XVI Międzynarodowa Konferencja Naukowa im. Profesora Aleksandra Zeliasia "Modelowanie i prognozowanie zjawisk społeczno-gospodarczych" Zakopane 8-11 Maja 2023

HOST Hochschule Stralsund

Can’t See the Forest for the Trees

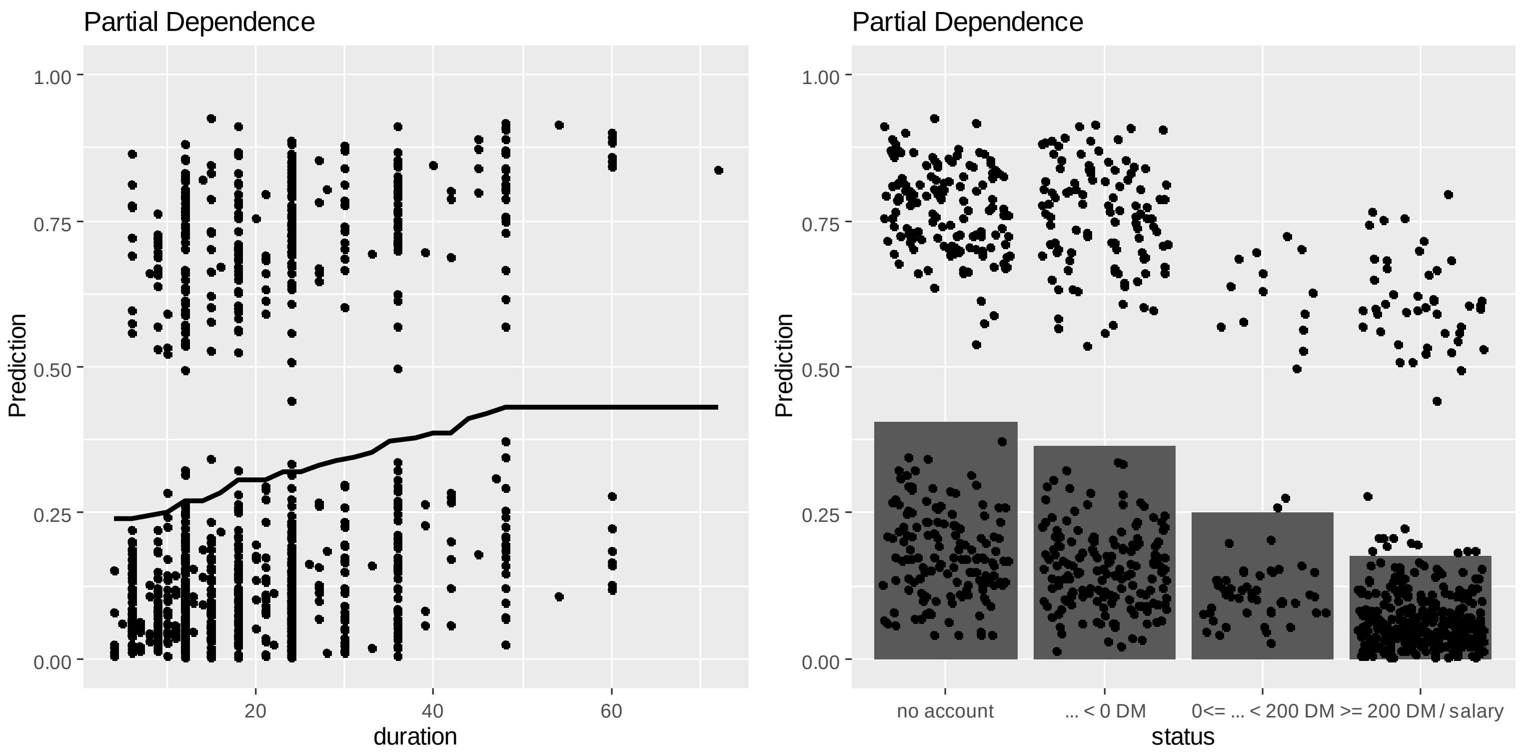
…using Groves to Explain Random Forests

Gero Szepannek
Stralsund University of Applied Sciences, Germany
Björn-Hergen von Holt
Institut für Medizinische Biometrie und Statistik, Lübeck, Germany

Explanation groves:
Tool to analyze complexity vs. explainability

Variable
status.account
duration
credit.history
credit.amount
purpose
savings
age
employment.since
property
other.installments
rate.to.income
personal.status
job
residence.since
housing
other.debtors
num.credits
telephone
numb.people.liable
foreign.worker

