

# Explaining Tree Models

Bernhard Jaeger

University of Tübingen

Tübingen, Germany

bernhard.jaeger@student.uni-tuebingen.de

**Abstract**—Understanding the decisions made by machine learning models can be challenging due to their black box nature. We investigate the additive feature attribution method SHAP (SHapley Additive exPlanations) that addresses this problem [2]. While SHAP can be used to explain the decisions of any model, the generic KernelSHAP algorithm for it can be intractable for larger datasets. For tree models there exists a specialized algorithm (Tree SHAP) that can compute exact SHAP values in polynomial time [1]. We use the SHAP framework to investigate the relationship between Decision Trees and Hoeffding Trees [3]. Theory tells us that a Hoeffding Tree should converge towards a Decision Tree when trained. In our experiments however we could not observe convergence. The reason for this remains unclear.

**Index Terms**—Tree SHAP, SHAP, Decision Tree, Hoeffding Tree

## I. INTRODUCTION

This . This and [3] document is a model and instructions for L<sup>A</sup>T<sub>E</sub>X. Please observe the conference page limits.

## II. SHAP VALUES

## III. DECISION AND Hoeffding TREES

### A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

### B. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \quad (1)$$

### C. L<sup>A</sup>T<sub>E</sub>X-Specific Advice

Please use “soft” (e.g., `\eqref{Eq}`) cross references instead of “hard” references (e.g., (1)). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

### D. Figures and Tables

a) *Positioning Figures and Tables:* Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

TABLE I  
TABLE TYPE STYLES

Table Head	Table Column Head		
	Table column subhead	Subhead	Subhead
copy	More table copy <sup>a</sup>		

<sup>a</sup>Sample of a Table footnote.



Fig. 1. Example of a figure caption.

## IV. METHODOLOGY

## V. EVALUATION

## VI. CONCLUSION

## VII. REPRODUCIBILITY CONSIDERATIONS

### A. Software Libraries:

The software library versions we used (read out using pip –freeze) are documented in table II.

All experiments were conducted on a single personal computer. The hardware specs can be found in table III.

## ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

TABLE II  
SOFTWARE LIBRARY VERSIONS

Software library	Version
Python	3.7.4
anaconda-client	1.7.2
anaconda-navigator	1.9.7
anaconda-project	0.8.3
conda	4.7.12
conda-build	3.18.9
conda-package-handling	1.6.
0 conda-verify	3.4.2
ipykernel	5.1.2
ipython	7.8.0
ipython-genutils	0.2.0
jupyter	1.0.0
jupyter-client	5.3.3
jupyter-console	6.0.0
jupyter-core	4.5.0
jupyterlab	1.1.4
jupyterlab-server	1.0.6
matplotlib	3.1.1
notebook	6.0.1
numpy	1.16.5
pandas	0.25.1
scikit-image	0.15.0
scikit-learn	0.21.3
scikit-multiflow	0.5.0
scipy	1.3.1
shap	0.35.0
sklearn	0.0
torch	1.5.0+cu101
torchvision	0.6.0+cu101

TABLE III  
HARDWARE SPECIFICATIONS:

Component	Model
Processor	Intel Core i7-6700 CPU
RAM	16 GB Corsair Vengeance LPX DDR4-2400
GPU	NVIDIA GeForce GTX 980
Motherboard	ASUS Z170 Pro Gaming
Operating system	Microsoft Windows 10 Education N
Hard Drive	SSD Crucial MX200

## REFERENCES

## REFERENCES

- [1] Scott M. Lundberg, Gabriel G. Erion, Su-In Lee: Consistent Individualized Feature Attribution for Tree Ensembles. CoRR abs/1802.03888 (2018).
- [2] Scott M. Lundberg, Su-In Lee: A Unified Approach to Interpreting Model Predictions. NIPS 2017: 4765-4774.
- [3] Pedro M. Domingos, Geoff Hulten: Mining high-speed data streams. KDD 2000: 71-80