Interaction forests: Identifying and exploiting interpretable quantitative and qualitative interaction effects

Roman Hornung, Anne-Laure Boulesteix

Institute for Medical Information Processing, **B**iometry and **E**pidemiology, University of Munich

ROeS 2021, Salzburg

September, 7th, 2021

Introduction

- **Interaction effects allow** valuable **insights** into the interplay between covariates – e.g., a medical treatment may have a strong effect for a subgroup of patients.

- Modeling these effects can also **improve automatic prediction** rules.

- **Most tree-based approaches** to modeling interaction effects **use univariable splitting**.

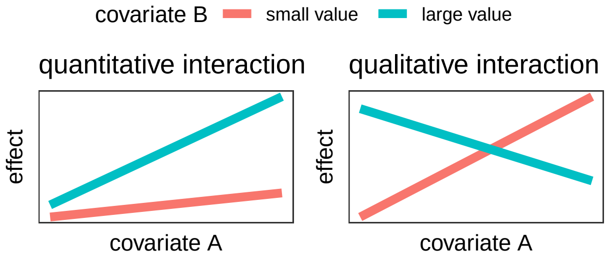
  $\Rightarrow$ **Interaction effects** of covariate pairs **without strong marginal effects not modeled** effectively.

  – In **interaction forests (IFs)** we use **bivariable splitting** to model interaction effects.

Introduction

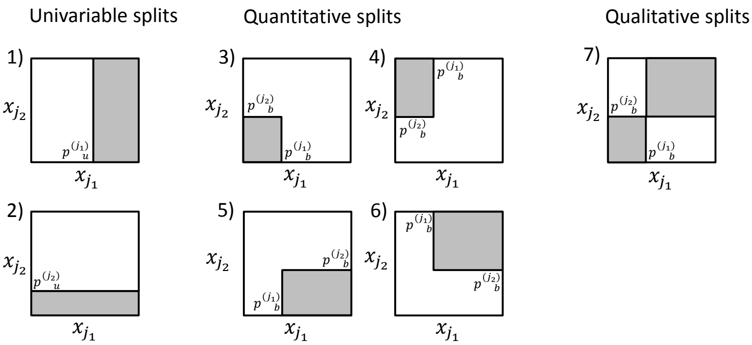
- Interaction forests (**IF**) model **well interpretable** and **communicable** interaction **effects** (keyword: **interpretable machine learning**).

- The Effect Important Measure (**EIM**) of IF **ranks covariate pairs** separately **with respect to** the predictive importance of their **quantitative** and **qualitative interaction effects**.



Interaction forest algorithm: Split types

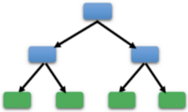
Seven split types considered in the trees of IF:



(x_{j_1}, x_{j_2}): specific **pair of covariates**; $p_u^{(j_1)}$, $p_u^{(j_2)}$: **univariable split points**; $(p_b^{(j_1)}, p_b^{(j_2)})$: **bivariable split points**.

Interaction forest algorithm: Tree growing

Trees are grown using **recursive binary splitting** (as in conventional random forests (**RF**)).

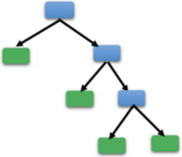


Each split in the trees is **found as follows**:

1. **Candidate split sampling**.

   For $pair = 1, \ldots, npairs$:

   1. Sample **one covariate pair** (x_{j_1}, x_{j_2}).
   2. Sample four **split points** in (x_{j_1}, x_{j_2}): $p_u^{(j_1)}$, $p_u^{(j_2)}$, $(p_b^{(j_1)}, p_b^{(j_2)})$
   3. **Add** to the candidate split set **seven splits - one of each of the seven split types** - associated with the split points sampled in 2.



2. **Select** the **best candidate split** out of 1 (i.e., the one associated with the best split criterion value).

