# How to learn faster: towards better adaptation in Meta-Learning

Abstract for the general public


In recent years – especially the last decade – humanity has observed an enormous machine improvement that enabled more and more computationally demanding experiments. This rise of computational power and gathering enormous amounts of data was the game-changer in deep learning, making it the state-of-the-art approach to Artificial Intelligence.

The power of the standard deep learning approaches usually comes from the mixture of three components – deep and complex models, large datasets, and a specific, well-defined problem to solve. All of these components have a common root - our data have a large and sufficient number of examples per datapoint. However, this condition is not always true for most real-world problems and does not reflect the way humans learn. Moreover, contemporary machine learning models are usually designed and trained to solve one specific problem. Although such a common practice framework in deep learning results in undoubted successes in different subfields, it has limitations. For example, our models need expensive computations and large amounts of well-structured data. They are often vulnerable to long-tailed data distributions in real-world use cases, such as medicine, where large amounts of balanced data are a rarity. Finally, the growing dependency of state-of-the-art deep learning models on enormous computational resources means that only a few well-funded labs can develop such models.

Meta-Learning introduces the ”learning how to learn” paradigm, an alternative strategy where the model learns to adapt to numerous related tasks and gains experience over multiple different learning episodes. The model uses previous experiences to improve its learning performance when learning the new tasks. Most importantly, the ”learning how to learn” paradigm is analogous to humans and animals gaining experience and, as such, is a step toward bridging the gap between the biological and machine way of learning. Moreover, the paradigm is easily adaptable to real-world scenarios where large datasets are unavailable, but knowledge could be gained by learning a variety of slightly different tasks. Finally, the Meta-Learning approach allows for fewer data and lowers the number of computations needed to run the models, leading to the democratization of AI research and a positive impact on the environment.

The most typical problem of Meta-Learning is Few-Shot Learning, where we are given a collection of small tasks. Each task consists of a support set of annotated examples and a query set with examples to classify. The aim is to train a model that can adapt to various such tasks. Moreover, during evaluation, the examples in the task may come from classes never encountered by the model during training. When trying to solve this natural Meta-Learning problem, we come across common obstacles like the problem of flexible adaptation to the new specific tasks and novel out-of-distribution data.

The obstacles mentioned above are sources of the ideas for fast adaptation and broader generalization – the two principal directions of research in Meta-Learning. In this research project, we will focus on the first one, enhancing the speed and flexibility of models for the adaptation to different tasks. Driven by the idea behind the well-known Model-Agnostic Meta-Learning (MAML) method and the impressive results in Transfer Learning, we want to focus on the fine-tuning approaches. Recent studies suggest that the knowledge transfer via fine-tuning from the huge models could significantly outperform the existing Meta-Learning models. The most severe obstacle with such a procedure is the necessity of intensive pretraining of the models, requiring enormous amounts of data and computational power. However, despite being heavily utilized by MAML, fine-tuning is still unexplored in Meta-Learning. In particular, there is no general study showing the most effective way of better adaptation via fine-tuning.

In the experimental part of the research, we will utilize the broad spectrum of the contemporary deep learning frameworks such as, e.g., Bayesian approaches (Bayesian Neural Networks or Gaussian Processes), conditional Normalizing Flows, HyperNetworks, Transformers, and attention mechanism. We will present the applications of created methods on the Few-Shot learning and Continual Meta-Learning problems. Moreover, we hypothesize that fine-tuning could be helpful also for overcoming catastrophic forgetting in the Continual Learning.