Week 8 Paper Review

Niharika Shrivastava
School of Computing
National University of Singapore
Singapore, 119077
niharika@comp.nus.edu.sg

Abstract

This is a brief review of [1], [2].

1 Multi-Task Reinforcement Learning with Soft Modularization

The authors introduce a method called soft-modularization for multi-task learning that helps to generate soft combinations of all possible modules for different tasks instead of designing individual modules explicitly for each sub-policy. The implementation consists of a base network composed of multiple modules, and a routing network that takes a task embedding as input along with the current state and outputs a probability distribution over all the modules, thereby estimating a routing strategy. This enables to avoid naive sharing of parameters across tasks. Additionally, soft-modularization results in lesser variance than hard-modularization techniques for multi-task RL due to RL's temporal nature. Soft-modularization also enables automatic inference on what modules to share across different tasks. This is aided by the fact that both the base and routing networks are differentiable and hence can be trained together end-to-end. For cases where easier tasks converge faster than hard ones, weights of tasks with high confidence is reduced comparably. This avoids exploitation of these tasks in the long horizon.

1.1 Experiments

It is shown that the probability distribution of modules over similar tasks are closer. Moreover, for problems having small number of tasks, the performance improvement is comparable to other methods. However, convergence using soft-modularization is faster. However, for problems with a large number of tasks, this method significantly outperforms. Experiments also show that this method provides comparable performance to individual task training with the added advantage of better generalization over multiple tasks. Another advantage shown emprically is that even a small network architecture for soft-modularization performs better than strong baselines.

The authors also claim that their method to explicitly design modules for different tasks is better than "optimization relying on the gradient similarity is usually unstable, especially when there is a large gradient variance within each task itself". However, this claim is not shown to be supported in any way.

2 Gradient Surgery for Multi-Task Learning

The authors of this paper address the problem of inefficient optimization of multi-task learning problems due to conflicting gradients of different tasks. A conflicting gradient is identified as gradients having negative cosine similarity.

2.1 Implemetation

They identify three conditions, namely, conflicting gradients of different tasks, gradients coinciding with high positive curvature and gradients with a high difference in their magnitudes (leading to some task gradients dominating others), that if present together, are a cause of this detrimental performance. To this effect, the paper introduces Projecting Conflicting Gradient (PCGrad) that projects a task's gradient onto the normal plane of the gradient of other tasks that has a conflicting gradient, in order to aid efficient learning. The authors also show via ablation studies that conflicting gradients on their own are not detrimental and the correct solution can be achieved by simply averaging the different task gradients. Moreover, the authors prefer a large learning rate for faster optimization. However, this contributes to one of their 3 conditions of gradients having large differences in magnitude, hence, making PCGrad a good solution for this design choice.

In their formulation, only the conflicting gradients are projected normally to the gradient of other tasks while the rest are left untouched. Since the PCGrad procedure is only modifying the gradients of shared parameters in the optimization step, it is model-agnostic and can be applied to any architecture with shared parameters.

2.2 Experiments

The experiments show that PCGrad is indeed model-agnostic and can be combined with multiple supervised learning tasks and deep RL tasks to provide significant performance improvement. It also outperforms in terms of success rates and data efficiency. Ablation studies show that changing both the direction and maginitude of gradients is important for efficient learning.

References

[1] Yang, R., Xu, H., Wu, Y., & Wang, X. (2020). Multi-Task Reinforcement Learning with Soft Modularization. ArXiv. https://doi.org/10.48550/arXiv.2003.13661

[2] Yu, T., Kumar, S., Gupta, A., Levine, S., Hausman, K., & Finn, C. (2020). Gradient Surgery for Multi-Task Learning. ArXiv. https://doi.org/10.48550/arXiv.2001.06782