

Maschinelles Lernen und Datenanalyse erfreuen sich derzeit großer Beliebtheit und entwickeln sich rasant weiter. Ich habe während des größten Teils meines Berufslebens mit Python und Daten gearbeitet und mir dabei eine Referenz in Form eines gedruckten Buchs gewünscht: eine Beschreibung der gängigen Verfahren, die ich bei der Arbeit eingesetzt und in Workshops gelehrt habe, um Probleme aus dem Bereich des maschinellen Lernens mit strukturierten Daten zu lösen.

Dieses Buch enthält die aus meiner Sicht beste Sammlung von Ressourcen und Beispielen, um eine Modellierungsaufgabe für Vorhersagen anhand strukturierter Daten anzugehen. Es gibt viele Bibliotheken, die jeweils einen Teil der erforderlichen Schritte durchführen, und ich habe versucht, diejenigen aufzunehmen, die ich bei der praktischen Anwendung dieser Techniken nützlich gefunden habe.

Viele Leser mögen bedauern, dass Deep-Learning-Techniken fehlen. Diese könnten allerdings ein eigenes Buch füllen. Außerdem bevorzuge ich einfachere Techniken – wie offenbar auch andere professionelle Nutzer: Deep Learning für unstrukturierte Daten (Video- und Tonmaterial, Bilder) und mächtige Werkzeuge wie XGBoost für strukturierte Daten.

Ich hoffe, dass Ihnen das Buch als nützliche Referenz bei der Lösung drängender Probleme dienen kann.

Was Sie erwartet

Dieses Buch zeigt anhand detaillierter Beispiele, wie man übliche Aufgaben löst, bei denen es um strukturierte Daten geht. Es führt durch verschiedene Bibliotheken und Modelle, ihre jeweiligen Vor- und Nachteile, Anpassungsmöglichkeiten und Interpretationen.

Die Codeausschnitte sind so gestaltet, dass Sie sie für Ihre eigenen Projekte nutzen und anpassen können.

Für wen dieses Buch gedacht ist

Ob Sie gerade erst damit beginnen, sich mit Machine Learning zu beschäftigen, oder schon seit Jahren damit arbeiten – dieses Buch kann als wertvolle Referenz dienen. Es setzt einige Kenntnis in Python voraus und geht überhaupt nicht auf Fragen der Syntax ein. Vielmehr führt es vor, wie man mithilfe verschiedener Bibliotheken wirkliche Probleme löst.

Dieses Buch ist kein Ersatz für einen tiefergehenden Kurs, sollte aber eine gute Referenz für die Themen darstellen, die ein Machine-Learning-Kurs abdecken kann. (Der Autor benutzt es übrigens als Referenz für seine eigenen Kurse in Datenanalyse und maschinellem Lernen.)

Konventionen in diesem Buch

Folgende typografische Konventionen werden in diesem Buch angewandt:

Kursiv

Zeigt neue Begriffe, URLs, E-Mail-Adressen, Dateinamen und -erweiterungen an.

Konstante Zeichenbreite

Wird für Programmlistings verwendet, aber auch im Fließtext, um Programmelemente wie Variablen oder Funktionsnamen, Datenbanken, Datentypen, Umgebungsvariablen, Befehle und Schlüsselwörter auszuzeichnen.



Dieses Element zeigt einen Hinweis oder Vorschlag an.



Dieses Element zeigt eine allgemeine Bemerkung an.



Dieses Element zeigt einen Warnhinweis an.

Verwendung von Codebeispielen

Weiterführende Materialien (Codebeispiele, Übungen usw.) sind verfügbar unter https://github.com/mattharrison/ml_pocket_reference.

Dieses Buch soll Ihnen helfen, Ihre Arbeit zu erledigen. Im Allgemeinen dürfen Sie Codebeispiele, die zu diesem Buch gehören, in Ihren Programmen und Dokumentationen verwenden. Sie müssen uns nicht um Erlaubnis fragen, es sei denn, Sie wollen einen nennenswerten Anteil des Codes reproduzieren. Um ein Programm zu schreiben, das ein paar Codeabschnitte aus diesem Buch einsetzt, benötigen Sie beispielsweise keine Erlaubnis. Für den Verkauf oder die Verteilung einer CD-ROM mit Beispielen aus O'Reilly-Büchern müssen Sie jedoch eine Genehmigung einholen. Beantworten Sie eine Frage, indem Sie auf dieses Buch verweisen und Beispielcode zitieren, brauchen Sie dafür keine Erlaubnis. Jedoch benötigen Sie eine Genehmigung, um eine beträchtliche Menge Beispielcode aus diesem Buch in die Dokumentation Ihres Produkts zu übernehmen.

Wir freuen uns darüber, zitiert zu werden, verlangen es aber nicht. Ein Zitat umfasst gewöhnlich Titel, Autor, Verlag und ISBN. Ein Beispiel: »*Machine-Learning – Die Referenz* von Matt Harrison (O'Reilly). Copyright 2021 dpunkt.verlag, ISBN 978-3-96009-135-6.«

Wenn Sie vermuten, dass Ihr Einsatz von Codebeispielen nicht mehr unter das Prinzip der fairen Nutzung oder die oben gegebene Erlaubnis fällt, schreiben Sie uns gern unter folgender Adresse an: komentar@oreilly.de.

Dank

Vielen Dank an meine Frau und meine Familie für ihre Unterstützung. Ich bin der Python-Community dankbar für eine wunderbare Programmiersprache und die Auswahl an Werkzeugen. Die Zusammenarbeit mit Nicole Tache war sehr angenehm und lieferte ausgezeichnetes Feedback. Meinen technischen Gutachtern Mikio Braun, Natalino Busa und Justin Francis verdanke ich es, ehrlich geblieben zu sein. Danke!