AutoML and Meta-learning for Multimedia

Wenwu Zhu
wwzhu@tsinghua.edu.cn
Tsinghua University

Xin Wang
xin_wang@tsinghua.edu.cn
Tsinghua University

Wenpeng Zhang
zhangwenpeng0@gmail.com
Tsinghua University

ABSTRACT

AutoML and meta-learning are exciting and fast-growing research directions to the research community in both academia and industry. This tutorial is to disseminate and promote the recent research achievements on AutoML and meta-learning as well as their potential applications for multimedia. Specifically, we will first advocate novel, high-quality research findings and innovative solutions to the challenging problems in AutoML and meta-learning. Then we will discuss scenarios of multimedia where AutoML and meta-learning serve as candidates for solutions. Finally, we will point out future research directions on AutoML and meta-learning as well as their potential new applications for multimedia.

ACM Reference Format:
Wenwu Zhu, Xin Wang, and Wenpeng Zhang. 2019. AutoML and Meta-learning for Multimedia. In *Proceedings of the 27th ACM Int'l Conf. on Multimedia (MM'19), Oct. 21–25, 2019, Nice, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3343031.3350541

1 THEME AND TOPICS

AutoML and metal-learning have various applications in quite a few areas of multimedia and they should be of interest to many of the conference attendees ranging from students, researchers, to practitioners. The tutorial is planned for half-day and its topic areas include (but are not limited to) the following:

- The academic and industrial motivation
- Hyperparameter optimization for multimedia applications
- Neural architecture search and its applications in multimedia
- Meta-learning for multimedia
- Future directions

2 TARGET AUDIENCE AND PREREQUISITES

This tutorial will be highly accessible to the whole multimedia community, including researchers, students and practitioners who are interested in AutoML, meta-learning and their applications in multimedia related tasks. The tutorial will be self-contained and designed for introductory and intermediate audiences. No special prerequisite knowledge is required to attend this tutorial.

3 TUTORIAL OVERVIEW

All the slides and our experiences in winning the second place in NeuralIPS 2018 AutoML Competition will be shared with all the audiences.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
MM '19, October 21–25, 2019, Nice, France
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6889-6/19/10.
https://doi.org/10.1145/3343031.3350541

3.1 Hyperparameter Optimization

Every machine learning system has hyperparameters. The choice of the hyperparameters significantly affects the effectiveness of the learning system. Especially in deep learning systems, there can be tens of thousands of hyperparameters regarding neural architecture, regularization and optimization. Finding suitable hyperparameters often requires expert knowledge and sufficient experience in the ML system, which prohibits the feasibility and efficiency of ML system in real-world application fields. Hyperparameter Optimization (HPO) aims to automatically select the optimal configurations of hyperparameters. It can reduce the human effort and improve the performance and reproducibility of the machine learning systems.

3.1.1 Model-free Methods. The most basic model-free HPO methods include grid search and random search. Population-based methods are another important branch of model-free HPO methods, which include genetic algorithms, evolutionary algorithms, etc. Model-free methods are usually simple for implementations and free from specific assumptions on the machine learning system being optimized.

3.1.2 Bayesian Optimization. Bayesian optimization (BO) is a state-of-the-art framework for optimizing blackbox functions. It includes a probabilistic model which is fitted to the observations, as well as an acquisition function that determines the candidate optimal hyperparameters by using the probabilistic model. Commonly used probabilistic models include Gaussian processes (GP) and the Tree Parzen Estimator (TPE). The most common choice of the acquisition function is the expected improvement (EI). BO forms the basis of several prominent AutoML frameworks, such as Auto-WEKA and Auto-sklearn.

3.1.3 Bandit-Based Methods. In blackbox HPO methods such as BO, the performance evaluation can be very expensive, especially for large dataset sizes and complex models: each evaluation requires an entire running procedure of the machine learning system being optimized. To tackle this problem, bandit-based algorithms are proposed to cut off the less promising configurations early and focus on configurations that are expected to be optimal. Bandit-based HPO methods are built based on the assumption that a small subset of data is sufficient to indicate the final performance of the candidate configurations. A simple yet powerful method is the successive halving (SH). HyperBand further considers different trade-off between the total budget and the number of candidate configurations to improve effectiveness of SH. BOHB combines Bayesian optimization and HyperBand to improve anytime performance and final performance simultaneously.

3.2 Neural Architecture Search

Deep learning methods are very successful in solving tasks in machine translation, image/speech recognition and reinforcement learning in general. Neural Architecture Search (NAS), the process of automating architecture engineering, is important in automating machine learning. NAS can be seen as a subfield of AutoML and has significant overlap with hyperparameter optimization and meta-learning. The methods for NAS can be categorized according to three dimensions: search space, optimization methods and performance estimation strategy. The NAS part of this tutorial is structured with respect to these three dimensions.

