Automated Machine Learning and Meta-Learning for Multimedia

Wenwu Zhu • Xin Wang

Automated Machine Learning and Meta-Learning for Multimedia

Springer

Wenwu Zhu
Department of Computer Science
and Technology
Tsinghua University
Beijing, China

Xin Wang
Department of Computer Science
and Technology
Tsinghua University
Beijing, China

ISBN 978-3-030-88131-3 ISBN 978-3-030-88132-0 (eBook)
https://doi.org/10.1007/978-3-030-88132-0

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2021
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

Foreword by Martin Ester

Machine learning (ML) has achieved spectacular breakthroughs in many real-life applications, including image classification, machine translation, and robotics. Deep neural networks have been particularly successful, due to their great capacity to approximate extremely complex, non-linear functions. Neural networks achieve this capacity through increasingly diverse and complex network architectures. However, the great potential comes at a hefty price: the development of machine learning models does not only require suitable training datasets and knowledge of the application domain but also deep knowledge of the machine learning methods to actually leverage their full potential.

Automated machine learning (AutoML) is a research direction that has recently emerged in response to the ever-increasing complexity of ML models and their development, aiming to automate the development process as much as possible. Goals are, in particular, to automatically tune the many hyperparameters, e.g., the number and size of layers, and to determine the most appropriate architecture of a neural network, e.g., a convolutional network for feature extraction combined with a fully connected network for classification. A related goal is meta-learning, i.e., learning to learn, which promises to reduce the effort of model development by transferring a model from a source domain to a target domain. While AutoML is still a fairly young research area, neural networks developed through AutoML have already achieved performance comparable to that of neural networks handcrafted by data scientists in some applications. Since 2015, the research community has organized the AutoML Challenge, which has provided a benchmark and much stimulation to the field. Several ML development tools, including RapidMiner and Microsoft Azure, have already implemented the features of AutoML. In conclusion, AutoML is a promising direction in ML that is expected to mature in the years to come.

This timely book by two experts in the field introduces the state-of-the-art in AutoML with a focus on it for multimedia data. Wenwu Zhu is a Professor at the Department of Computer Science at Tsinghua University and is widely recognized for his research in the areas of multimedia networking and computing as well as multimedia big data. Xin Wang is an Assistant Professor at the Department of

Computer Science at Tsinghua University. Multimedia data, including image, video, audio, and text data, is much more complex in nature than structured data such as records stored in a relational database, and multimedia has been the domain where deep neural networks have had the greatest impact. In addition to being unstructured, multimedia data is typically very large and multimodal, i.e., combines various types of multimedia data, e.g., images and text. Finally, multimedia data is often served in a streaming fashion, i.e., large amounts of data arrive rapidly and have to be processed online. I highly recommend this book to anyone who wants to understand the state-of-the-art in AutoML, in particular the special challenges and methods for multimedia data.

Professor, Simon Fraser University Martin Ester
Burnaby, BC, Canada

Foreword by Steven C.H. Hoi

As a fundamental subset of AI techniques, machine learning has drawn popular attention from both academia and industry and made significant impact in real-world applications. This book covers two important and closely related subfields of machine learning, automated machine learning (AutoML) and meta-learning, which have been actively studied in recent years.

AutoML aims to automate the task of applying machine learning to solve a real-world problem. For example, one popular technique in AutoML is hyper-parameter optimization (HPO), which aims to automatically choose optimal hyperparameters for a learning algorithm. Another well-known AutoML technique is neural architecture search (NAS) for deep learning or deep neural network (DNN), which aims to automate the design of deep learning architectures.

Meta-learning, also known as learning to learn, aims to design a model that can learn new skills or adapt to new environments rapidly with limited training data. Meta-learning can be applied to tackle the AutoML tasks, such as HPO and NAS. In addition, meta-learning can be used for several other kinds of machine learning tasks and real-world applications, such as cold-start recommendation in multimedia and few-shot learning in computer vision and NLP.