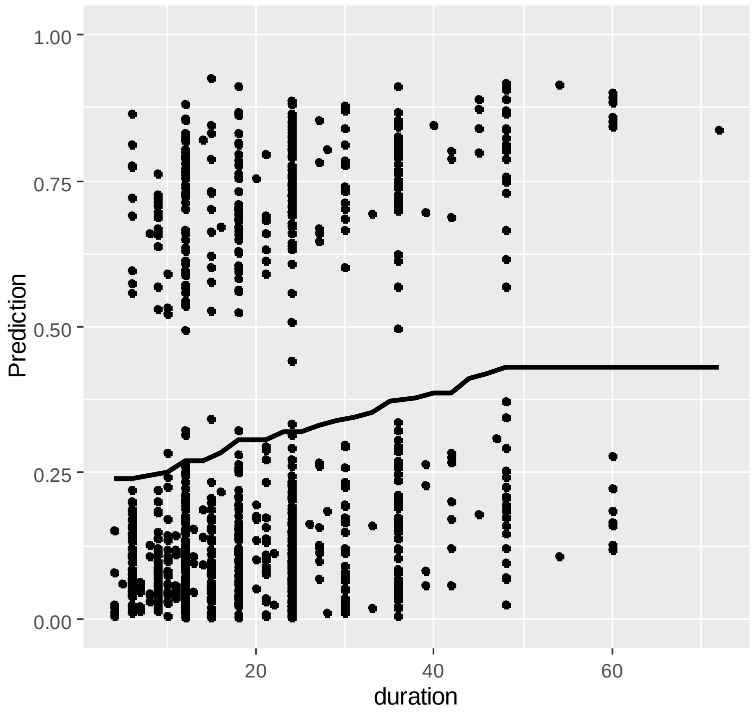
Groemping (2019)



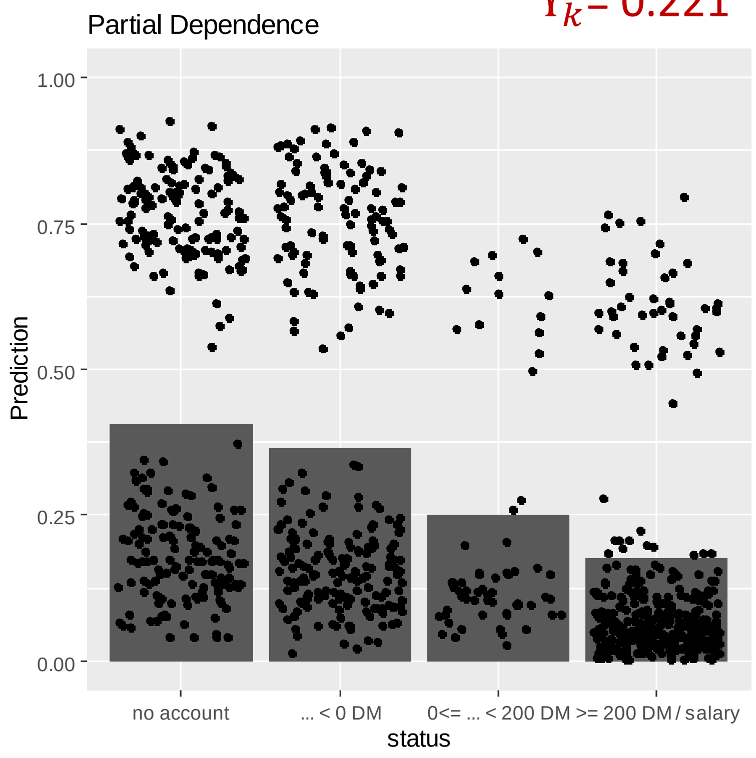
Model agnostic partial dependence plots

Friedman (2001)

