On Learning to Learn: Meta-Learning in Machine Learning

Mikhail M.Meskhi

April 20, 2020

Pattern Analysis Laboratory
University of Houston

Introduction

Defining Meta-Learning i

Meta-learning refers to the study of the process of how learning algorithms improve over a distribution of different tasks based on past experience. [6]

In a typical machine learning setting, given a data set D, solve for a task t. Learning algorithm selection is often chosen based on the following:

1. Domain knowledge
2. Expert knowledge
3. Technical constraints
4. Other reasons

Defining Meta-Learning ii

The meta-learning framework consists of:

1. base-learner
2. meta-learner

The **base-learner** has a fixed bias or user defined parameters (i.e. neural network). The **meta-learner** focuses on choosing the right bias dynamically. For example, given a data set D of plane images, the base-learner focuses on solving given tasks t_n (i.e. classifying input as planes or not). The meta-learner focuses on adapting to the task distribution and modifying the base-learner appropriately (e.g. images now contain birds so the base-learner must learn to differentiate planes).

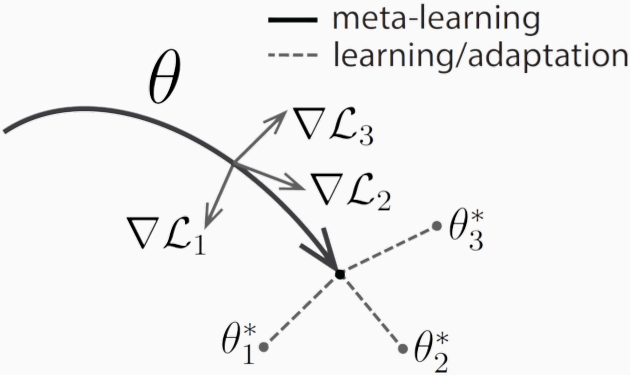

Figure 1: Meta-learning adaptation

Applications of Meta-Learning

Machine learning problems have recently seen approaches designed accordingly to meta-learning methodology. Below are a few of interesting implementations:

1. Learning Algorithm Selection (OpenML [4])
2. Few-Shot Learning (MAML [2])
3. Active Learning (SCA [1])
4. Architecture Search (NAS [3])

Recent Frameworks

Few-Shot Learning

In 2017, Model-Agnostic Meta-Learning [2] was introduced as a practical framework for few-shot learning problems. Few-shot learning is a supervised learning problem that exploit past information to adapt to newly unseen tasks with few examples as possible. MAML introduced a two-loop architecture where the base-learner is the inner loop (focuses on solving a task) and the outer loop is the meta-learner that learns an optimization for the base-learner as the distribution of tasks changes.

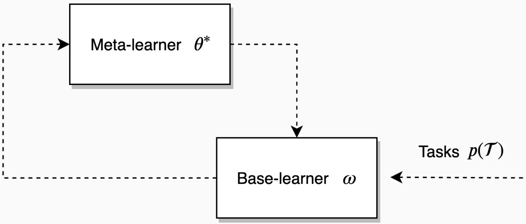

Figure 2: Meta-learning adaptation

OpenML

Vanschoren et al. [4] developed a scientific collaboration platform named OpenML. Platform Users can:

1. upload data,
2. define tasks,
3. run experiments, and
4. share results/failures

This is not a direct meta-learning framework, instead it is a knowledge base. Thousands of data sets along with tasks and their respective characteristics, various model performance measures and parameters are available to anyone. This is an excellent source of knowledge that a meta-learning framework can benefit from. Exploiting this knowledge base to recommend or even propose a new learning algorithm given a new data set and task, is an interesting approach that yet has to be fully implemented.

My Research

Data Characteristics

A data set D, can be described via a vector of **meta-features**. Meta-features are descriptive measures about a data set. Examples of types of meta-features:

- statistical (mean, kurtosis, skewness, ...)
- concept (concept concentration, concept variation, ...) [5]
- data complexity (entropy, decision boundary shape, ...)

They are able to offer insight into how complex the data set is and how the classes might be spread. This information can help identify what kind of learning algorithm is more appropriate to solve a task more efficiently.

Self-Adaptive Learning

My current research focus is on developing methods to eventually build a self-adaptive system that is able to when given a new data/task:

- Exploit past experience to build an appropriate learning algorithm for this new problem
- Modify internal components with respect to the task it is working on
- Assess learning quality and outcomes during and after solving a task
- Modify the knowledge base accordingly after solving current problem

Summary

Self-adaptive learning systems require:

- experience/knowledge bases
- learning quality assessment mechanism
- ability to characterize problems adequately
- dynamic reasoning based on evidence

A. Antoniou and A. J. Storkey.
Learning to learn by self-critique.
In *Advances in Neural Information Processing Systems*, pages 9936–9946, 2019.

C. Finn, P. Abbeel, and S. Levine.
Model-agnostic meta-learning for fast adaptation of deep networks.
In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1126–1135. JMLR. org, 2017.

J. G. Robles and J. Vanschoren.
Learning to reinforcement learn for neural architecture search.
arXiv preprint arXiv:1911.03769, 2019.

References ii

J. Vanschoren, J. N. Van Rijn, B. Bischl, and L. Torgo.
Openml: networked science in machine learning.
ACM SIGKDD Explorations Newsletter, 15(2):49–60, 2014.

R. Vilalta and Y. Drissi.
A characterization of difficult problems in classification.
2002.

R. Vilalta and Y. Drissi.
A perspective view and survey of meta-learning.
Artificial intelligence review, 18(2):77–95, 2002.

Thank you!

If a decision tree in a forest makes marginal improvements, and no one is around to publish it, is it really "state-of-the-art"?

Anonymous