Interaction forest algorithm: Effect Importance Measure (**EIM**)

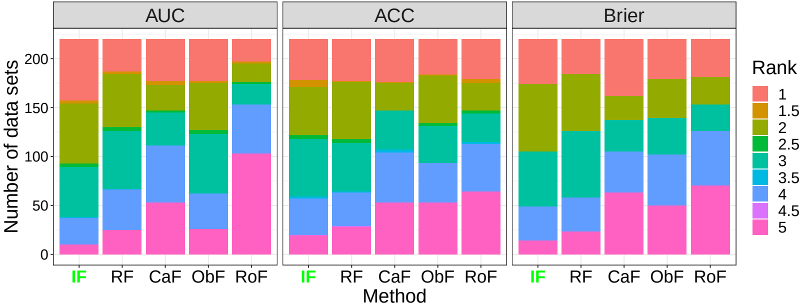
Sketch of the procedure for calculating the **EIM values**:

1. For each **covariate** / **covariate pair**, **measure** its **importance** (Hapfelmeier et al., 2014) **separately** with respect to **each split type**.
2. Obtain **three lists**:
   1. **univariable EIM** values: **Rank covariates** with respect to predictive importance (as in conventional RF).
   2. **quantitative EIM** values: **Rank** covariate **pairs** with respect to predictive importance of **quantitative interaction effects**.
   3. **qualitative EIM** values: **Rank** covariate **pairs** with respect to predictive importance of **qualitative interaction effects**.

Comparison study: Prediction performance – **real data study** design

- **methods**:
  - interaction forests (**IF**) (Hornung and Boulesteix, 2021)
  - random forests (**RF**) (Breiman, 2001)
  - canonical correlation forests (**CaF**) (Rainforth and Wood, 2015)
  - oblique random forests (**ObF**) (Menze et al., 2011)
  - rotation forests (**RoF**) (Rodríguez et al., 2006)
- **220** publicly available **data sets** with **binary outcome** obtained from OpenML (Vanschoren et al., 2013)
- **performance metrics**: area under the ROC curve (**AUC**), accuracy (**ACC**), Brier score (**Brier**)
- **validation scheme**: 5 times repeated stratified 5-fold **cross-validation**

Data set specific ranks of **each method among the other methods** in terms of the respective performance metric.



Comparison study: Performance in interaction detection – **simulation study** design

- **methods**:
  1. quantitative and qualitative EIM of IF (**IF-EIM-quant**, **IF-EIM-qual**)
  2. paired association measure (PA) (Ishwaran, 2007)
  3. Interaction Minimal Depth Maximal Subtree measure (IMDMS) (Dazard et al., 2018)
  4. stability score of iterative random forests (iRF) (Basu et al., 2018)
  5. baseline method: naive RF based measure that uses marginal effects only (RF-V-pairs)
- **binary** balanced **outcome**; 68 covariates: 6 with **main effects only**, 3 pairs of covariates with **quantitative** / **qualitative interaction effects** each, 50 without effect
- three levels of strength for each effect type: **strong**, **moderate**, **weak**; $n = 100, 500, 1000$; repetitions: 200

Comparison study: Performance in interaction detection – results I

- **IF-EIM-quant** and **IF-EIM-qual** ranked the interacting covariate pairs **better** (median ranks, interquartile range) than the competing methods.

- $n = 100$: only **strong qualitative** interactions detected consistently

- $n = 500, 1000$: **all qualitative** interactions and **moderate and strong quantitative** interactions detected

Comparison study: Performance in interaction detection – results II

- **IF-EIM-qual** and **IF-EIM-quant specific** for qualitative and quantitative interactions, respectively

- **IF-EIM-quant** and, in particular, **IF-EIM-qual** attributed **bad ranks** to **non-interacting covariate pairs** with main effects only; ✓ in contrast, the competing methods tended to rank these pairs very low. ✗

Further things & Outlook

- **IFs are** specific **diversity forests** (Hornung, 2020): split sampling **allows** using **complex split procedures**.

- **Pre-processing** steps **of** the **IF** algorithm:

  1. **ordering of categories** for unordered **categorical covariates**
  2. **if $p > 100$: pre-selection of 5000** likely interacting covariate **pairs** using a screening procedure

Further things & Outlook

- **IF** implemented for **categorical**, **continuous**, and **survival** outcomes in the **R package 'diversityForest'** (closely based on **ranger** (Wright and Ziegler, 2017))

- **Important** analysis step: (flexible) **estimation** of the **forms of** the **interaction effects** identified using EIM
  - **functions for visualization** available in **diversityForest**

- Possible future work: **Testing** procedure for **EIM**

References – I

Basu, S., Kumbier, K., Brown, J.B., Yu, B., 2018.
Iterative random forests to discover predictive and stable high-order interactions.
Proceedings of the National Academy of Sciences (PNAS) 115 (8), 1943–1948.

