ANASTASIS KRATSIOS, ETH Zurich

Universal Probability Measure-Valued Deep Neural Networks

We introduce deep neural architecture types with inputs from a separable and locally-compact metric space X and outputs in the Wasserstein- 1 space over a separable metric space Y. We establish the density of our architecture type in C(X; P1(Y)), quantitatively. NB that our results are new even in the case where X and Y are Euclidean, in which case, we find that many commonly used types such as MDNs and MGANs are universal special cases of our model type. We show that our models approximate functions in C(X; P1(Y)) by implementing e-metric projections in the Wasserstein- metric onto the hull of certain finite families of measures therein. If the target function can be represented as a mixture of finitely many functions, each taking values in a finite-dimensional topological submanifold of the Wasserstein space, we find that the approximating networks can be assumed to have bounded width. As applications of our results, we address the following problems. We show that, under mild conditions, our architecture can approximate any regular conditional distribution of an X-valued random element X depending on a Y -valued random element Y with arbitrarily high probability. Consequentially, we show that once our approximation of this regular conditional distribution is learned, any conditional expectation of the form E[f(X; Y)|Y = y] for Caratheodory f with uniformly-Lipschitz first component and a uniformly-bounded second component, is approximable by standard Monte-Carlo sampling against the learned measure. We illustrate our theory in the context of stochastic processes.