Meta-learning Achieves High Accuracy with a Small Amount of Training Data

Tomoharu Iwata
Senior Distinguished Researcher, NTT Communication Science Laboratories

Abstract

Large language models and generative artificial intelligence for images are increasing in popularity. Their accuracy is improved by training them with a vast amount of documents and images. If vast amounts of data are unavailable, however, it is difficult to achieve high accuracy. Meta-learning for attaining high accuracy through learning to learn with a small amount of training data has been attracting attention as an approach to solving this problem. We interviewed Tomoharu Iwata, a senior distinguished researcher at NTT Communication Science Laboratories, about research trends and future directions concerning meta-learning and his mindset of staying interested in and enjoying research.

Keywords: artificial intelligence, meta-learning, machine learning

Meta-learning—learning how to learn—expands application areas of machine learning

—*Would you tell us about the research you are currently conducting?*

In the field of machine learning, I'm researching meta-learning for achieving high-performance artificial intelligence (AI) using a small amount of training data. Since I joined NTT in 2003, I have been researching machine learning, and since around 2018, I have been focusing on meta-learning. The concept of meta-learning has been around for some time; however, around 2018, when deep-learning research was progressing apace, computer performance was improving, and various deep-learning libraries became available, meta-learning research gradually gained more attention.

Deep learning is widely used in natural-language processing and image processing, and achieves high performance by learning with large amounts of data. In contrast to deep learning, meta-learning learns how to learn from other related data, so its performance in regard to new tasks can be improved even when only a small amount of training data are available.

I will explain the basic framework of meta-learning by using an image-classification task as an example. For meta-learning, data for various tasks are prepared. An example task is classifying images of cats and dogs or cars and bicycles. For each of these tasks, when task-specific model parameters are learned
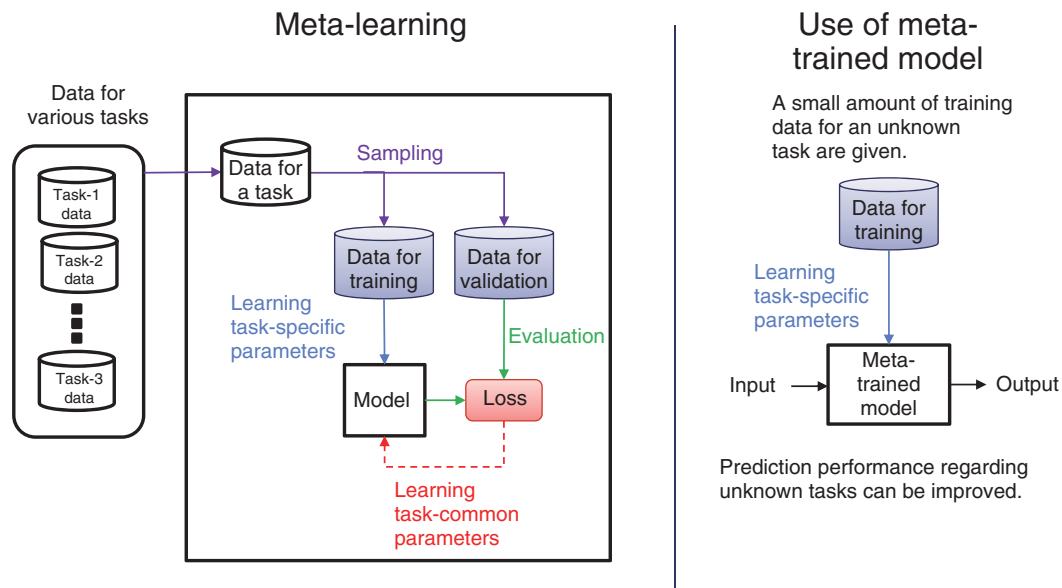
Meta-learning

Use of meta-trained model

Fig. 1.   Framework of meta-learning.

