Tree-based methods

Wenda Zhou

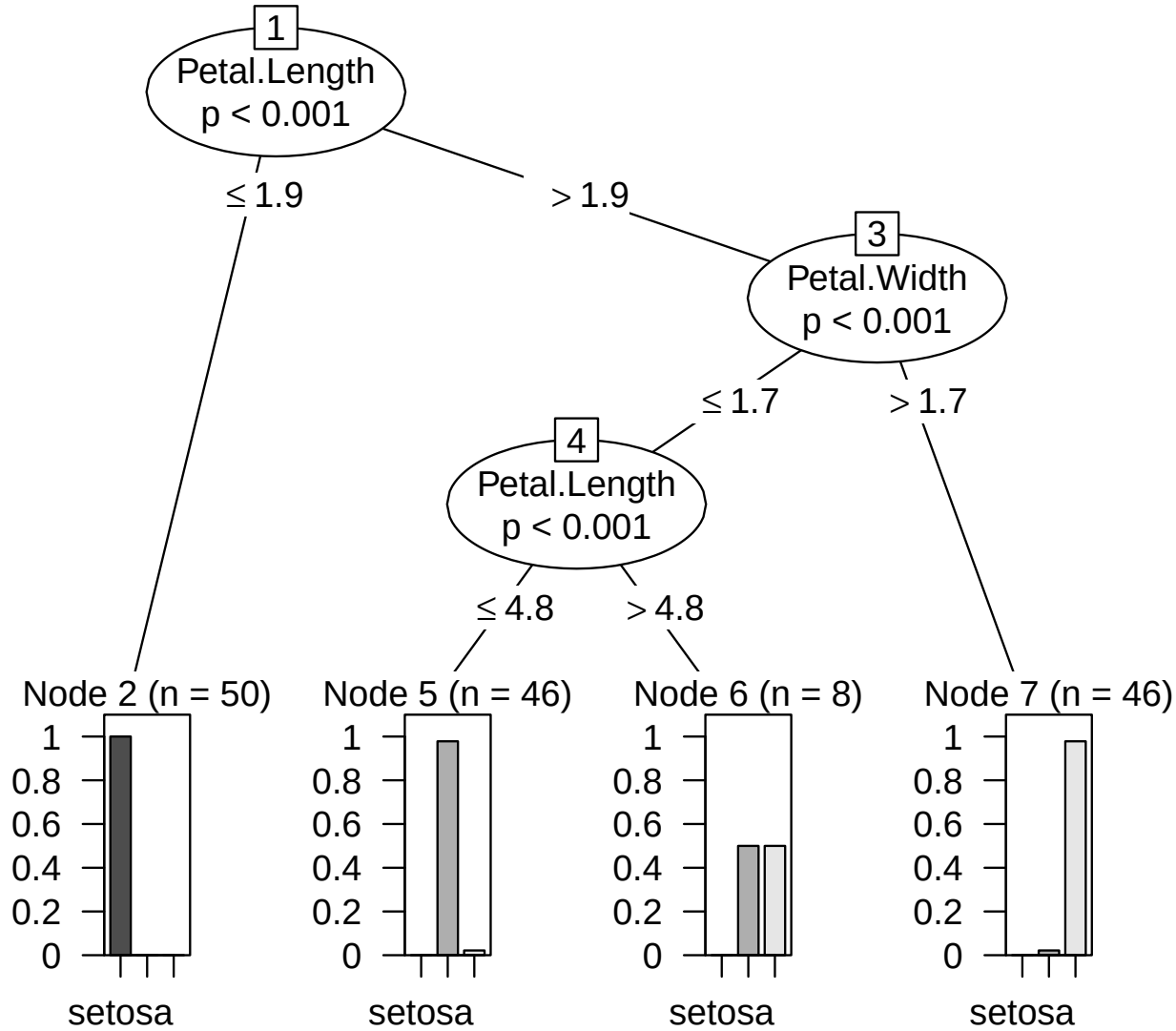
June 19, 2017

Tree-based methods

Tree-based methods centers around decision or regression trees.
Flexible and popular method for prediction.
Can be applied successfully to most prediction problems.