Breiman, L., 2001.
Random forests.
Machine Learning 45 (1), 5–32.

Dazard, J.E., Ishwaran, H., Mehlotra, R., Weinberg, A., Zimmerman, P., 2018.
Ensemble survival tree models to reveal pairwise interactions of variables with time-to-events outcomes in low-dimensional setting.
Statistical Applications in Genetics and Molecular Biology 17 (1), 20170038.

Hapfelmeier, A., Hothorn, T., Ulm, K., Strobl, C., 2014.
A new variable importance measure for random forests with missing data.
Statistics and Computing 24, 21–34.

Hornung, R., 2020.
Diversity forests: Using split sampling to allow for complex split procedures in random forest.
Technical report 234, Department of Statistics, University of Munich.

Hornung, R., Boulesteix, A.-L., 2021.
Interaction forests: Identifying and exploiting interpretable quantitative and qualitative interaction effects.
Technical report 237, Department of Statistics, University of Munich.

References – II

Ishwaran, H., 2007.
Variable importance in binary regression trees and forests.
Electronic Journal of Statistics 1, 519–537.

Menze, B.H., Kelm, B.M., Splitthoff, D.N., Koethe, U., Hamprecht, F.A., 2011.
On oblique random forests.
In: D. Gunopulos, T. Hofmann, D. Malerba, M. Vazirgiannis (Eds.), European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, pp. 453–469.

Rainforth, T., Wood, F., 2015.
Canonical correlation forests.
arXiv:1507.05444.

Rodríguez, J.J., Kuncheva, L.I., Alonso, C.J., 2006.
Rotation forest: A new classifier ensemble method.
IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (10), 1619–1630.

Vanschoren, J., van Rijn, J.N., Bischl, B., Torgo, L., 2013.
OpenML: Networked Science in Machine Learning.
SIGKDD Explorations 15 (2), 49–60.

Wright, M.N., Ziegler, A., 2017.
ranger: A fast implementation of random forests for high dimensional data in C++ and R.
Journal of Statistical Software 77, 1–17.

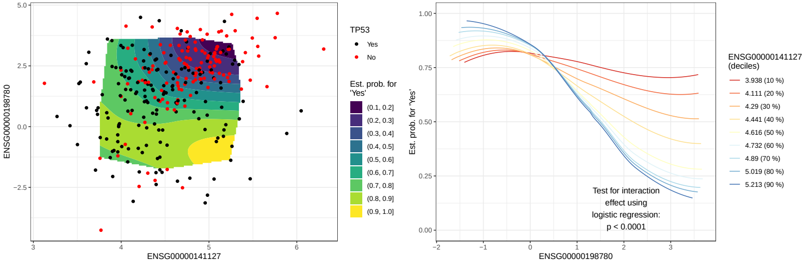
Thank you for your attention!

Recommended reading:

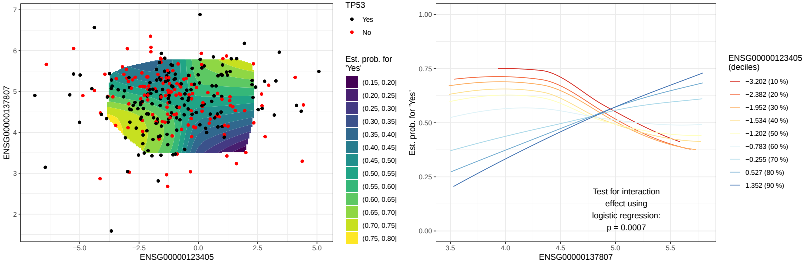
Hornung, R., Boulesteix, A.-L., 2021.
Interaction forests: Identifying and exploiting interpretable quantitative and qualitative interaction effects.
Technical report 237, Department of Statistics, University of Munich.

