

Transparency of Machine Learning Models in Credit Scoring

A major requirement for Credit Scoring models is of course to provide a risk prediction that is as accurate as possible. In addition, regulators demand these models to be transparent and auditable. Thus, in Credit Scoring very simple Predictive Models such as Logistic Regression or Decision Trees are still widely used and the superior predictive power of modern Machine Learning algorithms (cf. e.g. Lessmann et al., 2016; Bischl, Kühn and Szepannek, 2016) cannot be fully leveraged. A lot of potential is therefore missed, leading to higher reserves or more credit defaults (cf. Szepannek, 2017).

This paper presents an overview of techniques that are able to make "black box" machine learning models transparent and demonstrate how they can be applied in Credit Scoring. We use the DALEX set of tools (cf. Biecek, 2018) to compare a traditional scoring approach with state of the art Machine Learning models and asses both approaches in terms of interpretability and predictive power. Results show that a comparable degree of interpretability can be achieved while machine learning techniques keep their ability to improve predictive power.

References

- Bischl, B., Kühn, T. and Szepannek, G. (2016): On Class Imbalance Correction for Classification Algorithms in Credit Scoring, In: Löbbecke, M., Koster, A., Letmathe, P., Madlenerm R., Peis, B. und Walther, G. (Eds): Operations Research Proceedings 2014, 37-43, Springer, DOI: 10.1007/978-3-319-28697-6_6.
- Lessmann, S., Baesens, B., Seov, H. and Thomas, L. (2015): Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research, EJOR 247(1), 124-136.
- Szepannek, G. (2017): On the Practical Relevance of Modern Machine Learning Algorithms for Credit Scoring Applications, In: Mucha, H.: Big Data Clustering: Data Preprocessing, Variable Selection and Dimension Reduction, WIAS Report 29, S. 88-96, DOI: 10.20347/WIAS.REPORT.29.
- Biecek P (2018). "DALEX: Explainers for Complex Predictive Models in R." Journal of Machine Learning Research, 19(84), 1-5. <http://jmlr.org/papers/v19/18-416.html>.