



Contribution ID: 78

Type: **not specified**

Plots for XAI: FANOVA Graph and FaithShapGraph

Explainable AI (XAI) approaches, most notably Shapley values, have become increasingly popular because they reveal how individual features contribute to a model's predictions. At the same time, global sensitivity analysis (GSA) techniques, especially Sobol indices, have long been used to quantify how uncertainty in each input (and combinations of inputs) propagates to uncertainty in the model's output. Prior work (e.g., Owen 2014) highlighted the theoretical connections between Shapley-based explanations and Sobol-based sensitivity measures.

FANOVA (functional ANOVA) graphs were created to visualize main-effect Sobol indices and total interaction terms in a straightforward graphic form, making it easy to see which inputs (and pairs of inputs) drive model behavior. In this work, we apply the same general concept to Shapley values and the recently introduced Shapley interaction indices. By translating complex machine-learning models into analogous "Shapley graphs", we provide equally intuitive visual representations of both individual feature contributions and feature interactions. Through several real-world case studies, we show that these Shapley-based graphs are just as clear and user-friendly as FANOVA graphs, and we discuss how the two methods compare in terms of interpretability.

References

- Owen, A. B. (2014). Sobol' indices and Shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2, 245–251.
- Fruth, J., Roustant, O. & Kuhnt, S. (2014). Sensitivity Analysis and FANOVA Graphs for Computer Experiments. *Journal of Statistical Planning and Inference*, 147, 212–223
- Muschalik, M., Baniecki, H., Fumagalli, F., Kolpaczki, P., Hammer, B., & Hüllermeier, E. (2024). SHAP-IQ: Shapley Interactions for Machine Learning. In *Proceedings of the Thirty-Eighth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*

Special/ Invited session

Classification

Both methodology and application

Keywords

XAI, Shapley values, Sobol Indices

Primary authors: KUHLMANN DE CANAVIRI, Lara (Fachhochschule Dortmund); KUHNT, Sonja (Dortmund University of Applied Sciences and Arts)

Presenter: KUHLMANN DE CANAVIRI, Lara (Fachhochschule Dortmund)

Track Classification: AI: Interpretability and Trustworthiness