Daten extrahieren, umwandeln und laden	188
Die Daten auf dem Laufenden halten	189
Das Testen und die Testdaten erläutern	189
Kapitel 10	
<i>Ein Vorhersagemodell erstellen</i>	191
Der Einstieg	191
Ihre Geschäftsziele bestimmen	193
Vorbereitung Ihrer Daten	194
Einen Algorithmus wählen	196
Das Modell entwickeln und testen	197
Entwicklung des Modells	197
Test des Modells	198
Das Modell beurteilen	200
Den Regelbetrieb aufnehmen	201
Einsatz des Modells	201
Das Modell überwachen und erhalten	202
Kapitel 11	
<i>Visualisierung analytischer Ergebnisse</i>	203
Visualisierungen als vorhersagendes Instrument	203
Warum die Visualisierung eine Rolle spielt	204
Die Vorteile der Visualisierung nutzen	205
Mit der Vielschichtigkeit arbeiten	206
Ihre Visualisierung bewerten	207
Wie relevant ist diese Abbildung?	207
Wie aussagekräftig ist diese Abbildung?	207
Ist die Abbildung einfach genug?	208
Führt die Abbildung zu neuen Erkenntnissen?	208
Die analytischen Ergebnisse Ihres Modells visualisieren	208
Die Visualisierung verborgener Gruppierungen	208
Die Visualisierung der Ergebnisse einer Datenklassifizierung	210
Die Visualisierung von Ausreißern in Ihren Daten	211
Die Visualisierung von Entscheidungsbäumen	211
Vorhersagen visualisieren	213
Weitere Arten der Visualisierung in Predictive Analytics	215
Das Schwarmverhalten visualisieren	215

Teil IV	
Predictive Analytics programmieren	221
Kapitel 12	
Grundlegende vorhersagende Beispiele erzeugen	223
Das Softwarepaket installieren	223
Python installieren	224
Das Modul zum maschinellen Lernen installieren	224
Die Zusatzdateien installieren	228
Aufbereitung der Daten	232
Den Probedatensatz bekommen	233
Bezeichnung Ihrer Daten	233
Mithilfe von Klassifizierungsalgorithmen Vorhersagen machen	234
Überwachtes Lernen mithilfe von Stützvektoren	235
Ein Modell des überwachten Lernens mit der logistischen Regression erzeugen	244
Zwei Klassifizierungsmodelle vergleichen	248
Kapitel 13	
Grundlegende Beispiele für unüberwachte Vorhersagen	251
Den Probedatensatz herunterladen	251
Die Verwendung von Clusteralgorithmen zur Vorhersage	252
Vergleich zweier Clustermodelle	252
Ein Modell des unüberwachten Lernens mit k-Means erzeugen	253
Mit ein Modell des unüberwachten Lernens erzeugen	263
Kapitel 14	
Die vorhersagende Modellierung in R	269
In R programmieren	270
R installieren	271
RStudio installieren	271
Sich mit der Umgebung vertraut machen	271
Ein wenig R lernen	273
Vorhersagen mithilfe von R machen	277
Mithilfe der Regression vorhersagen	278
Die Klassifikation zur Vorhersage verwenden	287

Kapitel 15
Fallen bei der Analyse vermeiden 295

Die Herausforderungen	296
Die Grenzen der Daten skizzieren	297
Extremfälle (Ausreißer) behandeln	299
Die Daten glätten	302
Kurven anpassen	306
Die Annahmen auf ein Minimum beschränken	308
Analytische Herausforderungen	309
Die überwachte Analyse	310
Auf nur eine Analyse vertrauen	310
Die Grenzen des Modells beschreiben	311
Nichtskalierbare Modelle vermeiden	312
Ihre Vorhersagen genau bewerten	313

Kapitel 16
Auf Big Data zielen 315

Wesentliche technologische Trends bei Predictive Analytics	316
Der Einsatz von Predictive Analytics als Service	316
Verteilte Daten für die Analyse verbinden	317
Datengesteuerte Analyse in Echtzeit	318
Frei zugängliche Tools auf Big Data anwenden	319
Apache Hadoop	319
Apache Mahout	322
Einen schnellen Prototyp Ihres Vorhersagemodells erstellen	323
Musterentwicklung für Predictive Analytics	323
Ihr Vorhersagemodell testen	326

Teil V
Der Top-Ten-Teil 329
Kapitel 17
Zehn Gründe für die Durchführung von Predictive Analytics 331

Die Geschäftsziele erläutern	331
Über Ihre Daten Bescheid wissen	332
Ihre Daten organisieren	333
Ihre Kunden zufriedenstellen	334
Die Betriebskosten reduzieren	336
Die Kapitalrendite erhöhen	336
Wachsendes Vertrauen	337
Kundige Entscheidungen treffen	338
Wettbewerbsvorteile gewinnen	339
Das Geschäft verbessern	340

Kapitel 18

Zehn Schritte zum Erstellen eines Vorhersagemodells	341
Die Bildung eines Teams für die Predictive Analytics	342
Geschäftskompetenz mit ins Boot holen	342
Mit IT- und Mathematikkenntnissen einheizen	343
Die Geschäftsziele festlegen	343
Ihre Daten aufbereiten	344
Die Auswahl Ihrer Daten	345
»Wo man Müll hineinsteckt, kommt auch Müll heraus« – vermeiden Sie das	345
Einfach ist nicht dumm	346
Die Aufbereitung der Daten fördert das gute Material zutage	346
Schnelle Erfolge erzielen	346
Änderungen in Ihrem Unternehmen unterstützen	347
Einsetzbare Modelle erstellen	349
Ihr Modell überprüfen	350
Ihr Modell aktualisieren	350
 Stichwortverzeichnis	 353