Applied AI (with Python)

Decision Trees

FH-Ass. Prof. DI Martin Uray, BSc

Salzburg University of Applied Sciences
Josef Ressel Centre for Intelligent and Secure Industrial Automation (ISIA)
Applied Data Science Lab (ADS Lab)

January 7, 2025

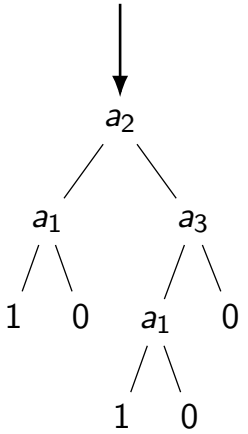
Literature

[RN21] Stuart Russell and Peter Norvig. *Artificial Intelligence, A Modern Approach*. Pearson, 2021. ISBN: 978-1-292-40113-3.

[HTF09] Trevor Hastie, Robert Tibshirani, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. Springer series in statistics. New York, NY: Springer, 2009. ISBN: 978-0-387-84857-0.

Idea

Intelligence can be captured in a set of **if-then-else** Rules that provide **branching** for classification.



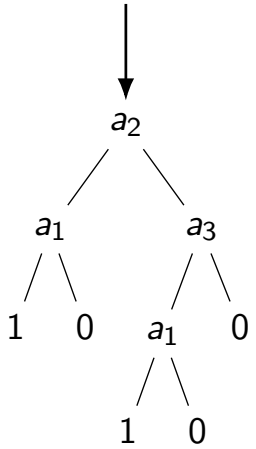
- output is discrete
- when NO large data is available
- if data is noisy
- classes are disjoint

Shannon Entropy

Entropy quantifies the (expected) number of bits to encode a class of randomly drawn samples.

Day	Outlook	Temperature	Play
1	sunny	hot	No
2	sunny	hot	No
3	overcast	hot	Yes
4	rain	mild	Yes
5	rain	cool	Yes
6	rain	cool	No
7	overcast	cool	Yes
8	sunny	mild	No
9	sunny	cool	Yes
10	rain	mild	Yes
11	sunny	mild	Yes
12	overcast	mild	Yes
13	overcast	hot	Yes
14	rain	mild	No

- min. Leaf size
- max. depth
- max. number of nodes
- min. decrease in loss
- Pruning



+ Easy to Explain
+ Interpretability
+ Categorical Variables
+ Fast
+ Missing Values

– High Variance
– bad at additive structure
– low predictive accuracy

References

[RN21] Stuart Russell and Peter Norvig. *Artificial Intelligence, A Modern Approach*. Pearson, 2021. ISBN: 978-1-292-40113-3.

[HTF09] Trevor Hastie, Robert Tibshirani, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. Springer series in statistics. New York, NY: Springer, 2009. ISBN: 978-0-387-84857-0.