Conventional meta-learning:
Meta-learning is possible only
with data in the same format

Meta-learning we aim for:
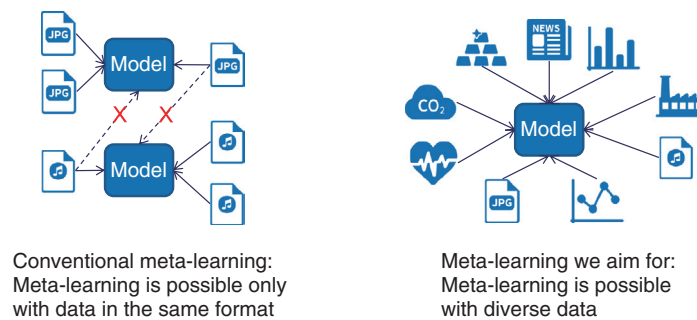Meta-learning is possible
with diverse data

Fig. 2.   Meta-learning from tasks in different feature spaces.

using a small amount of training data, the model parameters common to the tasks are updated to improve classification performance with the validation data. This process is referred to as "learning how to learn." Even if the task is a new one that has not been used during meta-learning (for example, classification of images of apples and oranges), it is possible to achieve high classification performance by simply learning task-specific model parameters with a small amount of training data (**Fig. 1**). In addition to the above example, meta-learning is also used for speaker recognition, which involves identifying who is speaking from audio data, and speakers can be classified even when a new speaker suddenly starts to speak; analyzing a language for which little data

exist; and automatically classifying books by using only a small amount of data with a new book-classification system that differs from the previous systems.

In the course of my research on meta-learning, my colleague and I proposed a deep-learning model that enables meta-learning from tasks in different feature spaces (**Fig. 2**)—in contrast to existing meta-learning that assumes meta-learning from tasks in the same feature space—and presented the model at the Annual Conference on Neural Information Processing Systems (NeurIPS), a top machine-learning conference, in 2020 [1]. We also proposed a deep-learning model that combines large language models with meta-learning in a manner that enables the use of

human-accumulated knowledge in machine learning and presented that model at NeurIPS 2022 [2]. I've also been conducting research to improve machine-learning performance by applying meta-learning to a variety of tasks such as clustering, uncertainty estimation, spatial analysis, causal inference, anomaly detection, and feature selection.

—*Meta-learning can overcome the weaknesses of deep learning. What direction will its research and applications take in the future?*

Machine learning and AI have made remarkable progress, and generative AI, such as ChatGPT, has emerged, but their success is based on having a large amount of data for training them. However, there are many cases in which sufficient data cannot be obtained due to the high cost of requiring expert knowledge, such as in medical images, and due to privacy protection. There is also a case in which there are almost no data because it is a new product or service in recommendation systems. Meta-learning makes it possible to use machine learning in application areas where only such a small amount of data are available.

Our meta-learning can currently handle a variety of data that can be expressed in tabular form but cannot handle more complex data such as data obtained on the factory floor or biometric data. Therefore, we are investigating meta-learning to expand the range of data that it can handle, which should allow us to expand the application area of meta-learning. In regard to the above-mentioned research on meta-learning using large language models, our research on meta-learning also aims to create AI that can address new challenges by using a variety of experiences and knowledge in the same manner as humans do.

Enjoying research while staying interested in it

—*What do you keep in mind when choosing a research theme?*

I choose research themes on the basis of my own curiosity. The meta-learning that I'm currently focusing on is also relevant to the evolution of life, which I was interested in and studied when I was a student. I hope that research on meta-learning, by which AI learns how to learn from various tasks, will also lead to understanding how humans have evolved to be able to learn from various things.

I used to read papers in areas of my interest and found problems to turn them into new research themes. I also combined things I was good at with things my colleagues were good at to create new themes. I sometimes found problems when I looked at new data and turned them into themes. When I was previously researching recommendation systems, for example, such systems were designed to recommend people to buy more products or use more services. However, in recommendation systems for subscription-type services or flat-rate services, a different perspective emerged, that is, the goal of recommendation systems is not to get people to buy more but increase their satisfaction.