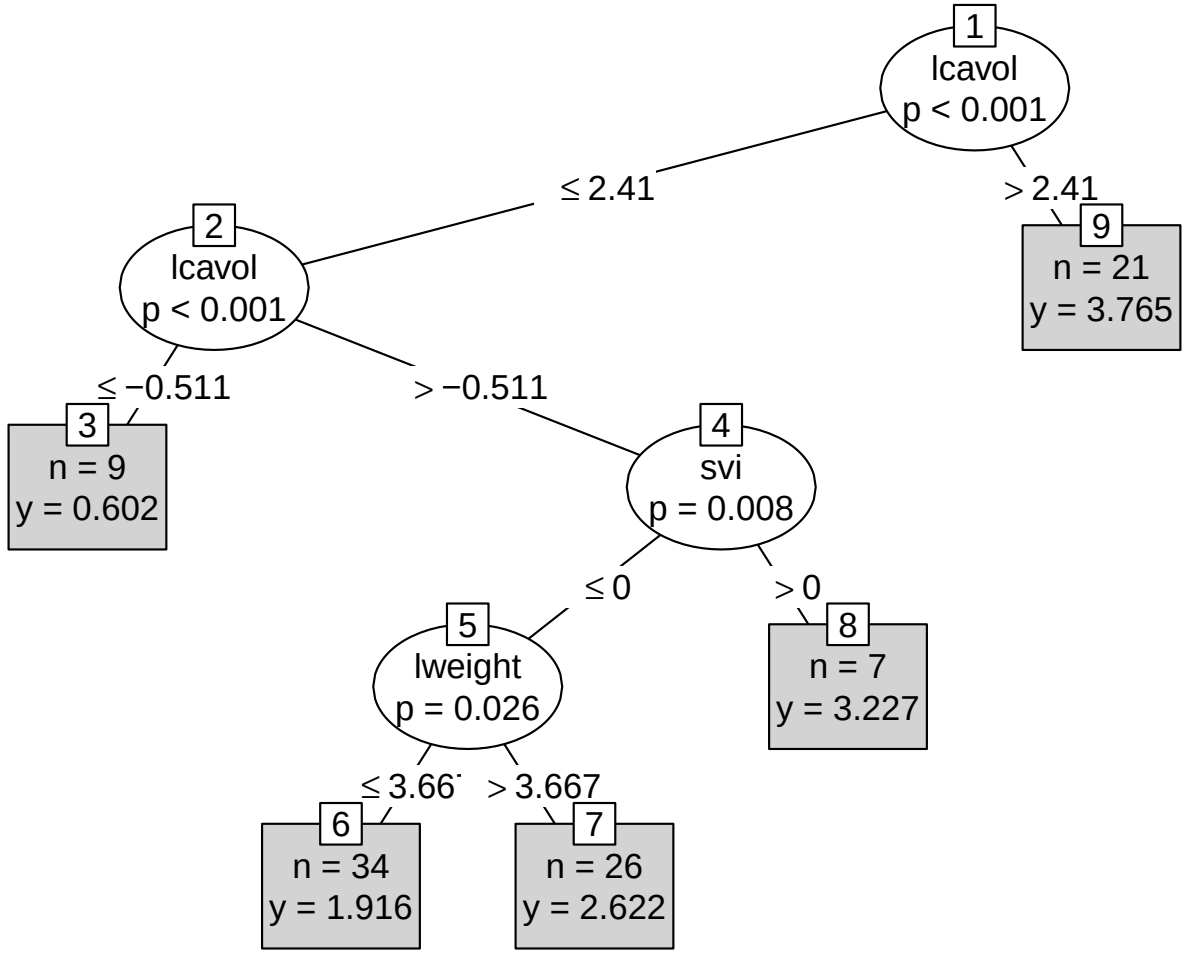
What is a decision tree?

A decision tree consists of nodes, that create a dichotomy on a given variable.



Regression trees

We can apply the same concept to create a regression tree to predict a continuous quantity.



Each leaf node gives the average value in that node.