$\Upsilon_k = 0.077$



$\Upsilon_k = 0.221$

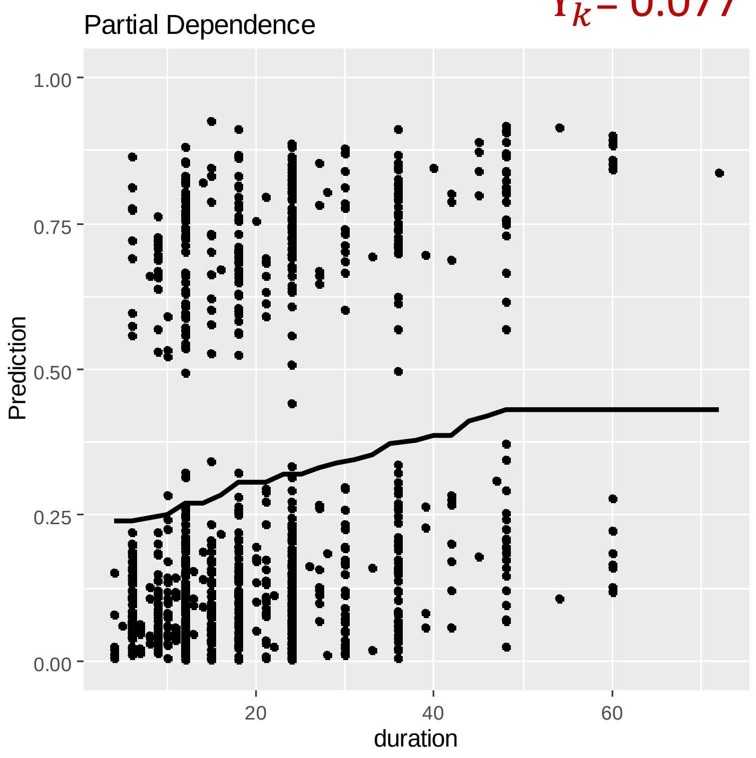


$$ESD(T) = \int (RF(X) - T(X))^2 \, dP(X)$$

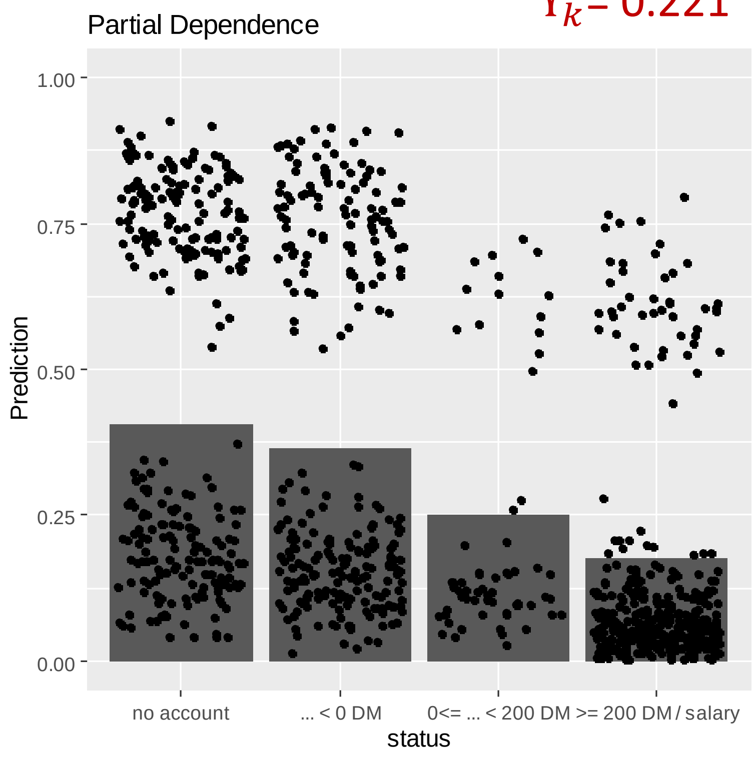
$$\Upsilon(T) = 1 - \frac{ESD(T)}{ESD(\emptyset)}.$$

...it is: $\Upsilon \leq 1$

$\Upsilon_k = 0.077$



$\Upsilon_k = 0.221$

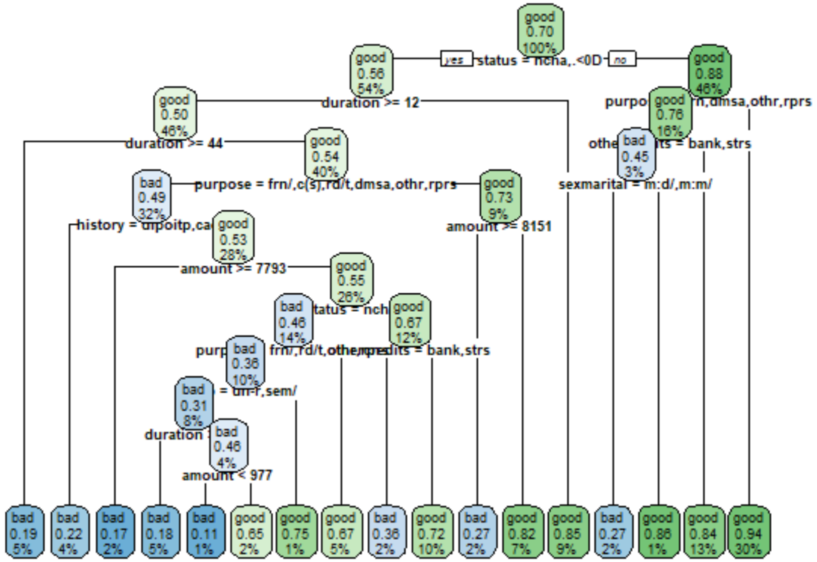


Variable	Step		
status.account	1	0,2210	0,2210
duration	2	0,0770	0,3040
credit.history	3	0,0740	0,3660
credit.amount	4	0,0540	0,4340
purpose	5	0,0390	0,5210
savings	6	0,0440	0,5950
age	7	0,0230	0,6710
employment.since	8	0,0180	0,7420
property	9	0,0170	0,8050
other.installments	10	0,0170	0,8430
rate.to.income	11	0,0040	0,8780
personal.status	12	0,0010	0,9100
job	13	0,0000	0,9370
residence.since	14	0,0000	0,9600
housing	15	0,0130	0,9770
other.debtors	16	0,0040	0,9890
num.credits	17	-0,0010	0,9950
telephone	18	0,0010	0,9980
numb.people.liable	19	0,0000	1,0000
foreign.worker	20	0,0010	1,0000

Random Forest

vs.

