



Lifelong learning with deep neural networks

Lawrence Carin

Duke University

lcarin@duke.edu

Abstract:

Neural networks trained via conventional techniques tend to experience catastrophic forgetting in sequential task settings (e.g., sequential sensing missions of interest to the DOE), where data from previous tasks are unavailable. A number of methods, using various model expansion strategies, have been proposed recently as possible solutions. However, determining how much to expand the model is left to the practitioner, and often a constant schedule is chosen for simplicity, regardless of how complex the incoming task is. Instead, we propose a principled Bayesian nonparametric approach based on the Indian Buffet Process (IBP) prior, letting the data determine how much to expand the model complexity. We pair this with a factorization of the neural network's weight matrices. Such an approach allows us to scale the number of factors of each weight matrix to the complexity of the task, while the IBP prior encourages sparse weight factor selection and factor reuse, promoting positive knowledge transfer between tasks. We demonstrate the effectiveness of our method on a number of continual learning benchmarks and analyze how weight factors are allocated and reused throughout the training.