This book provides a comprehensive understanding of AutoML and meta-learning methods and their applications. It is organized into two parts. Part I covers the subjects on the fundamentals of AutoML and meta-learning methodologies, including basics of some popular algorithms and recent advances in machine learning. Part II covers the subjects on applying AutoML and meta-learning techniques for a range of application domains, such as computer vision, natural language processing, multimedia, data mining, and recommender systems.

The authors are established AI experts and researchers with extensive experiences in investigating machine learning techniques for real-world applications. This book is strongly recommended for AI researchers, engineers, graduate students, and any readers who are interested in learning advanced machine learning subjects in AutoML and meta-learning.

Managing Director, Salesforce Research Asia Steven C.H. Hoi
Professor, Singapore Management University
Singapore

Foreword by Tong Zhang

I have known professor Wenwu Zhu for many years, and we have collaborated on a number of problems. Professor Zhu is a highly regarded scientist in multimedia and big data research. He has not only published many influential scientific papers but also worked on practical problems in the industrial setting.

Both automated machine learning and meta-learning are emerging topics in machine learning, which have drawn significant attention in recent years due to their many practical applications. This book presents a comprehensive overview of the recent advances in these subjects as well as their applications. The book contains two parts. Part I presents a unified view of basic concepts and many recently proposed algorithms that are scattered in the literature. This helps the readers to quickly grasp the basic concepts and algorithmic foundations. Part II contains case studies and applications that help the readers to understand how these methods can be applied to real-world problems. Although examples in this book focused on multimedia applications, the material should greatly benefit all researchers and practitioners who want to learn these advanced machine learning methods through practical examples.

Chair Professor
The Hong Kong University of Science and Technology
Hong Kong, China

Tong Zhang

Preface

This book is to disseminate and promote the recent research progress and frontier development in AutoML and meta-learning as well as their applications in computer vision, natural language processing, multimedia, and data mining-related fields, which are exciting and fast-growing research directions in the general field of machine learning. We will advocate novel, high-quality research findings and innovative solutions to the challenging problems in AutoML and meta-learning. This topic is at the core of the scope of artificial intelligence and is attractive to audience from both academia and industry.

Our efforts in writing this book is motivated by the following reasons. First of all, the topics on meta-learning and AutoML are very new emerging topics, which urgently requires a well-organized monograph on these topics. Second, several current viewpoints may treat neural architecture search (NAS) and Bayesian optimization (BO), two important techniques in AutoML, as components in meta-learning. Our book differs from them by regarding AutoML and meta-learning as two parallel tools that can enhance each other. Third, in this book, we will discuss more recent advances in AutoML and meta-learning, such as continual learning, hardware-aware architecture search, and automated graph learning. Last but not the least, this book also focuses on the applications of AutoML and meta-learning in many research fields, such as computer vision, natural language processing, and multimedia etc.

Therefore, we deeply hope that this book can benefit interested readers from both academy and industry, covering the needs from junior starters in research to senior practitioners in IT companies.

Beijing, China Wenwu Zhu

Beijing, China Xin Wang
June, 2021

Acknowledgments

We thank Wenpeng Zhang for participating in the NeurIPS contest on AutoML in 2018, which initiated our research on AutoML and meta-learning. Many thanks to Jie Cai, Yudong Chen, Kun Cheng, Chaoyu Guan, Jiyan Jiang, Yuhang Jiang, Xiaohan Lan, Haoyang Li, Yue Liu, Yijian Qin, Guangyao Shen, Zheng Xiong, Yitian Yuan, Zeyang Zhang, Ziwei Zhang, and Shiji Zhou from the Tsinghua University for their contributions during the writing of this book.

We would also like to express our sincere thanks to Martin Ester from the Simon Fraser University, Steven C.H. Hoi from the Salesforce Asia/Singapore Management University, and Tong Zhang from The Hong Kong University of Science and Technology for writing forewords for this book.

Xin Wang is the corresponding author.

Contents

List of Figures

List of Tables