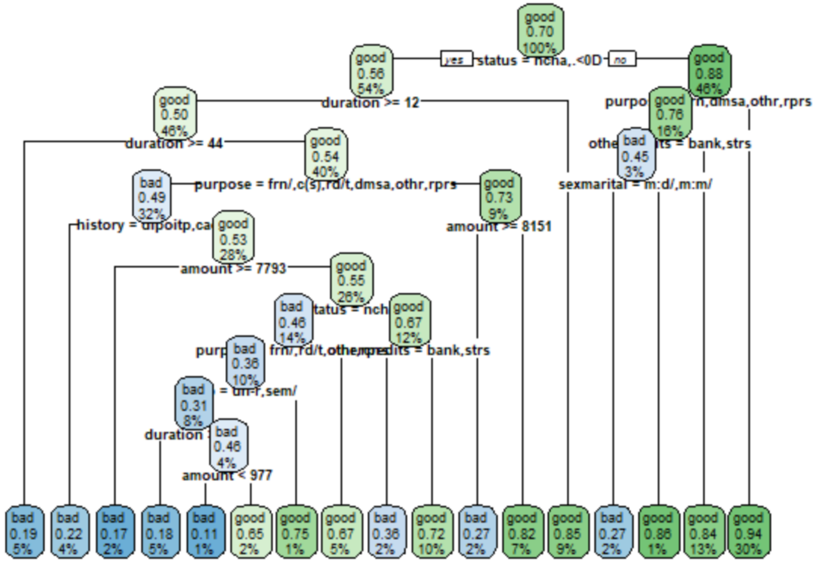
Decision Tree



Random Forest

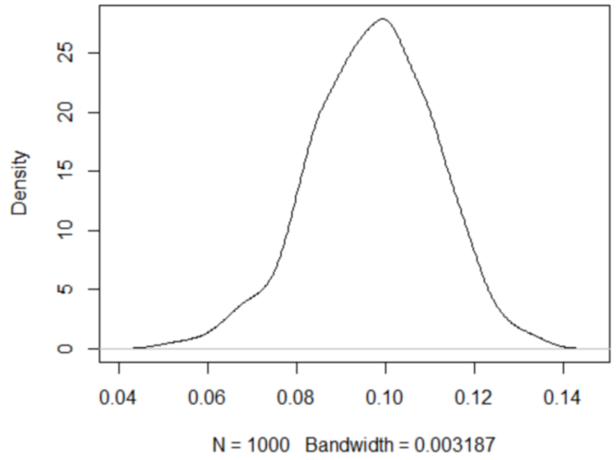
vs.

Decision Tree



	Tree	Forest
AUC	0.679	0.776
CI(AUC)		[0.716, 0.827]
# rules	16	40272

AUC(RF) - AUC(Tree)



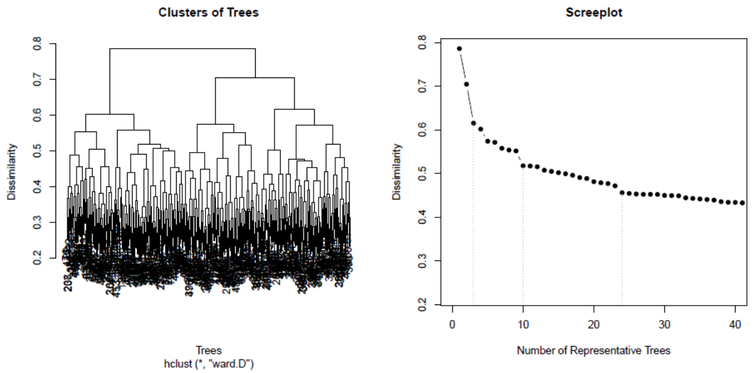
Miller, G. (1956),
Cowan, N. (2010),
DeLong, E., DeLong D. and Clarke-Pearson, D. (1988).

Random Forest

(I) Most Representative Trees
(II) Surrogate Trees
(III) Explanation Groves

Bannerjee, M., Ding, Y., Noone, A.-M. (2012), Molnar (2018).

(I) Most Representative Trees



1. Compute similarity d_j between any two trees of the forest.
2. Select most representative tree with minimum average similarity to all other trees.

$$d_1(Tree_1, Tree_2) = \frac{\#\ of\ covariates\ in\ only\ ony\ of\ both\ trees}{\#\ of\ covariates\ in\ data}$$

$$d_2(Tree_1, Tree_2) = \frac{\#\ pairs\ of\ obs.\ that\ are\ (not)\ in\ the\ same\ terminal\ node\ in\ both\ trees}{\#\ pairs\ of\ obs.\ in\ data,\ i.e.\ \binom{n}{2}}$$

$$d_3(Tree_1, Tree_2) = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_{i,Tree_1} - \hat{y}_{i,Tree_2}\right)^2$$

$$d_4(Tree_1, Tree_2) = \sum_j \left(U_{Tree_2}(X_j) - U_{Tree_2}(X_j)\right)^2$$

$$U_{Tree_k}(X_j) = \sum_n \frac{I_{\{X_j\ is\ used\ in\ split\ @\ node\ n\ at\ depth\ d\ of\ tree\ k\}}}{2^{d-1}} / \max(d)$$

?

Bannerjee, M., Ding, Y., Noone, A.-M. (2012),
von Holt, B. (2020).

1. Compute similarity d_j between any two trees of the forest.
2. Select most representative tree with minimum average similarity to all other trees.

