Chapter 17
Striking the Right Balance Between Greed and Force:
Is It Time to Revise Traditional Decision Tree Algorithms?

Maurice Saluschke
Carl von Ossietzky Universität Oldenburg, Germany

Michael Schulz
http://orcid.org/0000-0001-6868-788X
NORDAKADEMIE Hochschule der Wirtschaft, Germany

ABSTRACT

This article explores the integration of Brute Force (BF) methods and Greedy algorithms to develop a versatile decision tree (DT) algorithm that balances efficiency, complexity, and accuracy. The study introduces a parameterized search procedure that harnesses the strengths of both approaches, applicable to ensemble methods like AdaBoost and Random Forests (RFs). The primary research question investigates whether enhanced utilization of BF approaches can improve the quality of greedy DT algorithms without increasing their complexity. Key hypotheses address the trade-off between tree complexity and user interpretability, the potential for reduced complexity through BF methods, and the impact on computational time and accuracy. A qualitative field study supports these hypotheses, revealing that simpler DT structures are easier to interpret. Additionally, algorithmic evaluations suggest that incorporating BF methods can enhance the accuracy of greedy DTs without significantly extending computation time.

INTRODUCTION

The proliferation of smartphones and the rise of social media have led to an unprecedented increase in data volumes (Costa & Pedreira, 2023), enhancing the importance of data science and machine learning for extracting and operationalizing knowledge (Ebel et al., 2023; Saha et al., 2022).

Two principal categories of algorithms have emerged for model training: Brute Force (BF) methods and Greedy algorithms. BF methods exhaustively explore all possibilities, often resulting in inefficiency for large search spaces (Rosen, 2013; Zharmagambetov et al., 2021). Greedy algorithms, by contrast,

DOI: 10.4018/979-8-3693-5638-8.ch017

iteratively seek locally optimal solutions, achieving greater efficiency but potentially missing the global optimum – a trade-off particularly evident in decision tree (DT) construction (Williamson & Shmoys, 2011; Zharmagambetov et al., 2021).

DTs are widely used in domains where fast, transparent, and interpretable decision-making is critical. For instance, in healthcare, DTs support diagnostic decisions by providing clear, rule-based pathways that clinicians can easily follow and validate (Podgorelec et al., 2002). In financial services, they are employed for credit scoring and fraud detection, where regulatory compliance and explainability are essential (Paul et al., 2025). These applications benefit from the inherent interpretability and low latency of DTs, making them especially suitable for environments where decisions must be both rapid and justifiable. Other use cases are described in (Mienye & Jere, 2024), for example.

This paper explores a hybrid approach that combines Greedy and BF strategies to develop a flexible DT algorithm, aiming to balance efficiency, complexity, and accuracy. We introduce a parameterized search procedure that leverages both approaches and supports integration into ensemble methods such as AdaBoost and RFs. The central research question is:

> **RQ:** Can more intensive use of BF approaches improve the quality of results of greedy DT algorithms without increasing their application complexity?

Result quality is measured by classification accuracy and model complexity, where complexity refers to tree size. However, a reduction in size is only beneficial if it enhances model usability. (Rokach & Maimon, 2014)

The interpretability of DT models - an often-cited strength (Allahyari & Lavesson, 2011; Mingers, 1989) - is essential for practical application (Freitas, 2014). This raises the need to examine whether lower model complexity improves understandability.

The research question is operationalized through the following hypotheses:

> **H1:** DT with lower complexity can be interpreted more quickly by users.

In this paper, the complexity of a DT is measured by its number of levels.

> **H2:** DT with lower complexity are perceived as more understandable by users.

The verification of H1 and H2 through a qualitative field study is described. In addition to the user perspective, it is also necessary to consider the algorithmic perspective, which results in two further hypotheses:

> **H3:** DT modelled with an algorithm based on a BF approach can have lower complexity than DT modelled with an algorithm based on a greedy approach.

This hypothesis is confirmed by an example (Chapter 3.2). However, the potential benefit of models created by revised greedy DT algorithms must not have an unacceptable impact on the required computing time.

Related Content