Model Antagonistic Meta-Learning (MAML) and First Order Model Antagonistic Meta-Learning (FOMAML) on few-shots classification problem

by Saumya Vilas Roy

1. INTRODUCTION

1.0.1. What is meta-learning?

In a nutshell Meta-learning is learning how to learn better using previous experience as it is 'learning to learn' algorithm, it utilizes metadata from the previous related or unrelated tasks to tune the training in a way which would be beneficial for other tasks as it would allow faster or more accurate convergence, it is quickly gaining a lot of research focus to develop the next generation of A.I.

1.0.0.1. Introduction to MAML

It is a model and task-agnostic algorithm for meta-learning that trains a model’s parameters such that a small number of gradient updates will lead to fast learning on a new task, it utilizes gradient decent algorithm to optimize the model parameter θ such that is highly adaptable to individual θ^* i</sub> for each task i.

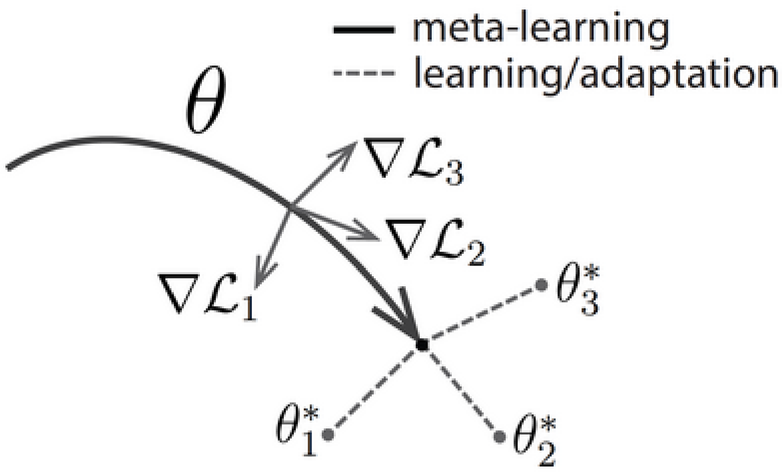


Diagram of the MAML approach.

1.0.1. What is first order approximation in MAML?

In MAML approach when we back propagate through back propagation it creates a hessian matrix term $\alpha\Delta_\theta^2 L_{T_i}(f_\theta)$, where L is differentiable, the second order derivative is ignored as assuming as the activations (maxout, leaky ReLU, ...) are piece-wise continuous so this shows if any composite function of such activations would also be locally linear, this allows us a very close approximation for MAML.

1.0.1. Experiment setup

For the this local experiment omniglot data-set is being used in batches to train, validate and test the respective models.

About Omniglot dataset

Omniglot data set for one-shot learning. This dataset contains 1623 different handwritten characters from 50 different alphabets. Each of the 1623 characters was drawn online via Amazon's Mechanical Turk by 20 different people. Each image is paired with stroke data, a sequences of [x,y,t] coordinates with time (t) in milliseconds.

To train learn2learn library was used

About learn2learn

learn2learn is a software library for meta-learning research.

learn2learn builds on top of PyTorch to accelerate two aspects of the meta-learning research cycle:

- *fast prototyping*, essential in letting researchers quickly try new ideas, and
- *correct reproducibility*, ensuring that these ideas are evaluated fairly.

learn2learn provides low-level utilities and unified interface to create new algorithms and domains, together with high-quality implementations of existing algorithms and standardized benchmarks. It retains compatibility with torchvision, torchaudio, torchtext, cherry, and any other PyTorch-based library you might be using.

With the hyperparamters: ways=5 (No. of classes to be classified), shots=5 (No. of images/samples in each class), meta_lr=0.003 (Meta-learning gradient decent update learning rate), fast_lr=0.5 (Learning rate while using MAML), meta_batch_size=32 (Batch size for the MAML training) , adaptation_steps=1 (No. of gradient steps for the fine tuning of the model parameter), num_iterations= 10000 (No. of Epochs for training), cuda=True (Setting the usage/availability of GPU), seed=42 (To generate same string of random numbers).

This training is done on NVIDIA GeForce GTX 1650 Mobile with Intel i7-9750H processor.

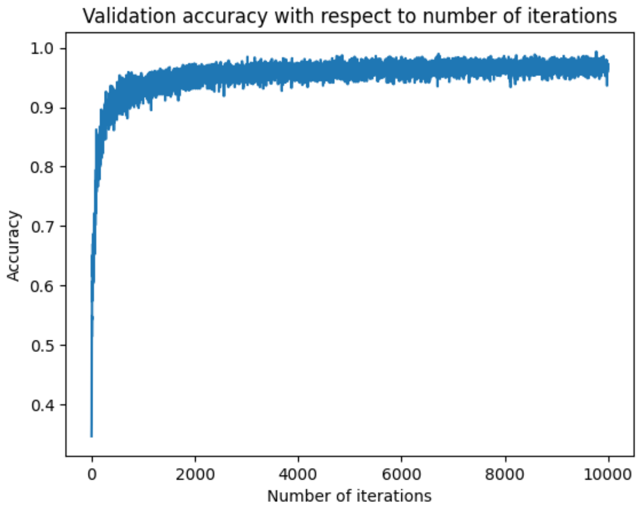
The running time has been sampled using the time library.

