OPTIMAL ROBUST DECISION TREES THROUGH SEARCH

Author
Kutlu Şan Demirören

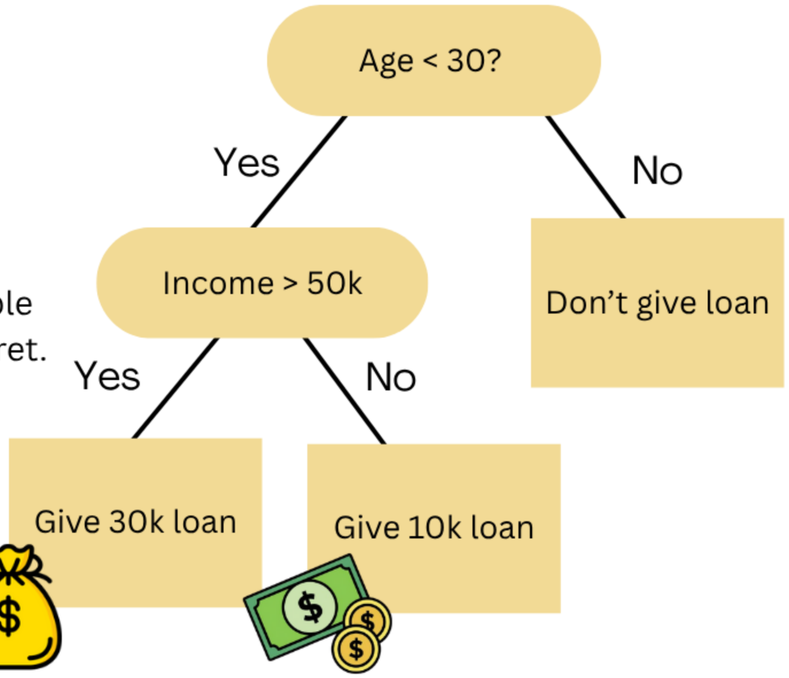
Supervisor
Ir. J.G.M. van der Linden

Responsible Professor
Dr. Emir Demirović

TU Delft

1 INTRODUCTION

Decision trees are desirable as they are easy to interpret.



...However, finding an optimal decision tree is NP Hard[1].

Name	Age	Income
Bob	30	49k

An adversary can change the data

Name	Age	Income
Bob	29	51k

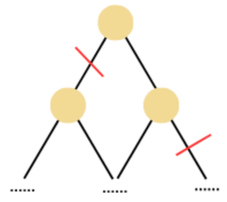
...Sadly decision trees are vulnerable (not robust) to data changes by an adversary[2].

2 MAIN QUESTION

To what extent are search-based methods more scalable than MILP methods for finding optimal robust decision trees?

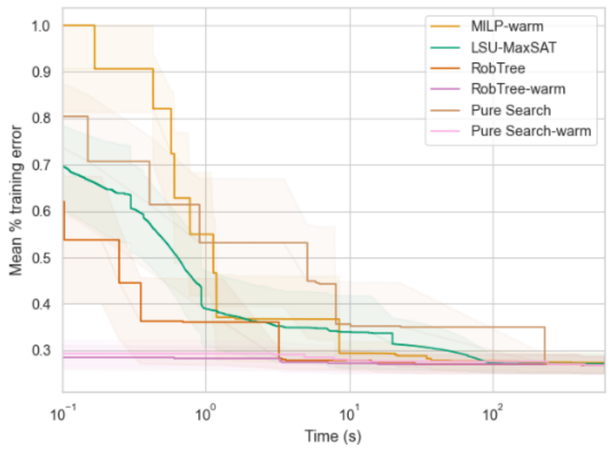
3 RESEARCH

- Looked into pre-existing methods for (robust) optimal decision trees.
- Came up with several theorems that reduce the search space
- Novel search algorithm that re-uses previous solutions



4 RESULTS

- *Up to two orders of magnitude speed-up.*
- *Converges to a close to optimal solution faster than the state-of-the-art*
- *Scales better with the number of data points than the state-of-the-art*



5 CONTACT

email: sandemiroren1@gmail.com

References

[1]Laurent Hyafil and Ronald L. Rivest. Constructing optimal binary decision trees is NP-complete. Information Processing Letters, 5(1):15–17, May 1976.

[2]Szegedy, C.; Zaremba, W.; Sutskever, I.; Bruna, J.; Erhan, D.; Goodfellow, I.; and Fergus, R. 2013. Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199