

SEMINAR ANNOUNCEMENT

Wednesday March 12, 2025

at 11:00 am Room "Sala Seminari" - Abacus Building (U14)

Optimization models and algorithms for sparse soft decision trees

Speaker Antonio Consolo

post-doc researcher DISCo

Abstract

Decision trees are supervised machine learning models widely used for classification and regression tasks. They are inherently interpretable since, for every input vector, they reveal to the domain experts the clear sequence of decisions leading to the tree response. Although building an opti- mal decision tree is known to be NP-hard, recent research has shown growing interest in globally optimized trees, where the entire tree is trained by optimizing an error function over all its pa- rameters. This talk focuses on soft decision trees, a variant of classical decision trees in which the traditional deterministic hard splitting rule at each branch (internal) node is replaced with a probabilistic soft rule. Specifically, we investigate and improve soft decision trees for classifi- cation and regression tasks by devising new model variants, establishing theoretical properties, and developing decomposition algorithms for training them. Particular attention is given to sparsity, to improve model interpretability, and group fairness, to mitigate potential algorithmic biases affecting sensitive groups.

Concerning soft classification trees, we propose alternative sparsification methods based on concave approximation of the l0 norm. Experiments on various benchmark datasets show that our approach yield sparser trees than standard l1 and $l\infty$ regularizations without compromising accuracy. We determine bounds on the Vapnik-Chervonenkis dimension and present a general node-based decomposition scheme. As to soft regression trees, we propose a model variant which provide for every input vector a single leaf node prediction, satisfying the conditional computation property. Our nonlinear optimization formulation is well-suited to decomposition and to consider both group fairness and sparsity. We investigate the universal approximation property and present a convergent node-based decomposition algorithm which includes a heuristic for the reassignment of the data points along the tree and a specific initialization procedure. Experiments on several benchmark datasets show that our model trained with the decomposition algorithm outperforms two state-of-the-art soft and deterministic regression tree models in terms of both accuracy and/or computational time.

Short Bio

Antonio Consolo is a postdoctoral researcher in Operations Research and Machine Learning at De- partment of Informatics, Systems and Communication of Universit`a degli Studi di Milano Bicocca. He received the M.S. degree in Mathematical Engineering from the Politecnico di Milano in 2019. He obtained his Ph.D. in Information Technology at the Department of Electronics, Information and Bioengineering of the Politecnico di Milano in 2024 with a thesis titled "Sparse soft decision trees and kernel logistic regression: optimization models and algorithms". His research interests mainly concern the design of nonlinear optimization algorithms and models for the training of interpretable machine learning models.