Introduction to Learning to Learn

Meta-learning, Transfer Learning, AutoML and OpenML

Michael M.Meskhi

University of Houston - Department of Computer Science

Table of contents

Introduction

Learning to Learn

Let's start off with a simple demonstration of what is "learning".

Keywords:

- Knowledge acquisition
- Learning from examples
- Past experiences

Defining Learning

Definition: Learning is the process of acquiring new, or modifying existing, knowledge, behaviors, skills, values, or preferences [2].

How do humans learn? To put it in a simple way we learn about our environment but analyzing patterns and gluing things together as our brain tries to make sense out of the information acquired by our senses.

Learning Process

Sound $\mathcal{X} \rightarrow$ Processing in the brain $f(\mathcal{X}) \rightarrow \mathcal{Y}$ That's a plane!

Now, what does all of this have to do with machine learning? Well, the answer is quite simple! We are trying to build learning systems that have learning potentials such of humans. Are we not trying to achieve the ultimate goal of building a terminator?! I mean, AGI (Artificial General Intelligence).

So obviously you might ask, how can we learn from a human's learning process and transfer that into a machine learning process? I wish I had an answer... but what I do have are few research paths that look promising.

Types of Learning

Introduction

There are various learning to learn processes that machine learning is trying to solve. Today I will talk about these:

1. **Meta-learning**
2. **Transfer Learning**
3. **AutoML**

Each of these learning to learn methodologies try to solve various machine learning problems such as supervised, unsupervised and reinforcement learning.

Meta-learning

Enter **Meta-learning**. Meta-learning aims at discovering ways to dynamically search for the best learning strategy as the number of tasks increases [5].

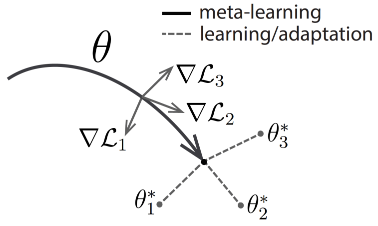

Figure 1: Given a learning algorithm $\mathcal{L}$ and θ, find $\nabla\mathcal{L}_k$ with updated θ_k^* in respect to given task distribution $\mathcal{T}$.

Meta-learning Example

Meta-learning has to do with past experience. How can we extract and store ”experience”? In our case experience refers to **meta-knowledge**. ”Meta” meaning itself and ”knowledge” making this ”knowledge of self”. We collect meta-knowledge in form of **meta-features**.

Take the iris data set. Extracting meta-features of the data set gives us this table. Using this meta-data we can build a mapping function between the meta-feature space $\mathcal{F}_{mf}$ and hypothesis space $\mathcal{H}_{\mathcal{L}}$. In simple terms, we want to find the optimal learning algorithm based on our inference of meta-features.

Meta-learning Examples

Algorithms:

- Model-Agnostic Meta-learning (MAML, MAML++ and ProMix-MAML)
- Reptile
- Meta-SGD

Tools:

- OpenML-Python
- Metalearn

Meta-learning Discussion

Now there is much more to meta-learning than this but for another day. To summarize what we discussed today:

- Meta-learning tries to create learning systems by inference based on extracted meta-features.
- Meta-learning has seen some advancement recently and seems promising.
- Meta-learning has been applied to supervised and reinforcement learning problems successfully.

To learn more about meta-learning, visit: MtL-Progress

Tansfer learning focuses on storing knowledge gained while solving one problem and applying it to a different but related problem [3].

To put it simply, transfer learning is when you have data set of poodle dog images $\mathcal{D}_{poodle}$ and train a learning system $\mathcal{L}$ on it. In return you have a built a model $\mathcal{L}_{poodle}$ with parameter θ_{poodle}. Now you want to use this model on a different data set with a similar problem (i.e. corgi dog breed). You would **transfer the model and parameters** and feed in $\mathcal{D}_{corgi}$ as a test data set. This is essentially transfer learning.

Transfer Learning Examples

There are various transfer learning methods to use:

- Inductive Transfer learning (Domain Adaptation)
- Unsupervised Transfer Learning (Zero-shot Learning)
- Transductive Transfer Learning (One-shot Learning)
- Instance transfer (Multitask Learning)
- Feature-representation transfer (Multitask Learning)
- Parameter transfer (Domain Adaptation)

Transfer Learning Discussion

Transfer learning is a popular approach in problems such as speech recognition, natural text processing and computer vision. The obvious answer is that when you train a complex model on big amount of data, it is near impossible for someone to retrain it on their own problem. This is where transfer learning offers to cut computational costs.

For example, Google can train a VGG-19 neural network on lots of images and offer this pre-trained model for free use. All one has to do is set up the experiment and feed it test data.

AutoML

Machine learning requires mastery from mathematics and computer science. It is quite simple to setup and build a simple logistic regression model but real world problems need way more than that. In the real world, you need to conduct a full scale problem research. Automated Machine Learning provides methods and processes to make Machine Learning available for non-Machine Learning experts, to improve efficiency of Machine Learning and to accelerate research on Machine Learning [1].