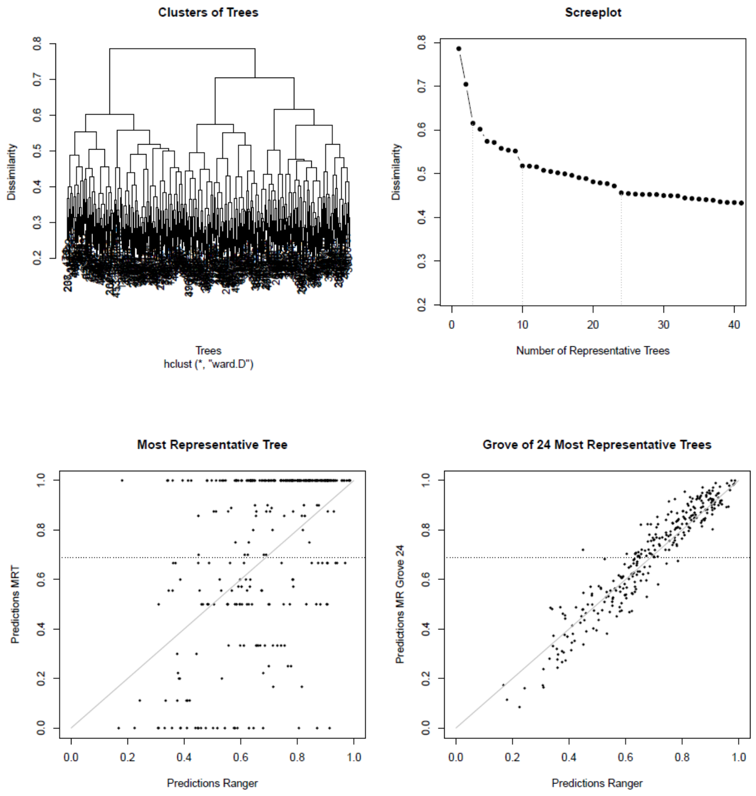
$$d_1(Tree_1, Tree_2) = \frac{\#\ of\ covariates\ in\ only\ ony\ of\ both\ trees}{\#\ of\ covariates\ in\ data}$$

$$d_2(Tree_1, Tree_2) = \frac{\#\ pairs\ of\ obs.\ that\ are\ (not)\ in\ the\ same\ terminal\ node\ in\ both\ trees}{\#\ pairs\ of\ obs.\ in\ data, i.e.\ \binom{n}{2}}$$

$$d_3(Tree_1, Tree_2) = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_{i,Tree_1} - \hat{y}_{i,Tree_2}\right)^2$$

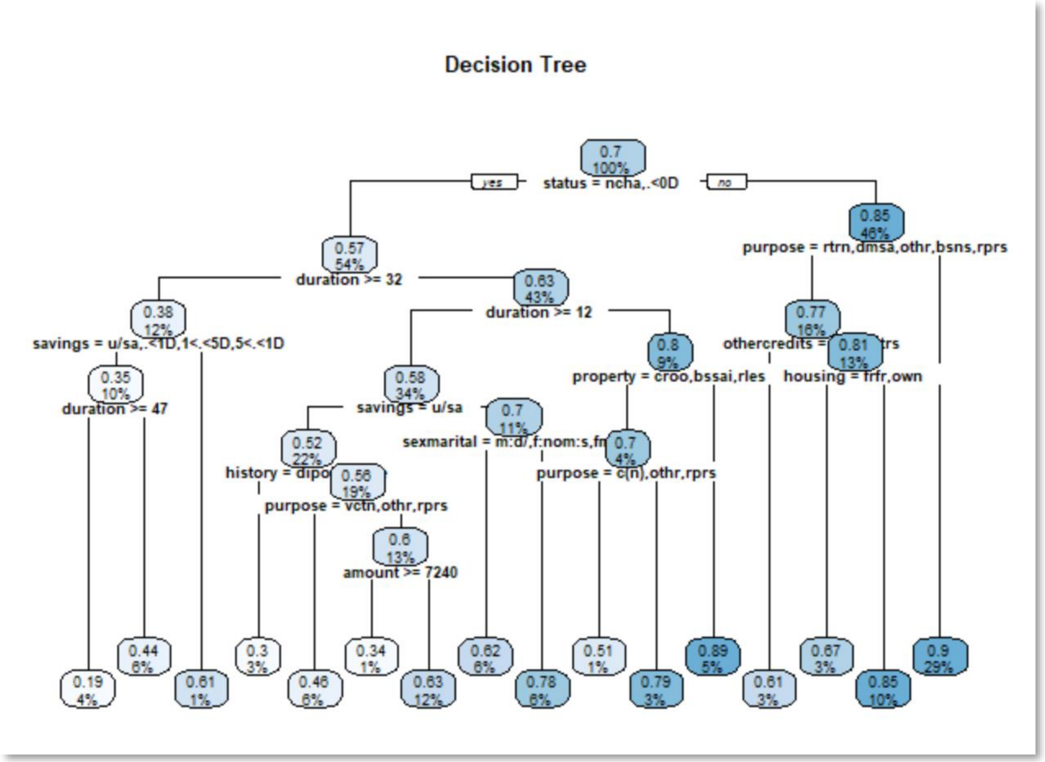
$$d_4(Tree_1, Tree_2) = \sum_j \left(U_{Tree_2}(X_j) - U_{Tree_2}(X_j)\right)^2$$

$$U_{Tree_k}(X_j) = \sum_n \frac{I_{\{X_j\ is\ used\ in\ split\ @\ node\ n\ at\ depth\ d\ of\ tree\ k\}}}{2^{d-1}} / \max(d)$$

