FAKULTÄT FÜR MATHEMATIK
Dekan Univ.-Prof. Dr. Radu Ioan Boţ

Einladung zur öffentlichen Defensio

Beryl Ramadhian ARIBOWO

Thema der Dissertation

Machine learning for chemical data

Abstract:

The most important chemical property of molecules is the equilibrium energy and that of their molecular configurations. This thesis deals with the prediction of these energies from given data using statistical and machine learning approaches.

The first part of the thesis deals with the construction of highly accurate potential energy surfaces for diatomic molecules using traditional functional approaches.

The second part deals with more complex potential energy functions for small molecules H_xO_y and their training with reverse mode automatic differentiation, which offers better efficiency compared to finite differences or symbolic differentiation. In addition, a novel semi-supervised learning technique based on Shepard interpolation is presented, which is useful for training chemical data with minimal information on the target.

The third part of the thesis discusses the QM9 challenge, the problem of predicting all 130K equilibrium energies of the QM9 dataset with chemical accuracy from the equilibrium energies for a tiny amount (less than 1%) of training geometries. By combining delta machine learning, hyperparameter tuning, and data selection techniques involving combinatorial optimization, an accuracy of 3.25 kcal/mol) is achieved with 100 training geometries. Using the training parameters found and a less expensive data selection strategy, chemical accuracy (1 kcal/mol) was achieved with 3750 training geometries.

Prüfungssenat

Univ.-Prof. Mag. Dr. Andreas Cap
(Vorsitz, Universität Wien)

emer. o. Univ.-Prof. Dr. Arnold Neumaier
(Universität Wien)

Prof. Dr. Nikolaos V. Sahinidis
(Georgia Institute of Technology)

Prof. Dr. Helmut Harbrecht
(Universität Basel)

Zeit und Ort

Mittwoch, 13. November 2024, 16:00 Uhr

Online:

> https://univienna.zoom.us/j/64014640223?pwd=4iXbBdoV4mvGKsUaWKtWkKy8czW
> NEF.1
>
> Meeting ID: 640 1464 0223
>
> Kenncode: 587067