Engineering Risk Analysis Group
Prof. Dr Daniel Straub

MSc thesis proposal

Proper scoring rules for probabilistic forecasting

Motivation and Objectives

Engineers and researchers utilize uncertainty quantification (UQ) to enrich classical deterministic model predictions by probabilistic statements. Domains such as weather forecasting and machine learning have a long tradition of training and also validating these probabilistic predictions based on data [1].

The key question, therefore, is: what constitutes a good probabilistic prediction? Ideally, it is both calibrated as well as sharp. Well-calibrated means that the model's predicted range of future values coincides with the observed frequencies in the long run. On the other hand, a sharp prediction is characterized by low prediction uncertainty, i.e., the predicted range for the future values is narrow. To balance calibration and sharpness during training, one should use a proper scoring rule [2], e.g., the cross-entropy or log-loss.

The topic of this work is the exploration of proper scoring rules for probabilistic forecasting with Deep Learning. The work could span the development of new scoring rules, and comparison with commonly used ones on simple toy problems or real-life datasets. This thesis offers a combination of method development with subsequent numerical implementations.

This is an open thesis, so the objectives, as well as the workflow, can be adjusted to the student's ideas, progress, results, etc.

Methodology

The suggested workflow is as follows:

- Getting familiar with probabilistic forecasting and the established methods
- Understanding the nuances of proper scoring rules
- Development of own scoring rule
- Implementation of the proposed solutions
- Benchmarking with commonly used scoring rules on toy problems or real-life datasets.

Requirements

We expect that the student has sufficient background / experience in:

- Probability theory and statistics, e.g., via completion of the course "Risk Analysis"
- Deep Learning, e.g., via completion of the course "Introduction to Deep Learning"

Starting date: Flexible, as soon as possible

Supervision

If you are interested in this topic, please contact Daniel Koutas (daniel.koutas@tum.de, ERA group, TUM) with your transcript of records and preferred approximate starting date.

References

[1] Gneiting, T., Balabdaoui, F., Raftery, A.E., 2007. Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society. Series B: Statistical Methodology 69, 243–268. https://doi.org/10.1111/j.1467-9868.2007.00587.x

[2] Murphy, K. P. (2022). *Probabilistic machine learning: an introduction*. MIT press.