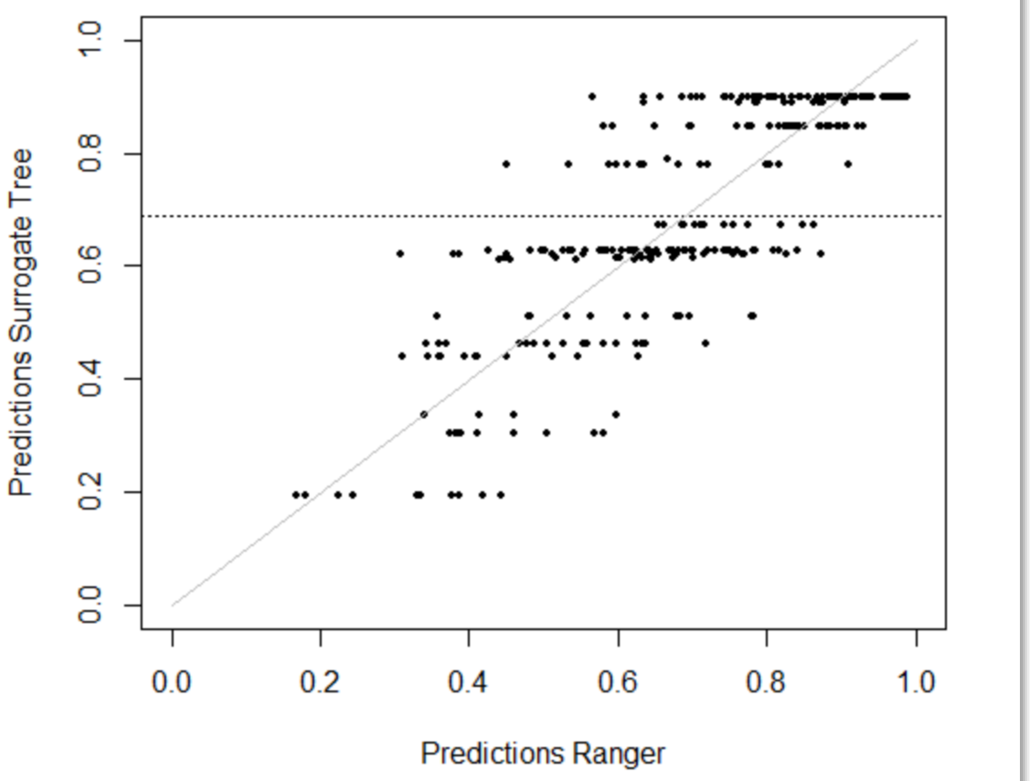


	MRT			Groves		
	d_1	d_3	d_4	3	10	24
Rules	80.54	80	78	252	822	1944
Υ	-2.605	-1.719	-1.859	0.095	0.687	0.871

Bannerjee, M., Ding, Y., Noone, A.-M. (2012),
von Holt, B. (2020).

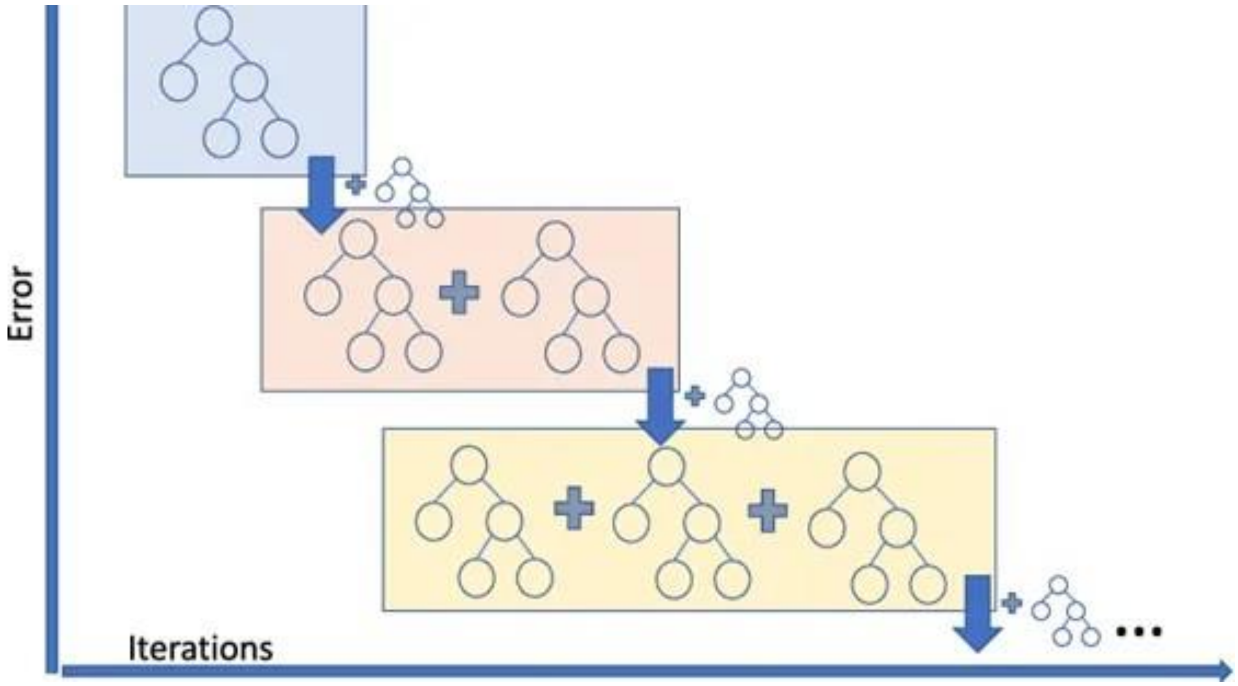


ϒ= 0.590



Molnar, C. (2019),
Therneau, T. and Atkinson, A. (1997).

Sequence of flat regression trees (= explainable rules) with predicted posterior probabilities of the forest as target variable via gradient boosting.



