Transferring Knowledge across Learning Processes

Abraham Jose
ID :5068109, CAP6614
abraham@knights.ucf.edu

1 Summary

The authors propose Leap, a framework that achieves transfer learning at a higher level abstraction by transferring knowledge across learning processes. They associated each task with a manifold on which the training process travels from initialization to final parameters and construct a meta-learning objective that minimizes the expected length of this path. ie, the framework is learned during the training process on the fly at negligible cost. It develops the principled approach for short horizontal bias problem in meta-learning.

2 Strengths of the proposal

1. It is a novel idea that is relatively well written. Also, they rely on best initialization from which the model converges asap. Leap in-fact finds the best initialization and learning gradient for transferring the learning.
2. The model has less number of parameters compared to state-of-the-art techniques to address the 'far transfer learning' problem or the transfer of high level abstraction.
3. Leap framework outperforms competing methods, both in meta-learning and transfer learning tasks which is tested in Multi-CV task and Omniglot data-set.

3 Weaknesses of the proposal

1. The experimental setting for Leap is somewhat non-standard. Also, the ability to reproduce the results are questionable since the parameter configuration is missing.
2. There exists several gaps between the proposed theory and algorithm in this far transfer learning technique called Leap. By flipping the sign, for an increasing change from Ps^i to Ps^{i+1} you can't expect the to trick the meta-learner that it is decreasing because it is a optimization trajectory.

4 Results

Transfer learning techniques typically ignores the learning process itself by restricting knowledge transfer to scenarios where the target and source task are similar. However, Leaps brings this out-of-the-box idea that the transfer learning of higher level abstraction can be done by minimizing the expected length of gradient path. Thus by learning how the source network reached the accuracy, we can make target network learn how it should learn. The proposed framework has superior generalizing properties to fine-tuning and competing meta-learners like HAT and Progressive Net.

5 Discussion

How does Leaps works for an entirely different architectures for source and target network? The learning parameters and the gradient back-propagation(for example ResNet has easy path to back-propagate error) will be different and so the nature of learning in both network.

CAP 6614: Current Topics in Machine Learning (Fall 2019), UCF.