Advances in Reliability, Safety and Security

ESREL 2024
Collection of Extended Abstracts

Importance Reliability Measures For Ensemble Models In Machine Learning

Jose Divasón, Francisco Javier Martinez-de-Pison, Laura Moreno-Resa, Idriss T. Sandoval, Eduardo Sáenz-de-Cabezón

Universidad de La Rioja, Logroño, Spain

Keywords: ensemble models, machine learning, artificial intelligence, importance measures, structural importance, series-parallel

Ensemble models in machine learning (Zhang and Ma, 2012) represent a powerful paradigm where the combination of diverse models leads to enhanced predictive performance and robustness. The fundamental idea is to leverage the collective extracted knowledge of multiple models to achieve more accurate and stable predictions than any individual model. This approach is particularly valuable in addressing complex problems and mitigating overfitting. Model diversity is a key principle in ensemble learning, ensuring that the constituent models capture different aspects of the data or employ distinct learning strategies. Popular ensemble techniques include bagging, boosting, and stacking. Bagging (Breiman, 1996), exemplified by Random Forests, involves training multiple instances of the same model on different subsets of the training data and aggregating their predictions. Boosting (Shapire, 2003), as seen in algorithms like AdaBoost and Gradient Boosting Machines, focuses on sequentially improving the performance of weak learners. Stacking (Pavlyshenko, 2018) combines predictions from diverse models using a meta-model, exploiting the complementary strengths of each base model.

In this contribution we use notions from classical system reliability theory (Trivedi and Bobbio, 2017) to study ensemble models in Machine Learning, and propose ensemble structures to improve the models' performance. In particular, we make use of importance measures (Kuo and Zhu, 2012) including structural importance and Birnbaum importance among others.

We consider model ensembles for classification problems as systems such that their components (i.e. the individual models) are arranged in a given structure (e.g. consensus ensembles are seen as series systems). The reliability of the so considered system gives us the probability of correct classification of a sample, given that it contains the feature of interest, i.e. it is a *true positive*, the unreliability of the system gives us the probability of a *false negative*. On the other hand, the reliability of the dual system gives us the probability of a *true negative*, and the unreliability of the dual system gives the probability of *false positives*. Using classical techniques on system design, based on importance measures, we can design ensembles that optimize a function on the balance of true positives/negatives, according to the problem needs.

Some examples of application of these techniques are given, in particular focusing on series-parallel systems/ensembles and k-out-of-n systems/ensembles.

Acknowledgements

This work is partially supported by project Inicia 2023/02 funded by La Rioja Government (Spain).

References

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Kuo, W., Zhu, X., 2012. Importance Measures in Reliability, Risk, and Optimization: Principles and Applications, Wiley and Sons, Chichester, UK.

Pavlyshenko, B. 2018. Using Stacking Approaches for Machine Learning Models, IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, pp. 255-258.

Schapire, R.E., 2003. The Boosting Approach to Machine Learning: An Overview. Denison, D.D., Hansen, M.H., Holmes, C.C., Mallick, B., Yu, B. (Ed.), Nonlinear Estimation and Classification. Lecture Notes in Statistics, vol 171. Springer, New York, 149-171.

Trivedi, K. S., Bobbio, A. 2017. Reliability and Availability Engineering: Modeling, Analysis, and Applications. Cambridge University Press, Cambridge.

Zhang, C., Ma, Y., 2012. Ensemble Machine Learning. Methods and Applications. Springer Verlag, Neew York.