Friedman, J. (2001).

Sequence of flat regression trees (= explainable rules) with predicted posterior probabilities of the forest as target variable via gradient boosting.

`interaction.depth`	Integer specifying the maximum depth of each tree (i.e., the highest level of variable interactions allowed). A value of 1 implies an additive model, a value of 2 implies a model with up to 2-way interactions, etc. Default is 1.

➡ Each tree corresponds to a single rule.

Friedman, J. (2001).

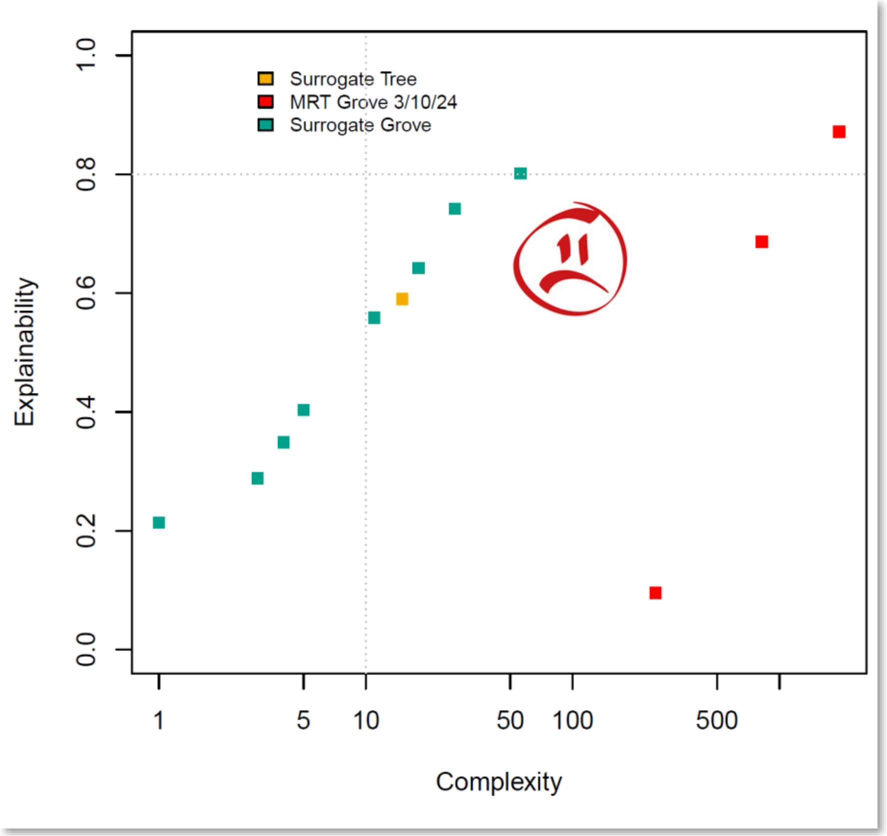
Sequence of <u>flat</u> regression trees (= explainable rules) with predicted posterior probabilities of the forest as target variable via gradient boosting.

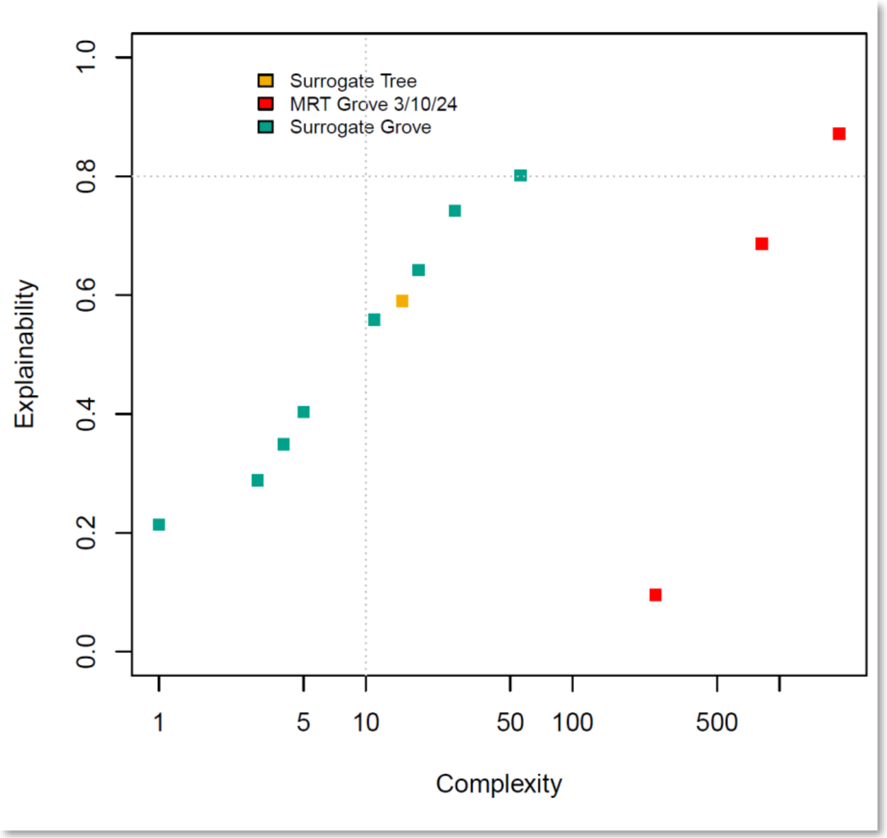
`interaction.depth`	Integer specifying the maximum depth of each tree (i.e., the highest level of variable interactions allowed). A value of 1 implies an additive model, a value of 2 implies a model with up to 2-way interactions, etc. Default is 1.

