universität innsbruck
Department of Mathematics

Mathematics Colloquium

Sjoerd Dirksen
Utrecht University

The separation capacity of random neural networks

Neural networks with random weights appear in a variety of machine learning applications, most prominently as the initialization of many deep learning algorithms and as a computationally cheap alternative to fully learned neural networks. The first goal of this talk is to enhance the theoretical understanding of random neural networks by showing that these objects can beneficially transform data for a binary classification task. Specifically, I will show that a sufficiently large two-layer ReLU-network with Gaussian weights and uniformly distributed biases makes two classes (with positive distance) linearly separable with high probability. Building on the insights behind this result, I will next present a simple randomized algorithm to produce a small interpolating neural net for a given dataset with two classes, which goes beyond worst case memorization capacity bounds. In both results, the size of the network is explicitly linked to geometric properties of the two classes and their mutual arrangement, which quantify the "intrinsic dimension" of the problem. This instance-specific viewpoint allows to overcome the curse of dimensionality. No prior knowledge of neural networks is assumed for this talk.

The talk is based on joint works with P. Finke (Utrecht University), M. Genzel (Utrecht University, Merantix Momentum), L. Jacques (UCLouvain), and A. Stollenwerk (UCLouvain, KPMG).

Jan 31, 2024
14:15 – 15:15

SR 1, ICT-Gebäude,
Technikerstraße 21a