PhD Colloquium

Vienna School of Mathematics

Daniel Haider

The injectivity of ReLU layers and other non-linear measurements

Abstract: Non-linearities frequently arise in applications, either due to technical constraints or as intentional elements of model design. A prominent instance of the latter is to use the composition of affine linear mappings and non-linear activation functions as layers of artificial neural networks. Among many layer designs, ReLU layers - i.e., layers using $\mathrm{ReLU}(t) = \max(0, t)$ as activation - are the most widely used layer types due to their simplicity and effectiveness. By performing hard thresholding, the ReLU function naturally acts as a sparsifier, where a black-box machinery determines which and how much information of the input is suppressed. Assessing whether the original input can be reconstructed from the output is therefore crucial for improving the interpretability and functionality of the associated models.

In this talk, we present a frame theoretic perspective to approach the injectivity of ReLU layers and show how it is linked to other situations where non-linearities occur. To check injectivity in practice, we derive an injectivity characterization via the bias vector of the ReLU layer, and to do the reconstruction, we modify the classic frame algorithm.

This is joint work with M. Ehler, D. Freeman, H. Eckert, and P. Balazs.

January 13 2024, 15:30-16:30
Seminarraum 13, Universität Wien
(Oskar-Morgenstern-Platz 1)