Visual exploration of interaction effects

Quantitative interaction effect



Qualitative interaction effect



Real data study results: Performances of the methods summarized across the 220 data sets

	AUC	ACC	Brier
IF	**0.9182 [0.7820, 0.9862]**	**0.8822 [0.7664, 0.9499]**	**0.0890 [0.0425, 0.1641]**
RF	0.9110 [0.7715, 0.9826]	0.8796 [0.7670, 0.9503]	0.0923 [0.0407, 0.1658]
CaF	0.8842 [0.7660, 0.9781]	0.8761 [0.7555, 0.9468]	0.0962 [0.0391, 0.1748]
ObF	0.9051 [0.7721, 0.9824]	0.8644 [0.7356, 0.9465]	0.0985 [0.0461, 0.1818]
RoF	0.8632 [0.7652, 0.9685]	0.8676 [0.7544, 0.9421]	0.1016 [0.0437, 0.1686]

The numbers show the medians of the cross-validated metrics across the data sets. The numbers in square brackets show the 25% quantiles and 75% quantiles (i.e., the first and third quartiles) of the cross-validated metrics obtained for each data set. Larger AUC values, larger ACC values, and smaller Brier values indicate a better performance.

Simulation study results: Quantitative interaction effects

Effect:	Strong	Moderate	Weak
	n = 100		
IF-EIM-quant	**19.0 [5.0, 75.8]**	**141.0 [33.0, 452.0]**	**675.0 [237.0, 1361.5]**
RF-V-pairs	199.0 [79.5, 285.2]	329.5 [208.8, 491.2]	704.0 [493.5, 1066.5]
PA	107.5 [28.8, 579.8]	324.5 [91.0, 756.5]	729.0 [288.5, 1411.2]
IMDMS	77.5 [20.0, 189.2]	259.5 [111.8, 442.5]	499.5 [300.2, 872.5]
iRF	16.0 [5.5, 25.5] (46%)	29.5 [18.2, 36.8] (17%)	37.0 [30.0, 50.0] (2%)
	n = 500		
IF-EIM-quant	**1.0 [1.0, 2.0]**	**7.0 [4.0, 20.0]**	**100.5 [35.0, 251.5]**
RF-V-pairs	138.0 [79.8, 156.2]	331.0 [268.0, 392.2]	532.5 [457.0, 593.0]
PA	11.0 [5.0, 21.2]	34.0 [17.0, 149.8]	294.0 [100.0, 946.2]
IMDMS	22.0 [16.8, 30.0]	147.5 [71.2, 257.2]	510.5 [382.2, 589.5]
iRF	26.0 [18.0, 38.0] (85%)	59.0 [43.5, 73.0] (18%)	– [–, –] (0%)
	n = 1000		
IF-EIM-quant	**1.0 [1.0, 1.0]**	**3.0 [2.0, 5.0]**	**43.0 [20.0, 108.0]**
RF-V-pairs	138.5 [86.8, 142.0]	332.0 [271.0, 389.0]	570.0 [513.0, 626.0]
PA	11.0 [5.8, 46.2]	35.0 [14.0, 186.2]	360.0 [117.5, 955.8]
IMDMS	24.0 [19.0, 29.0]	160.0 [86.5, 211.8]	515.0 [442.8, 592.5]
iRF	28.5 [17.0, 44.0] (99%)	77.0 [65.0, 91.0] (22%)	87.0 [87.0, 87.0] (0%)

The numbers show the median ranks the respective covariates obtained across the simulated data sets. The numbers in square brackets show the 25% quantiles and 75% quantiles of the ranks obtained for the simulated data sets.

