

Vorwort

Für wen dieses Buch gedacht ist

Einführende Bücher zum Machine Learning konzentrieren sich normalerweise auf das *Was* und *Wie* des maschinellen Lernens (ML). Sie erläutern dann die mathematischen Aspekte neuer Methoden aus KI-Forschungseinrichtungen und lehren, wie sich diese Methoden mithilfe von KI-Frameworks implementieren lassen. Dieses Buch hingegen bringt die hart erarbeiteten Erfahrungen rund um das *Warum* zusammen, das den Tipps und Tricks zugrunde liegt, auf die erfahrene ML-Praktiker:innen setzen, wenn sie maschinelles Lernen auf reale Probleme anwenden.

Wir gehen davon aus, dass Sie über Vorkenntnisse zu maschinellem Lernen und Datenverarbeitung verfügen. Dies ist kein Grundlagenlehrbuch für maschinelles Lernen. Vielmehr richtet sich dieses Buch an Data Scientists oder ML Engineers, die nach einem zweiten Buch über praktisches Machine Learning suchen. Wenn Sie die Grundlagen bereits kennen, präsentiert Ihnen dieses Buch einen Katalog von Ideen, von denen Sie (als ML-Praktiker:in) vielleicht einige wiedererkennen, und gibt diesen Ideen einen Namen, damit Sie zielsicher nach ihnen greifen können.

Wenn Sie Informatikstudent:in sind und einen Job in der Industrie anstreben, wird dieses Buch Ihr Wissen abrunden und Sie auf die Berufswelt vorbereiten. Es wird Ihnen helfen, zu lernen, wie man hochwertige ML-Systeme aufbaut.

Was Sie nicht im Buch finden

In erster Linie richtet sich das Buch an ML Engineers in Unternehmen, nicht an ML-Wissenschaftler:innen in akademischen oder industriellen Forschungslabors.

Wir gehen absichtlich nicht auf Bereiche der aktiven Forschung ein – so werden Sie hier sehr wenig zur Modellarchitektur des maschinellen Lernens finden (wie zum Beispiel bidirektionale Encoder, den Aufmerksamkeitsmechanismus oder Kurzschlusschichten), da wir annehmen, dass Sie mit einer vorgefertigten Modellarchitektur (wie etwa ResNet-50 oder GRUCell) arbeiten und nicht Ihr eigenes Bildklassifizierungs- oder rekurrentes neuronales Netz schreiben werden.

Die folgenden Punkte listen einige konkrete Beispiele für Bereiche auf, von denen wir uns absichtlich fernhalten, da wir glauben, dass diese Themen eher für Hochschulkurse und ML-Forscher geeignet sind:

ML-Algorithmen

Zum Beispiel behandeln wir nicht die Unterschiede zwischen Random Forests und neuronalen Netzen. Damit beschäftigen sich Lehrbücher, die in das maschinelle Lernen einführen.

Bausteine

Verschiedene Arten von Optimierern für den Gradientenabstieg oder Aktivierungsfunktionen behandeln wir ebenfalls nicht. Wir empfehlen hier Adam und ReLU – unserer Erfahrung nach ist das Potenzial für Verbesserungen der Performance durch verschiedene Auswahlen bei derartigen Dingen eher gering.

ML-Modellarchitekturen

Wenn Sie Bilder klassifizieren, verwenden Sie am besten ein Standardmodell wie ResNet oder was auch immer der letzte Schrei ist, wenn Sie dies lesen. Überlassen Sie den Entwurf neuer Modelle für die Bild- oder Textklassifizierung den Forschern, die sich auf derartige Probleme spezialisiert haben.

Modellebenen

In diesem Buch werden Sie keine Convolutional Neural Networks und keine rekurrenten neuronalen Netze finden, und zwar gleich aus zwei Gründen – erstens, weil sie Bausteine sind, und zweitens, weil sie zu den Dingen gehören, die bereits als Standardlösungen zur Verfügung stehen.

Benutzerdefinierte Trainingsschleifen

Der einfache Aufruf von `model.fit()` in Keras dürfte den Ansprüchen von Praktikerinnen und Praktikern genügen.

Wir haben versucht, in dieses Buch nur gängige Muster aufzunehmen – wie sie etwa ML Engineers in Unternehmen bei ihrer täglichen Arbeit verwenden.

Nehmen Sie als Analogie dazu Datenstrukturen. Während ein Hochschulkurs über Datenstrukturen eingehend die Implementierungen der verschiedenen Datenstrukturen erläutert und ein Forscher zu Datenstrukturen lernen muss, wie man ihre mathematischen Eigenschaften formal darstellt, können Fachleute pragmatischer herangehen. Entwickler:innen von Unternehmenssoftware müssen einfach wissen, wie sie effektiv mit Arrays, verketteten Listen, Mengen und Bäumen arbeiten können. Dieses Buch ist für pragmatische Fachleute des maschinellen Lernens geschrieben.

