Einladung

Würzburger Mathematisches Kolloquium

Julius-Maximilians-Universität Würzburg • Institut für Mathematik

Leon Bungert

Julius-Maximilians-Universität Würzburg - Antrittsvorlesung

The Mathematics of Adversarial Machine Learning

Dienstag, 16. Januar 2024 • 14:15 Uhr

Seminarraum SE41 • Forschungsbau (Emil-Fischer-Straße 41, 97074 Würzburg)

Der Vortrag wird auch Zoom-Meeting übertragen: go.uniwue.de/ifmcolloquium-zoom

Abstract. It is well-known that, despite their aptness for complicated tasks like image classification, modern neural networks are prone to insusceptible input perturbations (a.k.a. adversarial attacks) which can lead to severe misclassifications. In this talk I will speak about adversarial training, a robust optimization method designed for the training of adversarially robust classifiers. I will show that the method is equivalent to a nonlocal geometric regularization problem involving the decision boundary of classifiers. This allows for the application of tools from calculus of variations and geometric measure theory to study existence, regularity, and asymptotic behavior of minimizers. Furthermore, I will show Gamma-convergence of this perimeter to a local anisotropic perimeter as the strength of the adversary tends to zero, thereby establishing an asymptotic regularization effect of adversarial training.

Alle sind herzlich eingeladen.

Die Dozentinnen und Dozenten der Mathematik