3.2.1 Search Space. The search space defines which architectures can be represented in principle. The optimization problem of NAS is often non-continuous and high-dimensional, which can be largely simplified by a suitable choice of the search space since the size of the search space can be reduced by incorporating prior knowledge on the specific task.

3.2.2 Search Strategy. The search strategy is about the exploration of the search space. There have been several different search strategies for the exploration of the search space, such as reinforcement learning, one-shot architecture search and evolutionary algorithms.

3.2.3 Performance Estimation Strategy. To guide the search process, Performance Estimation Strategy is used to estimate this performance of a given architecture. The simplest way is to perform a standard training and validation of the architecture on data. However, it can consume thousands of GPU days for NAS to train each architecture from scratch, which is computationally expensive. Hence, many methods have been proposed to reduce the cost of performance estimations recently.

3.3 Meta-learning

Meta-learning refers to utilizing past experience from solving the related tasks to facilite the task being solved. In meta-learning, meta-data is collected to describe previous tasks and models. Then the meta-data is utilized to guide the search for optimal models for the new tasks. The meta-learning methods can be classified by which kind of meta-data is used.

3.3.1 Learning from Meta-features. The learning tasks can be characterized with various properties (i.e. meta-features) and represented by a vector. Then the task similarity can be measured by the distance between the vectors. With this approach, information from previous similar tasks can be transferred to the new task. Many existing works also investigate the relation between the properties of a task, its possible configurations and the corresponding utility. In these works, meta-models are learned to capture the relation and predict promising configurations on the specific tasks.

3.3.2 Learning from Previous Models. Transfer learning is a basic method for utilizing the knowledge of previous models on related tasks. In transfer learning, models are trained on source tasks and used as the initializations on a similar target task. This general framework can be fitted into a wide range of areas such as kernel methods, Bayesian models, reinforcement learning and especially deep learning. Model-agnostic meta-learning (MAML) directly learns a model initialization that generalizes better to similar tasks. It is especially suitable in few-shot learning scenarios where the training examples in the task being solved are very limited. Reptile accelerates MAML by replacing gradient steps with averaging when updating the parameters of the general model.

4 ORGANIZERS

Wenwu Zhu

Wenwu Zhu is currently a Professor and the Vice Chair of the Department of Computer Science and Technology at Tsinghua University, the Vice Dean of National Research Center for Information Science and Technology, and the Vice Director of Tsinghua Center for Big Data. Prior to his current post, he was a Senior Researcher and Research Manager at Microsoft Research Asia. He was the Chief Scientist and Director at Intel Research China from 2004 to 2008. He worked at Bell Labs New Jersey as Member of Technical Staff during 1996-1999. He received his Ph.D. degree from New York University in 1996 in Electrical and Computer Engineering.

Wenwu Zhu is an AAAS Fellow, IEEE Fellow, SPIE Fellow, and a member of The Academy of Europe (Academia Europaea). He has published over 300 referred papers in the areas of multimedia computing, communications and networking, and big data. He is inventor or co-inventor of over 50 patents. He received seven Best Paper Awards, including ACM Multimedia 2012 and IEEE Transactions on Circuits and Systems for Video Technology in 2001. His current research interests are in the area of Cyber-Physical-Human big data computing, and Cross-media big data and intelligence.

Wenwu Zhu currently serves as EiC for IEEE Transactions on Multimedia. He served as Guest Editors for the Proceedings of the IEEE, IEEE Journal on Selected Areas in Communications, ACM Transactions on Intelligent Systems and Technology, etc.; and Associate Editors for IEEE Transactions on Mobile Computing, ACM Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology, and IEEE Transactions on Big Data, etc. He served in the steering committee for IEEE Transactions on Multimedia (2015-2016) and IEEE Transactions on Mobile Computing (2007-2010), respectively. He served as TPC Co-chair for ACM Multimedia 2014 and IEEE ISCAS 2013, respectively. He serves as General Co-Chair for ACM Multimedia 2018 and ACM CIKM 2019, respectively.

Xin Wang

Xin Wang is currently a Postdoc researcher at the Department of Computer Science and Technology, Tsinghua University. He got both of his Ph.D. and B.E degrees in Computer Science and Technology from Zhejiang University, China. He also holds a Ph.D. degree in Computing Science from Simon Fraser University, Canada. His research interests include cross-modal multimedia intelligence and inferable recommendation in social media. He has published several high-quality research papers in top conferences including ICML, MM, KDD, WWW, SIGIR etc. He is the recipient of 2017 China Postdoctoral innovative talents supporting program.

Wenpeng Zhang

Wenpeng Zhang obtained his Ph.D degree in machine learning at Tsinghua University, China in 2018. He has published several papers in top tier conferences and journals, including ICML, WWW, KDD, TKDE etc. Now, his research interests lie in online learning, Meta-learning and AutoML. He led the team Meta_Learners that won the second place in the NeuralPS 2018 AutoML Competition.