The geometry of neural networks : a Riemannian foliation perspective on robustness.

Abstract : In this presentation we employ the tools of geometry and statistics to shed light on the relationship between data and neural network predictions. In particular, we take inspiration from the field of information geometry combines precisely these two approaches. We see the neural network's output as the parameter of a probability distribution. By using the Data Information Matrix (DIM), a variation of the Fisher Information Matrix (FIM), we investigate the network's input/output relationship and reveal its understanding of the data structure. This statistical framework yields a (degenerate) Riemannian metric that we use to analyze the geometry of the data. In particular, we lean on a foliation arising from the kernel of the DIM to conduct our study of the low dimensional data in the high dimensional input space. Unfortunately, in most practical cases of machine learning, the DIM has a non constant rank and is non smooth, making it difficult to yield a well defined foliation. To tackle this issue, we prove that for usual neural network architectures, this only happens in a nowhere dense set. Besides, we investigate these singularities as they may teach us about distances between datasets and efficiency of knowledge transfer. Finally, we apply this new geometrical framework given by the DIM to the analysis of the robustness of neural networks. We show that the curvature of the transverse to the kernel leaves can be utilized to improve adversarial attacks, indicating that the geometry of the data is key in the robustness of machine learning algorithms.