Federated Learning

Heiko Ludwig • Nathalie Baracaldo
Editors

Federated Learning

A Comprehensive Overview of Methods and Applications

Springer

Editors
Heiko Ludwig
IBM Research – Almaden
San Jose, CA, USA

Nathalie Baracaldo
IBM Research – Almaden
San Jose, CA, USA

ISBN 978-3-030-96895-3 ISBN 978-3-030-96896-0 (eBook)
https://doi.org/10.1007/978-3-030-96896-0

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Machine learning has made great strides over the past two decades and has been adopted in many application domains. Successful machine learning depends largely on access to quality data, both labeled and unlabeled.

Concerns related to data privacy, security, and sovereignty have caused public and technical discussion on how to use data for machine learning purposes consistent with regulatory and stakeholder interests. These concerns and legislation have led to the realization that collecting training data in large central repositories may be at odds with maintaining privacy for data owners.

While distributed learning or model fusion has been discussed since at least a decade, federated machine learning (FL) as a concept has been popularized by MacMahan and others since 2017. In the subsequent years, much research has been conducted – both, in academia and the industry – and, at the time of writing this book, the first viable commercial frameworks for federated learning are coming to the market.

This book aims to capture the research progress and state of the art that has been made in the past years, from the initial conception of the field to first applications and commercial use. To get this broad and deep overview, we invited leading researchers to address the different perspectives of federated learning: the core machine learning perspective, privacy and security, distributed systems, and specific application domains.

The book's title, *Federated Learning: A Comprehensive Overview of Methods and Applications*, outlines its scope. It presents in depth the most important issues and approaches to federated learning for researchers and practitioners. Some chapters contain a variety of technical content that is relevant to understand the intricacies of the algorithms and paradigms that make it possible to deploy federated learning in multiple enterprise settings. Other chapters focus on providing clarity on how to select privacy and security solutions in a way that can be tailored to specific use cases, while others take into consideration the pragmatics of the systems where the federated learning process will run.

Given the inherent cross-disciplinary nature of the topic, we encounter different notational conventions in different chapters of the book. What might be parties in

federated machine learning may be called clients in the distributed systems perspectives. In the introductory chapter of this book, we lay out the primary terminology we use, and each chapter explains how the discipline-specific terminology maps to the common one when it is introduced, if this is the case. With this approach, we make this book understandable to readers from diverse backgrounds while staying true to the conventions of the specific disciplines involved.

Taken as a whole, this book enables the reader to get a broad state-of-the-art summary of the most recent research developments.

Editing this book, and writing some of the chapters, required the help of many, who we want to acknowledge. IBM Research gave us the opportunity to work in this exciting field, not just academically but also to put this technology into practice and make it part of a product. We learned invaluable lessons on the journey, and we have much to thank to our colleagues at IBM. In particular, we want to acknowledge our director, *Sandeep Gopisetty*, for giving us the space to work on this book: Gegi Thomas, who made sure our research contributions make their way into the product: and our team members.

The chapter authors provide the substance of this book and were patient with us with requests for changes to their chapters.

We owe greatest thanks to our families, who patiently put up with us devoting time to the book rather than them over the year of writing and editing this book. Heiko is deeply thankful to his wife, *Beatriz Raggio*, for making these sacrifices and supporting him throughout. Nathalie is profoundly thankful to her husband and sons, *Santiago* and *Matthias Bock*, for their love and support and for cheering for all her projects, including this one. She also thanks her parents, Adriana and Jesus; this and many more achievements would not be possible without their amazing and continuous support.

San Jose, CA, USA
September 2021

Heiko Ludwig
Nathalie Baracaldo

Contents