
Vorwort

Dieses Buch zielt darauf ab, Einsteiger in die Verarbeitung natürlicher Sprache (Natural Language Processing, NLP) und Deep Learning an einen Probiertisch zu bringen, auf dem die wichtigsten Themen in beiden Bereichen präsentiert werden. Beide Themenbereiche wachsen exponentiell. Da dieses Buch sowohl in Deep Learning als auch in NLP mit dem Schwerpunkt auf Implementierung einführt, nimmt es eine wichtige Mittelstellung ein. Als wir das Buch geschrieben haben, mussten wir die schwierige und manchmal unbequeme Entscheidung treffen, welches Material wir weglassen sollten. Wir hoffen, dass das Buch insbesondere Anfängern solide Grundkenntnisse vermittelt und einen Einblick in das gibt, was möglich ist. Maschinelles Lernen und speziell Deep Learning sind eher empirische, auf viel Erfahrung beruhende Disziplinen als intellektuelle Wissenschaften. Die großzügigen geschlossenen Beispiele in jedem Kapitel laden Sie ein, an dieser Erfahrung teilzuhaben.

Als wir mit der Arbeit am Buch begannen, ging es los mit PyTorch 0.2. Die Beispiele wurden mit jedem PyTorch-Update von 0.2 bis 0.4 überarbeitet. PyTorch 1.2 ist bereits veröffentlicht, wenn dieses Buch in deutscher Übersetzung erscheint. Die Codebeispiele im Buch sind PyTorch 0.4-konform und funktionieren auch mit PyTorch 1.2.

Ein paar Worte zum Stil des Buchs. An den meisten Stellen haben wir bewusst auf Mathematik verzichtet, nicht weil die Mathematik von Deep Learning besonders schwierig wäre (ist sie nicht), sondern weil sie in vielen Situationen vom Hauptziel dieses Buchs – nämlich den Einsteiger zu befähigen – ablenken würde. In gleicher Weise haben wir sowohl im Code als auch im Text Ausführlichkeit gegenüber Prägnanz bevorzugt. Fortgeschrittene Leser und erfahrene Programmierer sehen wahrscheinlich Möglichkeiten, den Code zu straffen, doch wir wollten so deutlich wie möglich sein, um ein möglichst breites Publikum zu erreichen.

Konventionen in diesem Buch

In diesem Buch werden folgende typografische Konventionen verwendet:

Kursiv

Kennzeichnet neue Begriffe, URLs, E-Mail-Adressen, Dateien 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.

Codebeispiele verwenden

Ergänzungsmaterial (Codebeispiele, Übungen usw.) steht zum Download unter <https://nlproc.info/PyTorchNLPBook/repo/> bereit.

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, außer wenn Sie einen erheblichen Teil des Codes kopieren. Wenn Sie zum Beispiel ein Programm schreiben, das einige Codeblöcke aus diesem Buch verwendet, benötigen Sie keine Erlaubnis. Sollten Sie aber eine CD-ROM mit den Beispielen von

O'Reilly-Büchern verkaufen oder verteilen, ist eine Erlaubnis erforderlich. Wenn Sie eine Frage beantworten und dabei dieses Buch oder Beispielcode zu diesem Buch zitieren, brauchen Sie wiederum keine Erlaubnis. Aber wenn Sie erhebliche 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: »*Natural Language Processing mit PyTorch* von Delip Rao und Brian McMahan (O'Reilly). Copyright 2020, ISBN 978-3-96009-118-9.«

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.

Danksagungen

Dieses Buch hat eine Entwicklung durchgemacht, wobei jede Version des Buchs anders als die vorherige Version aussah. An jeder Version waren verschiedene Personen (und sogar verschiedene Frameworks für Deep Learning) beteiligt.

Die Autoren möchten Gokul Mohandas für sein anfängliches Mitwirken am Buch danken. Gokul hat viel Energie in das Projekt gesteckt, bevor er es aus beruflichen Gründen leider verlassen musste. Gokuls Enthusiasmus für PyTorch und seine positive Einstellung sind unübertroffen, und die Autoren haben ihn schmerzlich vermisst. Wir erwarten noch große Dinge von ihm!

Das Buch wäre nicht in bester technischer Form, wenn es nicht das freundliche und dennoch hochqualitative Feedback unserer technischen Gutachter, Liling Tan und Debasish Gosh, gegeben hätte. Liling hat seine Kompetenz bei der Entwicklung von Produkten mit modernem NLP beigetragen, während das wertvolle Feedback von Debasish aus der Sicht des Entwicklers stammt. Dankbar sind wir auch für die Unterstützung von Alfredo Canziani, Soumith Chintala und den vielen anderen tollen Menschen in den Entwicklerforen von PyTorch. Darüber hinaus haben wir von den täglichen, fruchtbaren NLP-Gesprächen unter der Twitter-Gemeinde *#nlproc* profitiert. Viele Erkenntnisse dieses Buchs sind sowohl dieser Community als auch unserer persönlichen Praxis zuzuschreiben.

Keinesfalls möchten wir versäumen, Jeff Bleiel für seine ausgezeichnete Unterstützung als unser Redakteur zu danken. Ohne seine Führung hatte dieses Buch nicht das Licht der Welt erblickt. Die Lektorate von Bob Russell und die Produktionsunterstützung durch Nan Barber haben dieses Manuskript von einem Rohentwurf zu einem druckfähigen Buch gemacht. Außerdem möchten wir Shannon Cutt für ihre Hilfe in den frühen Phasen des Buchs danken.

Ein Großteil des Materials im Buch entstand aus dem zweitägigen NLP-Training, das die Autoren auf den AI- und Strata-Konferenzen von O'Reilly abgehalten haben.

Wir möchten Ben Lorica, Jason Perdue und Sophia DeMartini dafür danken, dass sie mit uns an der Gestaltung der Trainings zusammengearbeitet haben.

Delip ist dankbar dafür, Brian McMahan als Co-Autor gewonnen zu haben. Brian tat alles, um die Entwicklung des Buchs zu unterstützen. Für beide waren es wertvolle Erfahrungen, sie haben die Freude und die Schmerzen an der Entwicklung geteilt! Delip möchte auch Ben Lorica bei O'Reilly dafür danken, dass er darauf beharrt hat, ein Buch über NLP zu schreiben.

Brian möchte Sara Manuel für ihre unendliche Unterstützung danken und Delip Rao dafür, dass er der Motor war, der dieses Buch zur Vollendung gebracht hat. Ohne seine nicht nachlassende Beharrlichkeit und Charakterstärke wäre dieses Buch nicht möglich gewesen.