Codebeispiele

Wir stellen Code für maschinelles Lernen (manchmal in Keras/TensorFlow, manchmal in scikit-learn oder BigQuery ML) und Datenverarbeitung (in SQL) zur Verfügung, um zu zeigen, wie die im Buch beschriebenen Techniken in der Praxis umgesetzt werden. Der gesamte Code, auf den im Buch verwiesen wird, ist Teil unseres

GitHub-Repositorys (<https://github.com/GoogleCloudPlatform/ml-design-patterns>), in dem Sie voll funktionsfähige ML-Modelle finden. Diese Codebeispiele sollten Sie unbedingt ausprobieren.

Gegenüber den behandelten Konzepten und Techniken spielt der Code nur eine untergeordnete Rolle. Unser Ziel ist es gewesen, dass Thema und Prinzipien unabhängig von Änderungen an TensorFlow oder Keras relevant bleiben. Wir können uns durchaus vorstellen, das GitHub-Repository zum Beispiel mit anderen ML-Frameworks zu ergänzen, während der Buchtext unverändert bleibt. Daher sollte das Buch für Sie genauso informativ sein, wenn Ihr primäres ML-Framework PyTorch oder sogar ein Nicht-Python-Framework wie H2O.ai oder R ist. Darüber hinaus begrüßen wir es, wenn Sie die Implementierung eines Musters (es dürfen auch mehrere sein) in Ihrem bevorzugten ML-Framework als Beitrag für das GitHub-Repository bereitstellen.

Dieses Buch soll Ihnen bei Ihrer täglichen Arbeit helfen. Falls Beispielcode zum Buch angeboten wird, dürfen Sie ihn im Allgemeinen in Ihren Programmen und für Dokumentationen verwenden. Sie müssen uns nicht um Erlaubnis bitten, es sei denn, Sie kopieren einen erheblichen Teil des Codes. Wenn Sie zum Beispiel ein Programm schreiben, das einige Codeblöcke aus diesem Buch verwendet, benötigen Sie keine Erlaubnis. Sollten Sie aber Beispiele aus O'Reilly-Büchern verkaufen oder verbreiten, ist eine Erlaubnis erforderlich. Wenn Sie eine Frage beantworten und dabei dieses Buch oder Beispielcode aus diesem Buch zitieren, brauchen Sie wiederum keine Erlaubnis. Aber wenn Sie große Teile des Beispielcodes aus diesem Buch in die Dokumentation Ihres Produkts einfließen lassen, ist eine Erlaubnis einzuholen.

Wir schätzen eine Quellenangabe, verlangen sie aber nicht. Eine Quellenangabe umfasst in der Regel Titel, Autor, Verlag und ISBN. Zum Beispiel: »*Design Patterns für Machine Learning* von Valliappa Lakshmanan, Sara Robinson und Michael Munn (O'Reilly). Copyright 2022 dpunkt.verlag, ISBN 978-3-96009-164-6.« Wenn Sie der Meinung sind, dass Sie die Codebeispiele in einer Weise verwenden, die über die oben erteilte Erlaubnis hinausgeht, kontaktieren Sie uns bitte unter komentar@oreilly.de.

Typografischen Konventionen

In diesem Buch folgen wir diesen typografischen Konventionen:

Kursiv

Kennzeichnet neue Begriffe, URLs, E-Mail-Adressen, Dateinamen und Dateierweiterungen.

Schreibmaschinenschrift

Wird in Programmlistings verwendet und im Fließtext für Programmelemente wie zum Beispiel Variablen- oder Funktionsnamen, Datenbanken, Datentypen, Umgebungsvariablen, Anweisungen und Schlüsselwörter.

Schreibmaschinenschrift **fett**

Kennzeichnet Befehle oder andere Texte, die vom Benutzer buchstäblich eingegeben werden sollen.

Schreibmaschinenschrift *kursiv*

Zeigt Text, der ersetzt werden soll, durch Werte, die der Benutzer bereitstellt, oder Werte, die sich aus dem Kontext ergeben.



Dieses Element kennzeichnet einen Tipp oder Vorschlag.



Dieses Element kennzeichnet einen allgemeinen Hinweis.



Dieses Element kennzeichnet eine Warnung oder einen Achtungshinweis.