When I read other researcher's papers, I naturally think about how the findings reported in those papers could be expanded or whether I could combine them with meta-learning, my specialty, or with the probabilistic models and topic models that I had investigated in the past. By looking at other researcher's papers and research on the basis of my specialty, I come up with new perspectives and new research themes. About ten years ago, I was posted to the UK for a year as a visiting researcher. Since I thought it was a great opportunity to collaborate with the people there, I asked students and post-doctoral researchers what they were researching, and I wondered if I could combine their research with my own specialties. I was therefore able to learn new things, write several papers on new themes, and enjoy life in the UK through researching with a variety of people.

—*What do you keep in mind as a researcher?*

I'm always willing to try different things. I like inventing and implementing new methods, so I come up with a lot of ideas, experiment with them, and repeat trial and error. Naturally, I often fail, but I believe that by trying new things, I can learn a lot and acquire unexpected knowledge that will lead to future research.

When researching machine learning, I want to solve problems as clearly and as simply as possible. I first clarify a goal in regard to the problem I want to solve, determine the loss that can be quantitatively evaluated, and minimize the loss. Solving problems neatly on the basis of theories, such as probability theory, will improve the performance of a machine-learning model. Its evolvability and extensibility will also increase in a way that leads to subsequent research and contributes to the research of others. Since that approach is easy to understand, I try to

create a model by identifying only the important points so that the model can deliver dependable performance with the minimum amount of work required.

When I joined NTT, I was assigned to the machine-learning group, which was a completely different research field from the one I studied at my graduate school. At first, I didn't know anything about machine learning; however, my mentor gave me step-by-step guidance, and I was able to give a presentation at a prestigious international conference in my second year. During this process, I was able to learn about the format of machine-learning research, which starts with problem setting, continues through development and implementation of a model, experimentation, and ends with writing a paper and presenting it at conferences. After that, I came up with my own ideas many times and presented them to my mentor. At first, my ideas were constantly rejected, but as I kept repeating this process, I was able to acquire the ability to research proactively. My encounters with my mentor had a great effect on my life as a researcher, and I hope to have encounters in which I can have a similar effect on others. I would be happy if younger researchers grow as they create their own research styles by referring to other researchers, come up with their own ideas, and have straightforward discussions.

—*Do you have a message for younger researchers?*

My message is enjoy your research by staying interested in it. Because I enjoy doing research that I'm interested in, I can concentrate when I'm researching and find it rewarding, even if I'm exhausted. I enjoy learning new things that interest me, and I think it is exceptional that researchers have the opportunity to learn things that are yet unknown by investigating problems that have not yet been solved. It has been over 20 years since I joined NTT, and I think I have been able to continue my research by enjoying learning new things along the way. If you choose a research theme that interests you and enjoy learning about things you do not know in the course of your research, you will keep finding new interests and enjoyment.

References

[1] T. Iwata and A. Kumagai, "Meta-learning from Tasks with Heterogeneous Attribute Spaces," Proc. of NeurIPS 2020, Virtual, Dec. 2020.

[2] T. Iwata and A. Kumagai, "Sharing Knowledge for Meta-learning with Feature Descriptions," Proc. of NeurIPS 2022, New Orleans, LA, USA, Nov./Dec. 2022.

■ **Interviewee profile**

Tomoharu Iwata received a B.S. in environmental information from Keio University, Kanagawa, in 2001, M.S. in arts and sciences from the University of Tokyo in 2003, and Ph.D. in informatics from Kyoto University in 2008. In 2003, he joined NTT Communication Science Laboratories. From 2012 to 2013, he was a visiting researcher at University of Cambridge, UK. His research interests include data mining and machine learning.