Let's give an example: Build a service that predicts office space usage.

Project Process:

1. Plan on how to collect data
2. Plan out data pipelines and how to ETL data
3. Data analysis and pattern detection
4. Feature engineering and model selection
5. Testing and validation
6. Deploy into production

AutoML Examples

Throughout recent years several off-the-shelf packages have been developed which provide automated machine learning. While there are more packages than the one listed below, we restrict ourselves to a subset of the most well-known ones:

- AutoWEKA
- Auto-sklearn
- TPOT
- H2O AutoML
- Google AutoML

The tools aforementioned automate the tedious and computationally exhaustive process of a complete machine learning pipeline.

AutoML Goals

What does AutoML actually try to achieve when they say ”*Automate the machine learning pipeline...*”? AutoML tries to automatically initialize, run and validate each of the following steps:

1. Data preparation and ingestion
2. Feature engineering
3. Model selection
4. Hyper-parameter optimization
5. Pipeline selection under time, memory, and complexity constraints
6. Selection of evaluation metrics and validation procedures
7. Problem checking
8. Analysis of results obtained

AutoML Discussion

AutoML's integration and take over is inevitable. This does not mean there will be less work for machine learning engineers or anyone else. This just means that lots of those annoying steps that we have to do (e.g. data cleaning, cross-validation, etc...) will be automated. This means we can refocus our time and effort onto more important things such as better problem understanding and building new algorithms.

Research Directions

Hot Topics

In **meta-learning**, currently developing a robust learning algorithm that can learn from experience and generalize well over a new unseen task is critical. Research areas are focused on the following:

- Develop ways to map learning algorithm performance to extracted meta-features of a given task
- Develop new class of topological meta-features that can unveil previously unseen relationships or data ”hot spots”
- A major challenge in the construction of self-adaptive learners is how to incorporate a flexible bias in both the base-learner and the meta-learner [6]

In **transfer learning**, transferring a model or its parameters is the fundamental step, but so far many things have been done on a small scale or on similarity assumptions between data sets and tasks. The following is open:

- Transfer knowledge across domains or tasks that have different feature spaces, and transfer from multiple such source domains
- Transfer learning techniques have been mainly applied to small scale applications with a limited variety, such as sensor-network-based localization, text classification and image classification problems. In the future, transfer learning techniques will be widely used to solve other challenging applications, such as video classification, social network analysis and logical inference [4]

In AutoML, the problems are more of a classic nature. In essence, the goal is to automate already existing solutions and methodologies in machine learning. Unfortunately, it is not as easy as it sounds because not only are you automating the technical aspect but also the decision making process that an individual expert carries.

- Making a science of model search argues that the performance of a given technique depends on both the fundamental quality of the algorithm and the details of its tuning and that it is sometimes difficult to know whether a given technique is genuinely better, or simply better tuned
- Hyperparameter optimization and algorithm configuration provide methods to automate the tedious, time-consuming and error-prone process of tuning hyperparameters to new tasks at hand [1]

Conclusion & Discussion

Conclusion

Future of machine learning is bright no doubt. Thanks to the Deep Learning explosion in 2010, machine learning has become a standard in our solutions and services. We managed to build algorithms and methods that take advantage of vast amounts of data our digital worlds produce on a daily basis. Couple weeks ago Tesla demonstrated their FSD (Full Self Driving) chip. The amounts of data they collect from their fleet and how they have build a complex web of neural networks that processes gigabytes of data per-second is amazing. Now imagine all of this being automated. Imagine AI building AI!

Discussion

I have a couple of questions for you:

1. What do you think will be a next breakthrough in machine learning?
2. Have you used AutoML? If so, what do you think of it?

OpenML

Check out OpenML. It is a online platform where you can update your data, set a task, do the experiment offline and upload results. It is made for result repdocution, meta-data collection and science community collaboration.

References

[1] AutoML.org. URL https://www.automl.org/.

[2] R. Gross. *Psychology: The Science of Mind and Behaviour 6th Edition*. Hodder Education, 2012. ISBN 9781444164367. URL https://books.google.com/books?id=Cle1Fcr_6_QC.

[3] Dan Ventura Jeremy, West and Sean Warnick. A theoretical foundation for inductive transfer. URL https://web.archive.org/web/20070801120743/http://cpms.byu.edu/springresearch/abstract-entry?id=861.

[4] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. URL https://www.uh.edu/~rvilalta/papers/icai01.pdf.

[5] Sebastian Thrun. *Lifelong Learning Algorithms*, pages 181–209. Springer US, Boston, MA, 1998. ISBN 978-1-4615-5529-2. doi: 10.1007/978-1-4615-5529-2_8. URL https://doi.org/10.1007/978-1-4615-5529-2_8.

[6] Ricardo Vilalta and Youssef Drissi. Research directions in meta-learning. URL https://www.uh.edu/~rvilalta/papers/icai01.pdf.