Variable	Split	Left	Δ_{left}	Δ_{right}
			$\hat{p}_0 = 0.7099$	
status		no checking account | ... < 0 DM	-0.012	0.013
status		no checking account | ... < 0 DM	-0.011	0.012
status		no checking account | ... < 0 DM	-0.012	0.013
status		no checking account | ... < 200 DM	-0.008	0.012
amount	4152		0.006	-0.016
status		no checking account	-0.014	0.007
status		no checking account | ... < 0 DM	-0.007	0.009
status		no checking account | ... < 0 DM	-0.008	0.010
duration	34.5		0.003	-0.018

Five effective rules!

Friedman, J. (2001).


Surrogate Tree
MRT Grove 3/10/24
Surrogate Grove
Explainability
0.0
0.2
0.4
0.6
0.8
1.0
1
5
10
50
100
500
Complexity



	Bank Marketing				Titanic			
	# rules	Υ	STAB	AUC	# rules	Υ	STAB	AUC
ranger	89857			0.917	38184			0.821
rpart	9			0.816	7			0.764
surrogate tree	11	0.784	0.495	0.859	6	0.882	0.025	0.774
surrogate_grove5	5	0.249	0.770	0.802	1	0.435	0.000	0.692
surrogate_grove10	8	0.400	0.649	0.810	3	0.592	0.000	0.754
surrogate_grove20	16	0.563	0.598	0.860	3	0.726	0.000	0.754
surrogate_grove50	36	0.758	0.493	0.894	16	0.806	0.001	0.783
surrogate_grove100	69	0.823	0.436	0.901	36	0.842	0.005	0.790
surrogate_grove500	284	0.858	0.416	0.910	125	0.865	0.012	0.800

- ✓ **Most representative trees** suffer from depth of single trees in forests.
- ✓ **Explanation groves (EGs)** extract set of explainable additive surrogate rules...
- ✓ It‘s questionable whether there exists an understandable explanation for a complex model.
- ✓ ...EGs can be used to **increase complexity** of an explanation and **relate it to the resulting explainability**.

https://github.com/g-rho/xgrove

1. Bannerjee, M., Ding, Y., Noone, A.-M. (2012). Identifying representative trees from ensembles. Stat in Med 31:1601-16.
2. Bücker, M., Szepannek, G., Gosiewska, A. and Biecek., P. (2021). Transparency, Auditability and explainability of machine learning models in credit scoring. Journal of the Operational Research Society 73: 1–21.
3. Cowan N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why?. Current directions in psychological science, 19(1), 51–57. https://doi.org/10.1177/0963721409359277
4. DeLong, E., DeLong D. and Clarke-Pearson, D. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837–845.
5. Fernández-Delgado, M., Cernadas, E. Barro and S. Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? JMLR 15, 1, 3133–3181.
6. Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. The Annals of Statistics,29:1189–1232.
7. Groemping, U. (2019). South German credit data: Correcting a widely used data set. Available at: http://www1.beuth-hochschule.de/ FB_II/reports/Report-2019-004.pdf.
8. von Holt, B. (2020). timbR: Tree interpretation methods based on range, R package version 0.1.0 https://github.com/imbs-hl/timbR.
9. von Holt, B., Westenberger, A., König, I. (2022): Identification of representative trees in random forests based on a new tree-based distance measure. https://www.biorxiv.org/content/10.1101/2022.05.15.492004v1
10. Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review 63 (2).
11. Molnar, C. (2019). Interpretable Machine Learning, Lulu, ISBN: 978-0244768522.
12. Probst, P., Boulesteix, A.-L. and Bischl, B. (2019). Tunability: importance of hyperparameters of machine learning algorithms. JMLR 20, 1934–1965.
13. Ridgeway, G. (2006). Generalized Boosted Models: A guide to the gbm package.
14. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1: 206–15.
15. Szepannek, G. (2017). On the practical relevance of modern Machine Learning Algorithms for Credit Scoring Applications. WIAS Report Series, 29, 88–96.
16. Szepannek, G. (2019). How Much Can We See? A Note on Quantifying Explainability of Machine Learning Models. https://arxiv.org/abs/1910.13376
17. Szepannek, G. (2022). An Overview on the Landscape of R Packages for Open Source Scorecard Modelling. Risks 10, no. 3: 67. https://doi.org/10.3390/risks10030067
18. Szepannek, G., Luebke, K. (2021). Facing the Challenges of Developing Fair Risk Scoring Models, Frontiers in Artificial Intelligence 4, DOI: 10.3389/frai.2021.681915.
19. Szepannek, G. and Luebke, K. (2023): How much do we see? On the explainability of partial dependence plots for credit risk scoring, Argumenta Oeconomica 50, DOI: 10.15611/aoe.2023.1.07.
20. Therneau, T. and Atkinson, A. (1997). An introduction to recursive partitioning using the rpart routines. Divsion of Biostatistics, TR61.
21. Wright M., Ziegler . (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1–17. doi: 10.18637/jss.v077.i01.

Thank you!