CS 330: Problem Set 2

Ji Won Park

October 16, 2020 (Due)

Colab notebook is available at `https://colab.research.google.com/drive/13ihNjVExjDOa24RupilHLy8SC8H361os?usp=sharing`.

1

1. See Colab.
2. See Colab.
3. Figure 3 shows the validation accuracy over iterations. Training was only done for 2,000 iterations, following TA comments on Piazza that training can be terminated once the benchmark performance (50%) on one of the learning rates is reached. After 2,000 iterations, the average test accuracy was 0.21, 0.52, and 0.57 for `inner_update_lr` of 0.04, 0.4, and 4.0, respectively. The inner-loop learning rate affects meta-training because the constant learning rate is used to optimize the task at hand. A fixed value of inner-loop learning rate that's too low (like 0.04) may not be able to get out of a local minimum efficiently. Here, the biggest learning rate of 4.0 led to fastest learning and highest final validation and test accuracy.

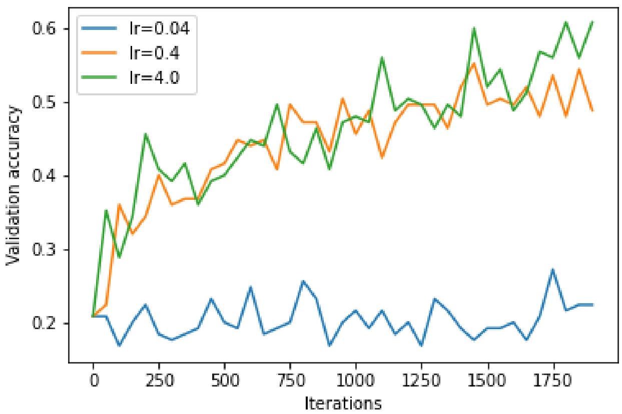

Figure 1: Meta-validation accuracy over meta-training iterations as the inner-loop learning rate is varied

4. See Colab for the implementation. Figure shows the validation accuracy over iterations. Again, training was only done for 2,000 iterations, following TA comments on Piazza that training can be terminated once the benchmark performance (50%) on one of the learning rates is reached. After 2,000 iterations, the average test accuracy was 0.51, 0.54, and 0.57 for `inner_update_lr` of 0.04, 0.4, and 4.0, respectively. The final validation and test accuracy was very similar when `inner_update_lr` was varied. Here, the learning rate is *learned* for different weights – so even when the initial learning rate is lower than optimal (0.04), the network eventually learns to grow the learning rate.

Compared to Prob 1.3 when the learning rate is fixed (Figure 3), the `inner_update_lr=0.04` case learns just as fast and just as well as the other cases.

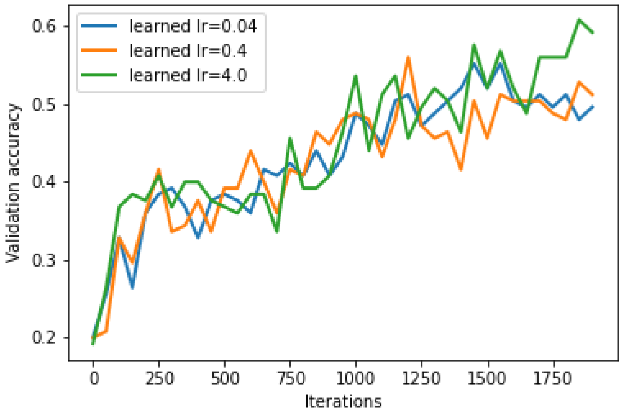

Figure 2: Meta-validation accuracy over meta-training iterations as the initial value of the optimized learning rate is varied

2

1. See Colab.

2. See Colab.

3. See Figure 3 for the validation accuracy over iterations. The average test accuracy with standard deviations was 0.85 ± 0.05.

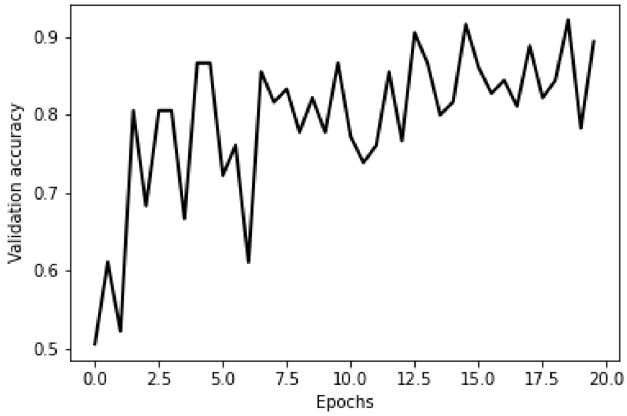

Figure 3: Meta-validation accuracy over meta-training iterations for ProtoNet

3

Figure shows the mean and standard deviation of meta-test accuracy of MAML and ProtoNet for meta-test values of $K = 4, 6, 8, 10$. Again, MAML was only trained for 2,000 iterations, following TA comments on Piazza that training can be terminated once the benchmark performance (50%) is reached. For MAML, I removed the division of shot number in the loss function following the TA comment on Piazza post @418 that the division should be removed as it leads to smaller-norm gradients (effectively, a smaller learning rate). Following the removal, the test accuracy seems to increase with increasing K for MAML, though the increase is not significant. This agrees with what students found in Piazza post @418. ProtoNet generally does well in few-shot classification because the model is natively defined by the concept of similarity to class prototypes.

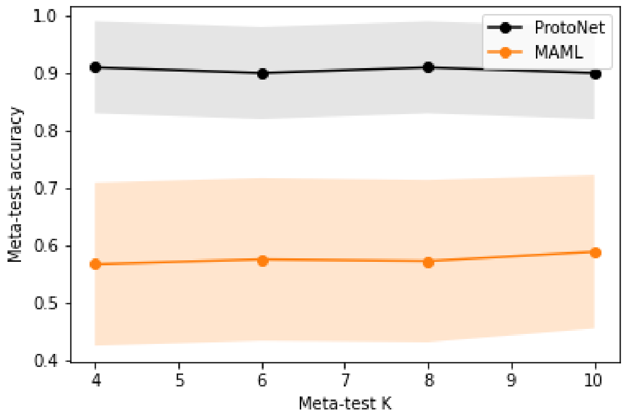

Figure 4: Mean and standard deviation of final meta-test accuracy for ProtoNet and MAML