

## Thesis topic (B.Sc. or M.Sc.)

### Lin's concordance coefficient and predictions of non-linear regression models

Concordance coefficients like Cohen's kappa or Fleiss' kappa are calculated when nominal or ordinal ratings from several judges is available and a measure of the judges' agreement is needed. Coefficients for concordance of continuous ratings (Lin, 1989; Lin et al., 2002), are less well-known, but are equally useful.

Recently, Bottai et al. (online first) discuss what changes in linear regression if instead of optimizing the least squares criterion, the concordance coefficient of observed dependent variables and fitted values is maximized. This is an interesting, non-standard application of Lin's concordance coefficient, as there are no raters, but the dependent variable and the fitted values are treated as raters. It turns out that the resulting fitted values do not coincide with the least squares fit, but have larger variance.

Based on this observation, this thesis explores regression models beyond those studied by Bottai et al., including the class of generalized linear models. Apart from an adaptation of the existing approach, the candidate could base the investigation on a generalization of Lin's coefficient for models with a threshold form. Another direction of this thesis could be to study predictions of tree-based machine learning models like the random forest or gradient boosting.

This thesis will be supervised by Prof. Dr. Philipp Doebler.

**Contact:** [doebler@statistik.tu-dortmund](mailto:doebler@statistik.tu-dortmund)

**Date:** 26.5.2022

### References

Matteo Bottai, Taeho Kim, Benjamin Lieberman, George Luta & Edsel Peña (online first) On Optimal Correlation-Based Prediction, *The American Statistician*, DOI: [10.1080/00031305.2022.2051604](https://doi.org/10.1080/00031305.2022.2051604)

Lin, L. (1989), A Concordance Correlation Coefficient to Evaluate Reproducibility, *Biometrics*, 45, 255–268.

Lin, L., Hedayat, A., Sinha, B., and Yang, M. (2002) Statistical Methods in Assessing Agreement: Models, Issues, and Tools, *Journal of the American Statistical Association*, 97, 257–270.