Simulation study results: Qualitative interaction effects

Effect:	Strong	Moderate	Weak
	n = 100		
IF-EIM-qual	**1.0 [1.0, 3.0]**	**10.0 [2.0, 217.5]**	**263.0 [21.8, 1058.8]**
RF-V-pairs	1265.5 [813.2, 1721.2]	1323.5 [957.8, 1786.5]	1439.5 [1028.2, 1853.8]
PA	145.5 [48.5, 428.2]	403.5 [111.8, 1076.8]	932.5 [437.0, 1676.0]
IMDMS	800.5 [582.0, 1170.2]	906.5 [629.8, 1394.5]	1129.5 [804.0, 1495.2]
iRF	35.0 [35.0, 35.0] (0%)	– [–, –] (0%)	– [–, –] (0%)
	n = 500		
IF-EIM-qual	**1.0 [1.0, 1.0]**	**2.0 [2.0, 2.0]**	**3.0 [3.0, 5.0]**
RF-V-pairs	837.5 [748.5, 1032.8]	1022.5 [802.2, 1311.0]	1256.0 [973.5, 1577.0]
PA	20.5 [15.8, 26.0]	37.0 [26.0, 72.0]	152.0 [79.0, 321.5]
IMDMS	740.0 [692.8, 796.5]	801.5 [739.0, 926.0]	919.0 [772.2, 1116.8]
iRF	– [–, –] (0%)	– [–, –] (0%)	– [–, –] (0%)
	n = 1000		
IF-EIM-qual	**1.0 [1.0, 1.0]**	**2.0 [2.0, 2.0]**	**3.0 [3.0, 3.0]**
RF-V-pairs	745.0 [739.0, 782.5]	825.0 [759.0, 924.2]	1149.5 [947.0, 1470.0]
PA	52.0 [26.0, 134.0]	188.5 [106.5, 374.5]	796.0 [297.5, 1538.5]
IMDMS	739.0 [739.0, 740.0]	740.0 [739.0, 779.2]	849.5 [770.0, 955.2]
iRF	– [–, –] (0%)	– [–, –] (0%)	– [–, –] (0%)

The numbers show the median ranks the respective covariates obtained across the simulated data sets. The numbers in square brackets show the 25% quantiles and 75% quantiles of the ranks obtained for the simulated data sets.

Simulation study results: Specificity of IF-EIM-qual and IF-EIM-quant

	n = 100	n = 500	n = 1000
IF-EIM-qual: Quantitative interaction effects			
Strong	533.0 [177.2, 1190.0]	235.0 [40.0, 744.0]	86.5 [14.0, 397.0]
Moderate	705.5 [249.8, 1269.5]	403.5 [92.0, 1020.2]	299.0 [34.0, 788.2]
Weak	1020.5 [470.0, 1640.0]	784.5 [339.8, 1360.2]	664.5 [233.8, 1266.2]
IF-EIM-quant: Qualitative interaction effects			
Strong	252.0 [110.0, 515.5]	177.0 [81.5, 307.2]	169.0 [72.5, 293.5]
Moderate	429.5 [170.5, 838.0]	254.5 [136.2, 517.0]	279.5 [165.8, 478.8]
Weak	938.5 [462.2, 1450.2]	466.0 [275.5, 800.5]	500.0 [334.8, 891.2]
IF-EIM-qual: Pairs with main effects only			
Strong	992.5 [591.2, 1414.2]	741.0 [382.0, 1117.5]	665.0 [372.8, 1031.0]
Moderate	1033.0 [606.2, 1460.5]	816.0 [411.5, 1350.8]	739.5 [345.0, 1134.2]
Weak	1004.5 [571.2, 1589.0]	951.0 [487.5, 1443.2]	788.5 [328.2, 1328.0]
IF-EIM-quant: Pairs with main effects only			
Strong	28.0 [6.0, 114.8]	23.5 [6.8, 124.0]	17.0 [5.0, 72.0]
Moderate	93.5 [20.0, 329.8]	52.0 [11.0, 211.2]	31.0 [12.0, 140.0]
Weak	338.0 [154.8, 1028.2]	192.0 [65.0, 465.5]	111.0 [29.8, 346.8]

The numbers show the median ranks the respective covariates obtained across the simulated data sets. The numbers in square brackets show the 25% quantiles and 75% quantiles of the ranks obtained for the simulated data sets.

Simulation study results: Median ranks obtained for covariate pairs with main effects, but without interaction effects

