**Review of MAML : Model-Agnostic Meta-Learning for Fast Adaption of Deep Networks**

Lee hyo jeong 20215539

1. Report

The goal of meta-learning is to train a model on a variety of learning tasks, such that it can solve new learning tasks using only a small number of training samples. In MAML, the parameters of the model are explicitly trained such that a small number of gradient steps with a small amount of training data from a new task will produce good generalization performance on that task. And MAML is model-agnostic which means that it is compatible with any model trained with gradient descent.

From a feature learning standpoint, MAML can be viewed as building an internal representation that is broadly suitable for many tasks and from a dynamical systems standpoint, MAML’s learning process can be viewed as maximizing the sensitivity of the loss functions of new tasks with respect to the parameters: when the sensitivity is high, small local changes to the parameters can lead to large improvements in the task loss.

In the aspect of meta learning, It can be said that the training examples of MAML are the tasks T. In the Inner loop training phase, the model parameter theta is fine-tuned for the task T with a few labeled samples in support set S. In the outer loop training phase, model classify query Q with the fine-tuned model theta apostrophe. Optimize theta to theta-star with the classification loss for Q. Then pass the optimized theta-star to the next task. In short, the model is improved by considering how t he test error on new data from query changes with respect to the parameters. In effect, the test error on sampled tasks serves as the training error of the meta-learning process.

In the results, MAML showed comparable results with prior methods like matching networks, Siamese networks and memory models. Some of these existing methods are designed with few-shot classification in mind, and are not readily applicable to domains such as reinforcement learning. Despite noted benefits of MAML, I personally found it little complicate to implement the Episodic-Training pipeline so I’d prefer another scheme of meta learning like metric-based meta learning if there are no noticeable performance differences.

1. Experiments

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 5-way | | 5-way (extra layers) | | 5-way (Finetunes only classifier) | |
| MiniImagenet | 1-shot | 5-shot | 1-shot | 5-shot | 1-shot | 5-shot |
| MAML | 45.68 | 60.55 | 49.46 | 65.06 | 23.51 | 46.9 |

Codes are heavily borrowed from <https://github.com/dragen1860/MAML-Pytorch>

1. Analysis.

For the basic MAML, my implementation showed accuracy drop about 3% compared to original paper. Accuracy difference between 1-shot and 5-shot classification is about 15%. The architecture for MAML with extra layers is as follows – 6 sequence of modules(convolution, relu, batch normalization, max pooling). Accuracy increases about 5% both on 1-shot and 5-shot compared to original implementation of MAML. Model that only finetunes the classifier weights showed inferior results, debugging is needed to check whether the code implementation reflect the idea correctly.