

Internship: Probabilistic modeling in chemical plant sensor monitoring

Description

Sensor monitoring is essential in chemical plants because process conditions such as temperature, pressure, and flow rate must remain within safe operating limits. In addition, sensor networks can provide early warning of deviations in product quality. Model-based anomaly and outlier detection can reduce the risk of failing market specifications.

In this internship, we will develop hybrid models that combine domain knowledge of the chemical plants with neural networks that fill gaps in the physical description. We will work with Bayesian methods and actual industrial data reflecting the process conditions measured over time in chemical plants operating at our NRW sites. The goal is a systematic comparison to identify the most expressive model.

To achieve this goal, we propose the following steps as part of this internship

- Development of approaches that combine pooling (complete, none, partial) with both super- and unsupervised learning
- Implementation of these models in numpyro and jax
- Bayesian hyperparameter search with respect to model architecture and solver parameters to control the trade-off between solver convergence and model expressiveness

Literature

[1] <https://arxiv.org/html/2412.20512v1>

[2] [\[2107.03342\] A Survey of Uncertainty in Deep Neural Networks \(arxiv.org\)](#)

Eligibility

- Enrollment at TU Dortmund University required
- Internship for Module “Project Work”; this part of the module should not have been completed yet

Must-haves

- Strong programming skills in python, incl. pandas, numpy, any of pytorch or jax, and any of matplotlib, plotly dash, or seaborn
- Background in hierarchical (mixed-effects) models and feed-forward MLPs
- Interest in Bayesian models
- Team mindset and creative modeling mindset

Nice-to-haves

- Knowledge of pyro, numpyro, ipython notebooks
- Knowledge of RNN, GRU, and LSTM
- Knowledge of HMC / NUTS, Gelman’s convergence criteria
- Background in hybrid modeling
- Experience with Git, Linux

Contacts

TU Dortmund University: Prof. Dr. Katja Ickstadt, Department of Statistics, ickstadt@statistik.tu-dortmund.de

Bayer AG: Dr. Moll Glass, Applied Mathematics, moll.glass@bayer.com

Applicants are asked to enclose their CV, transcript of records, preferred start and end date and preferred workload; earliest start date is January 2027. Please do not use Moll Glass' e-mail address for any other purposes than applying for this internship.

Miscellaneous

We recommend a duration of at least 3 months at 80% (4 days internship, 1 day off). This is a paid internship. Bayer will also provide the necessary hardware and a shared office. Working from home is possible.