Effect:	Strong	Moderate	Weak
	n = 100		
IF-EIM-qual	**992.5 [591.0, 1414.2]**	**1032.5 [606.2, 1460.5]**	**1004.5 [571.2, 1589.0]**
IF-EIM-quant	**28.0 [6.0, 114.8]**	**93.5 [20.0, 329.8]**	**338.0 [154.8, 1028.2]**
RF-V-pairs	2.0 [1.0, 6.0]	93.0 [15.8, 163.2]	336.5 [222.5, 475.5]
PA	5.0 [2.0, 33.8]	56.0 [13.8, 246.2]	497.5 [136.8, 1433.5]
IMDMS	2.0 [1.0, 6.0]	30.5 [11.0, 79.0]	297.5 [140.5, 491.2]
iRF	3.0 [1.0, 6.0] (99%)	15.0 [7.0, 24.0] (72%)	26.0 [16.5, 36.0] (14%)
	n = 500		
IF-EIM-qual	**741.0 [382.0, 1117.5]**	**816.0 [411.5, 1350.8]**	**951.0 [487.5, 1443.2]**
IF-EIM-quant	**23.5 [6.8, 124.0]**	**52.0 [11.0, 211.2]**	**192.0 [65.0, 465.5]**
RF-V-pairs	1.0 [1.0, 1.0]	77.0 [16.0, 136.0]	343.0 [271.0, 399.8]
PA	1.0 [1.0, 2.0]	15.5 [9.0, 30.2]	195.5 [47.8, 714.2]
IMDMS	1.0 [1.0, 1.0]	13.0 [9.0, 18.0]	233.0 [143.8, 316.2]
iRF	1.0 [1.0, 2.0] (100%)	17.0 [11.0, 26.0] (98%)	52.0 [45.0, 63.5] (18%)
	n = 1000		
IF-EIM-qual	**665.0 [372.8, 1031.0]**	**739.5 [345.0, 1134.2]**	**788.5 [328.2, 1328.0]**
IF-EIM-quant	**17.0 [5.0, 72.0]**	**31.0 [12.0, 140.0]**	**111.0 [29.8, 346.8]**
RF-V-pairs	1.0 [1.0, 1.0]	79.0 [18.0, 136.0]	339.0 [282.0, 395.0]
PA	1.0 [1.0, 2.0]	19.5 [8.0, 83.5]	201.0 [64.5, 553.8]
IMDMS	1.0 [1.0, 1.0]	13.0 [10.0, 18.0]	203.0 [155.8, 282.0]
iRF	1.0 [1.0, 2.0] (100%)	11.0 [7.0, 17.0] (100%)	78.5 [63.0, 92.8] (27%)

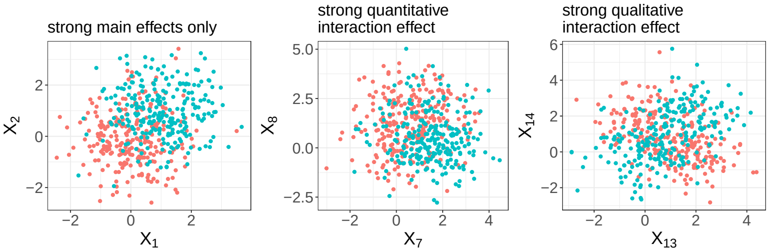
The numbers show the median ranks the respective covariates obtained across the simulated data sets. The numbers in square brackets show the 25% quantiles and 75% quantiles of the ranks obtained for the simulated data sets.

Simulation study results: Univariable effects

Effect:	Strong	Moderate	Weak
	n = 100		
IF-EIM-univ	**2.0 [1.0, 3.0]**	**4.0 [3.0, 7.0]**	**9.0 [6.0, 16.0]**
RF-V	2.0 [1.0, 3.0]	4.0 [3.0, 7.0]	10.0 [6.0, 17.0]
	n = 500		
IF-EIM-univ	**2.0 [1.0, 2.0]**	**4.0 [3.0, 5.0]**	**9.0 [8.0, 10.0]**
RF-V	2.0 [1.0, 2.0]	4.0 [3.0, 5.0]	9.0 [7.0, 10.0]
	n = 1000		
IF-EIM-univ	**2.0 [1.0, 2.0]**	**4.0 [3.0, 5.0]**	**9.0 [8.0, 10.0]**
RF-V	2.0 [1.0, 2.0]	4.0 [3.0, 5.0]	9.0 [8.0, 10.0]

The numbers show the median ranks the respective covariates obtained across the simulated data sets. The numbers in square brackets show the 25% quantiles and 75% quantiles of the ranks obtained for the simulated data sets.

Exemplary pairs of covariates with strong effects in a simulated data set (sample size: 500)



Each point corresponds to an observation in the data set. The two colors distinguish the two outcome classes, where red and green points show observations from the first and second class, respectively.