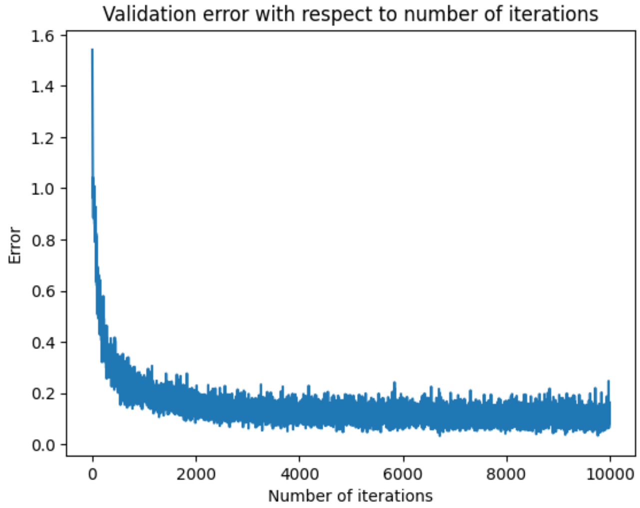
And the entire network has been built and deployed using PyTorch.

For FOMAML the tag first_order during MAML wrapper was set to True .

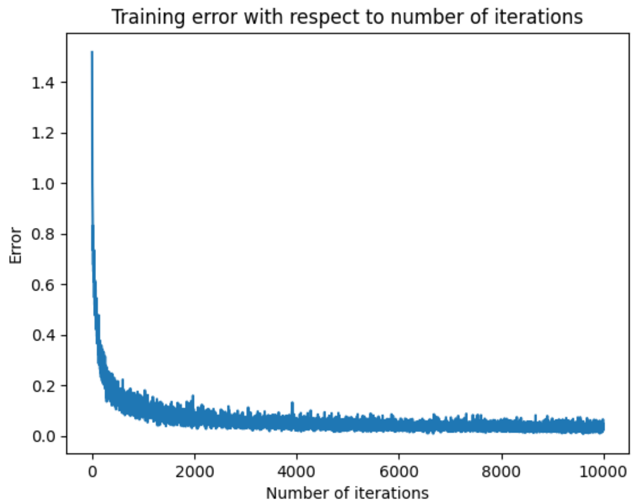
1.0.1. Results

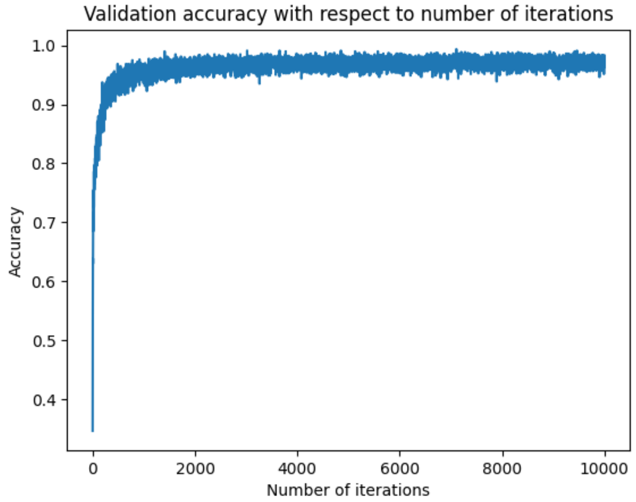
For FOMAML on Omniglot

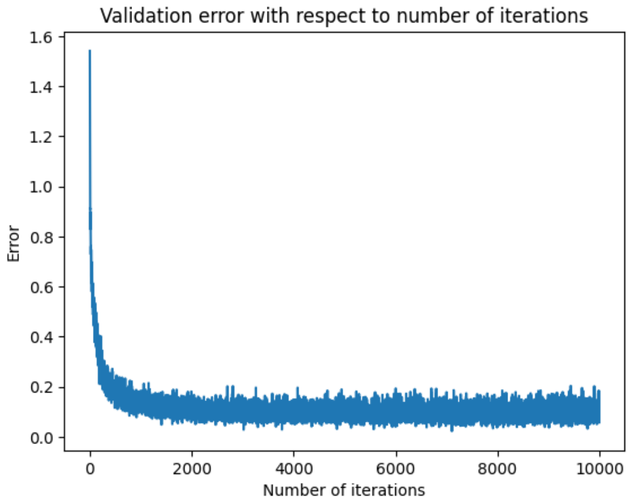




For MAML on Omniglot







1.0.1. Observations

Metrics	FOMAML	MAML
Avg. Test Accuracy	0.9699	0.9587
Avg. Test Error	0.10566	0.12392
Execution Time (s)	23617.31	24875.94

1.0.1. Conclusions

We can see FOMAML is producing higher accuracy for 10,000 epochs as due to the approximation it can converge much faster than MAML but if the number of epochs are increased the MAML is bound to have a much higher accuracy as it would lead to complete convergence.

We can also observe FOMAML to be faster by 5.05% when comparing it with MAML as the total number of Epochs are only 10,000 it doesn't have much significance but for systems with multiple and higher number of epochs it is very significant.

1.0.1. References

1. Arnold, Sebastien M. R., Praateek Mahajan, Debajyoti Datta, Ian Bunner, and Konstantinos Saitas Zarkias. 2020. “learn2learn: A Library for Meta-Learning Research.” arXiv [cs.LG]. http://arxiv.org/abs/2008.12284.

2. Giacomo Spigler. Meta-learnt priors slow down catastrophic forgetting in neural networks. arXiv preprint arXiv:1909.04170, 2019.

3. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks (Finn et al., ICML 2017)

4. An Interactive Introduction to Model-Agnostic Meta-Learning 👩‍🔬

5. https://www.ykilcher.com/