Expose - Explainability Machine Learning - Visualization of Random Forests

Fabio Rougier

April 9, 2022

Contents

1 Introduction

Random Forests (RF) are a powerful ensemble method with a low barrier of entry. Because of their ease of use and performance they are used in many applications. However, they fall short when it comes to transparency. RFs yield a multitude of valuable insights into their decision making and the data they process. Visualizing this is usually a challenge because of the inherent scale that of a RF. In this work we want to look at the approaches trying to overcome this limitation.

2 Visualizing Decision Trees

An obvious approach to visualizing a RF is to inspect the decision trees, constituting the RF. One example for the visualization of a decision tree is *BaobabView* (Van Den Elzen & Van Wijk, 2011). As many other approaches visualizing decision trees, it utilizes Node-Link Diagrams (NLDs) as its' main visualization. A confusion matrix yields additional insights into the relations of the underlying features. However this visualization does fall short when trees grow too large, as it becomes hard to inspect every individual node and branch of the tree. This is one of the problems that *TaxonTree* tries to overcome (Parr, Lee, Campbell, & Bederson, 2003). It uses a tree visualization approach that is scalable for large trees by adding the possibility to zoom, browse and search the tree. While this does deal help if a tree grows too large, it does not provide any insight on the general structure of the tree as a whole.

Generalizing either of the approaches towards RFs is also non trivial. Both simply lack the scalability in the desired dimension. It also leads into a dangerous of focusing too much on the structure of individual trees. RFs - as all ensemble methods - arrive at their decisions by combining the decisions of the individual trees. Inspecting and even understanding individual trees, will only yield limited insights over the RF.

3 Visualizing Random Forests

To fully understand the structure and decision making of a RF, many aspects of the RF have to be conveyed by the visualization. First of all the typical indicators like a *mean squared error* or a *mean average error* of any machine Learning model give an indication of the overall performance. Adiitionaly there are some metrics specific to RFs that should be included to get a deeper understanding of the specific instance of the RF, like the *mean impurity* of the individual trees or the *out of bag error*. With its' unique way of giving a distance measure for features, a confusion matrix can also yield valuable insights to the relations of the features and how the RF interprets them. As stated before, the visualization of RFs heavily relies on how it overcomes the scalability issues of RFs.

3.1 iForest

The *iForest* visualization is one of the most promising visualization approaches for RFs. (Zhao, Wu, Lee, & Cui, 2018) It focuses on the interpretability of the RF and raises the concern that many domains currently would not even consider using RFs because of their lack thereof. The authors also approach the understanding of RFs by coming from two different angles: *Feature Analysis* and *Case Based Reasoning*. Both require very different charts and give the user valuable insight into the dataset and the behaviour of the RF. The elaborate, dashboard-like web application utilizes the benefits of small multiples and has interconnected and interactive charts, each devoted to offer a specific insight. One especially unique part of the dashboard is the use of *Partial Dependence Plots* to display the RF's classification behaviour in regard to each feature. This is supported by a bar chart of the feature's distribution in order to support this view with more context. While the dashboard is a very powerful tool, it can be very overwhelming for users. It also requires some in-depth knowledge about RFs in order to come to conclusions or a required action instruction. This tool is clearly aimed at supporting data scientists to understand their own creations and not suited for a domain expert.

3.2 ExMatrix

The *ExMatrix* follows an out-of-the-box-thinking approach by breaking up machine learning models into a set of rules (Neto & Paulovich, 2020) (Ming, Qu, & Bertini, 2018). It is therefore very powerful because it is not limited to be used with RFs only, but can be applied to almost any kind of machine learning model. It is best suited to abstract ensemble models and simplify them. This is however not to be mistaken with creating surrogates from the given models, as it represents the underlying models exactly. By breaking down a RF in said rules, it is possible to represent the entire RF in a matrix structure, by represeting columns as features and rows as rules, resulting in cells as *rules predicates*. This entire rethinking of the RF as a whole allows for unique graphs and insights in the RFs' decision making. Thinking of the RF as a set of comparable rules and evaluating those, yields a very deep understanding, both on a case-based sample level, but also on a global level. While not supported in the original paper, this would also allow for intricate comparisons of models on the same data set. The biggest issue with this approach is, that while it does yield powerful insights, it can be very difficult to convey these insights to a domain expert. While a data scientist can be expected to wrap their head around the idea of disassembling a RF into a set of rules, this idea is not intuitive for a domain expert who might already have trouble understanding how the RF works in the first place.

4 Summary

While the main challenge of gaining access to the insights and details of the inner workings of a RF might already have been solved they seem to be hidden behind a kind of complexity layer. Powerful approaches for RF visualization exist, but there is a need to make them more accessible for domain experts. There is a lot of value to be derived from the existing work and they will surely be an influence to possible approaches aimed at domain experts.

References

Ming, Y., Qu, H., & Bertini, E. (2018). Rulematrix: Visualizing and understanding classifiers with rules. *IEEE transactions on visualization and computer graphics*, *25*(1), 342–352.

Neto, M. P., & Paulovich, F. V. (2020). Explainable matrix-visualization for global and local interpretability of random forest classification ensembles. *IEEE Transactions on Visualization and Computer Graphics*, *27*(2), 1427–1437.

Parr, C. S., Lee, B., Campbell, D., & Bederson, B. B. (2003). *Taxontree: Visualizing biodiversity information* (Tech. Rep.). MARYLAND UNIV COLLEGE PARK INST FOR ADVANCED COMPUTER STUDIES.

Van Den Elzen, S., & Van Wijk, J. J. (2011). Baobabview: Interactive construction and analysis of decision trees. In *2011 ieee conference on visual analytics science and technology (vast)* (pp. 151–160).

Zhao, X., Wu, Y., Lee, D. L., & Cui, W. (2018). iforest: Interpreting random forests via visual analytics. *IEEE transactions on visualization and computer graphics*, *25*(1), 407–416.