Danksagungen

Ein Buch wie dieses wäre nicht möglich ohne die Großzügigkeit zahlreicher Googler, insbesondere unserer Kolleg:innen aus den Teams Cloud AI, Solution Engineering, Professional Services und Developer Relations. Wir sind ihnen dankbar, dass wir ihre Lösungen für die herausfordernden Probleme beim Training, bei der Verbesserung und der Operationalisierung von ML-Modellen beobachten, analysieren und hinterfragen durften. Wir danken unseren Managern Karl Weinmeister, Steve Cellini, Hamidou Dia, Abdul Razack, Chris Hallenbeck, Patrick Cole, Louise Byrne und Rochana Golani dafür, dass sie den Geist der Offenheit bei Google fördern und uns die Freiheit geben, diese Muster zu katalogisieren und dieses Buch zu veröffentlichen.

Salem Haykal, Benoit Dherin und Khalid Salama haben jedes Muster und jedes Kapitel durchgesehen. Sal hat uns auf Feinheiten hingewiesen, die wir übersehen hatten, Benoit hat unsere Behauptungen eingegrenzt, und Khalid hat uns auf relevante Forschungsarbeiten aufmerksam gemacht. Ohne ihre Beiträge wäre dieses Buch bei Weitem nicht so gut geworden. Vielen Dank dafür! Amy Unruh, Rajesh Thallam, Robbie Haertel, Zhitao Li, Anusha Ramesh, Ming Fang, Parker Barnes, Andrew Zaldivar, James Wexler, Andrew Sellergren und David Kanter haben die Teile dieses Buchs, die in ihre Fachgebiete fallen, überprüft und zahlreiche Vorschläge dazu unterbreitet, wie die kurzfristige Roadmap unsere Empfehlungen be-

einflussen würde. Nitin Aggarwal und Matthew Yeager haben das Auge des Lesers in das Manuskript eingebracht und dessen Klarheit verbessert. Besonderer Dank gebührt Rajesh Thallam, der den Prototyp für das Design der allerletzten Abbildung in Kapitel 8 erstellt hat. Alle verbliebenen Fehler gehen natürlich auf unsere Kappe.

O'Reilly ist der Verlag der Wahl für technische Bücher, und die Professionalität unseres Teams zeigt, warum das so ist. Rebecca Novak hat uns an die Hand genommen, um eine überzeugende Gliederung zusammenzustellen, Kristen Brown hat die gesamte inhaltliche Entwicklung souverän gemeistert, Corbin Collins hat uns in jeder Phase hilfreiche Tipps gegeben, die Zusammenarbeit mit Elizabeth Kelly während der Produktion war eine Freude, und Charles Roumeliotis hat das Lektorat mit scharfem Blick begleitet. Danke an alle für diese Hilfe!

Michael: Ich danke meinen Eltern, die immer an mich geglaubt und meine Interessen gefördert haben, nicht nur meine akademischen. Sie werden das Cover, auf dem mein Name steht, genauso zu schätzen wissen wie ich. An Phil: Danke, dass du meinen kaum akzeptablen Zeitplan während der Arbeit an diesem Buch geduldig ertragen hast. Und jetzt werde ich schlafen gehen.

Sara: Jon – du bist ein wichtiger Grund dafür, dass es dieses Buch gibt. Danke, dass du mich ermutigt hast, es zu schreiben, dass du immer gewusst hast, wie du mich zum Lachen bringst, dass du meine Verrücktheit zu schätzen weißt und dass du an mich geglaubt hast, vor allem als ich es nicht tat. Meinen Eltern danke ich, dass sie vom ersten Tag, seit ich mich erinnern kann, meine größten Fans waren und meine Liebe zur Technik und zum Schreiben gefördert haben. An Ally, Katie, Randi und Sophie – danke, dass ihr in diesen unsicheren Zeiten eine ständige Quelle des Lichts und des Lachens seid.

Lak: Ich habe dieses Buch in Angriff genommen in der Annahme, ich könnte daran arbeiten, um Wartezeiten auf Flughäfen sinnvoll zu nutzen. COVID-19 hat es »möglich gemacht«, dass ich einen Großteil der Arbeit zu Hause erledigen konnte. Danke Abirami, Sidharth und Sarada für all eure Nachsicht, als ich mich wieder einmal zum Schreiben hingehockt habe. Mehr Wanderungen an den Wochenenden jetzt!

Wir drei spenden 100% der Tantiemen aus diesem Buch an »Girls Who Code« (<https://girlswhocode.com/>), eine Organisation, deren Mission es ist, die weltweit größte Pipeline an zukünftigen Ingenieurinnen aufzubauen. Vielfalt, Gleichberechtigung und Inklusion sind beim maschinellen Lernen besonders wichtig, um sicherzustellen, dass KI-Modelle bestehende Vorurteile in der menschlichen Gesellschaft nicht noch untermauern.