Universität Stuttgart

Institut für Aerodynamik und Gasdynamik
Wankelstraße 3, 70563 Stuttgart

NRG

Benchmarking Methods to Compute Lipschitz Bounds for Neural Networks

Master Thesis/ HiWi Position (m/w/d)

Background

Neural networks (NNs) are playing an increasingly important role in all our lives. They are also being used more and more frequently in technical systems. For such AI-based systems to be certifiable, it is crucial that they are verifiably robust against disturbances. Even the smallest changes in the input variables, possibly imperceptible to humans, should not cause the model's response to reverse. The Lipschitz constant (LC) of NNs is one metric to describe how sensitive they are to input disturbances. While its exact determination is generally too expensive, estimates of the LC with varying degrees of accuracy do exist. It has been shown that the accuracy of these estimates can vary considerably depending on the size of the NNs under investigation. Yet, for certifiability, a reliable estimate with known uncertainty is essential.

Project Description

In this thesis, various estimation methods are supposed to be implemented and benchmarked. This shall yield a selection rationale for which method to use in order to efficiently compute tight LC bounds in a particular verification scenario. To this end, the work is structured as follows:

- Literature review
- Merge existing implementations of estimation methods into a common Python library with a standardized interface
- Benchmark methods with respect to tightness of the bounds, computational efficiency, and flexibility to handle different NN architectures
- Development a guideline to choose an appropriate method based on the requirements
- Documentation of the results

Your Profile

- Solid Python programming skills
- Basic understanding of feed-forward and preferably convolutional NNs

Application

If you are interested, even if you don't fulfill all the listed requirements, please send your application (incl. CV and latest transcript) via e-mail.

Start: as soon as possible

Contact Person:
Hannes Mandler
Institut für Aerodynamik und Gasdynamik
E-Mail: hannes.mandler@iag.uni-stuttgart.de
Tel.: +49 711 685 63461