

Model Tree Forests

Werner Van der Merwe

Stellenbosch University

October 24, 2020

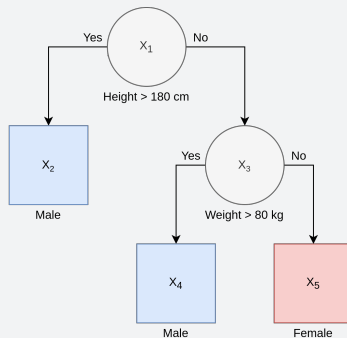
Summary

- 1 Decision trees
- 2 Ensemble learning
- 3 Proposed model tree forest (MTF)

Decision trees

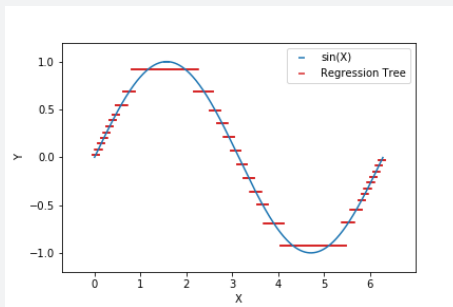
Decision trees

- Applications:
 - Classification
 - Regression
- Example:



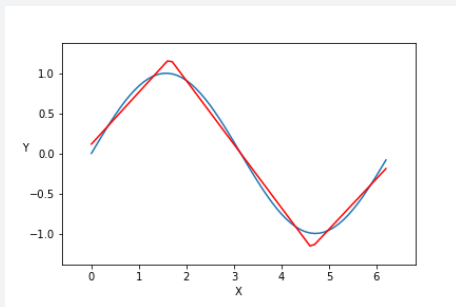
Shortcomings

- Regression trees fall short in capturing the nonlinear relationships within data.
- Regression trees are also limited to producing discontinuous outputs.



Model trees

- Model trees can produce continuous output signals.



- However, model trees only incorporate linear models.
- This is often unsatisfactory in modeling the nonlinear relationships within data.

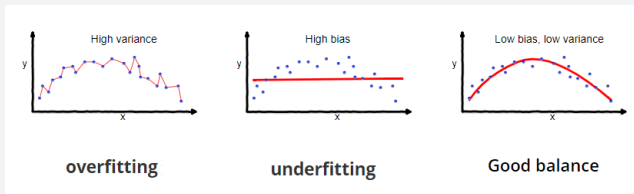
Ensemble learning

Ensemble strategies

- Types of ensemble methods:
 - Bagging
 - Boosting
- Success achieved by the random forest method.

Bias-Variance dilemma

- Each ensemble strategy addresses the bias-variance dilemma differently.



Proposed model tree forest (MTF)

Proposed model tree

- Leaf nodes:
 - GASOPE is used to produce higher order polynomials:
 - $f(\mathbf{x}) = \sum_{\tau=0, \sum_{j=1}^m \lambda_j = \tau}^n \left(w_{(\lambda_1, \lambda_2, \dots, \lambda_m)} \prod_{q=1}^m x_q^{\lambda_q} \right)$
- Employs a greedy induction approach.

Ensemble specifications

- Research different tree structures:
 - Tree stumps
 - Shallow trees
 - Deep trees
- Research effect of different ensemble sizes
- Random forest inspired ensemble method:
 - Trees induced on a subset of dataset
 - Splits chosen based on subset of input features.

The End