I will first highlight the role of metrics in the study of dynamical systems, for example in the context of evolution equations and random iterations. Nonexpansive maps frequently arise, sometimes at the outset and at other times in more hidden ways. From this perspective, I will then discuss a novel approach to training neural networks or optimization problems for which stochastic gradient descent (SGD) is employed. In recent joint work with Avelin (Uppsala), Dherin, Gonzalvo, Mazzawi, and Munn (Google Research), we introduce and investigate backward-SGD—a reverse training method that, in experiments such as training ResNet-50 on CIFAR-100, consistently demonstrates greater stability and faster convergence compared to standard SGD in the constant-learning-rate and small-batch-size regime. While computational challenges remain, we think that this approach presents promising opportunities for improving the efficiency of optimization algorithms.