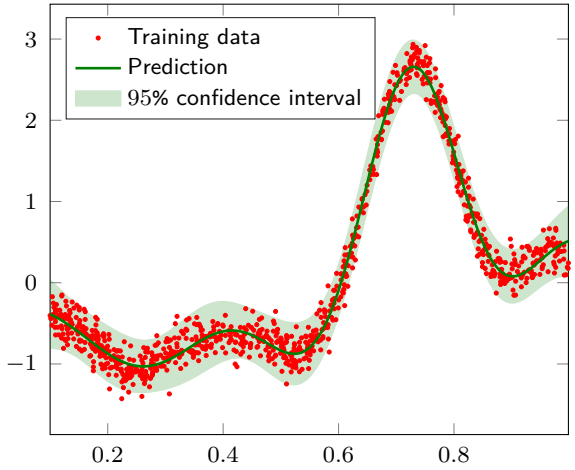
Gaussian Process Regression for Big Data

TECHNISCHE UNIVERSITÄT DARMSTADT

Master's Thesis

A Gaussian process (GP) is a supervised machine learning method that naturally provides an uncertainty measure on the prediction. Its objective is to approximate an unknown input-output mapping using observed data. To obtain a good mapping, the optimization of hyperparameters is crucial. The resulting posterior inference requires the inversion of a covariance matrix. The computational complexity of this inference increases cubically with the number of observed data points.

Hence, it is impractical to calculate the inverse directly for large data sets. There exists a variety of methods to deal with these large data sets.[1]



Your tasks will be:

1. Literature review on scalable Gaussian processes
2. Implementation of scalable algorithms for Gaussian processes in HILO-MPC[2]
3. Performance comparison, e.g., with neural networks

Experience with / knowledge about:	Gaussian process regression, big data, scalability
Programming skills:	Python (good to very good skills required)
Language:	English (thesis and presentation), German

Johannes Pohlodek

Room: S3|10 510
E-mail: johannes.pohlodek@iat.tu-darmstadt.de
Web: https://www.ccps.tu-darmstadt.de

1 https://doi.org/10.1109/TNNLS.2019.2957109
2 https://www.ccps.tu-darmstadt.de/research_ccps/hilo_mpc/index.en.jsp

CCPS Control and Cyber-Physical Systems