Trees

- Nested sequence of dichotomies
- Constant among leaf nodes
- Can be viewed as a piecewise constant regression
- Adapted to categorical and ordinal data

Fitting a tree

At each node, the algorithm decides:

1. Pick variable to split on
2. Decide split location
3. Decide whether to stop or continue

Fitting a tree

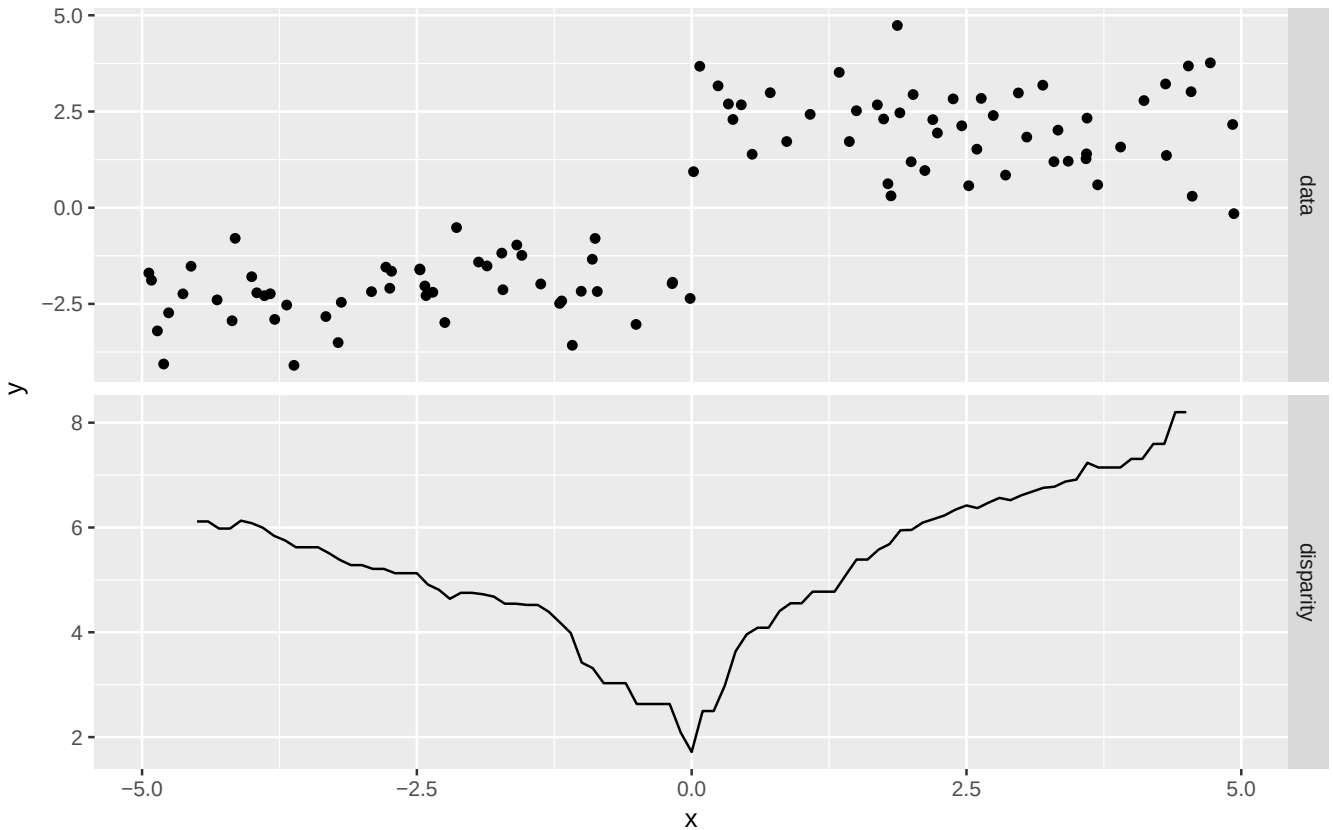
How to pick varibable and spit location?

Greedy algorithm

At each stage, pick the split that lowers disparity the most.
Does not necessarily pick optimal splits overall.

Fitting a tree

At each stage, pick the split that lowers disparity the most.



Random forests

Trees

Trees are simple and easy to understand.

However, they are somewhat limited.

Random forests

Trees

Trees are simple and easy to understand.
However, they are somewhat limited.

Random forest

- Combine multiple trees together (usually 100s).
- Each tree trained on a slightly randomised dataset
- Trees vote on prediction

Ensemble methods

General idea to combine weaker models into a stronger one.

Bagging

Bagging (= bootstrap aggregation): train a lot of models on artifical datasets, and average the models

Ensemble methods

General idea to combine weaker models into a stronger one.

Bagging

Bagging (= bootstrap aggregation): train a lot of models on artifical datasets, and average the models

Boosting

1. Train model 1, then evaluate its performance.
2. Train model 2 to do well where model 1 does not.
3. Train model 3 to do well where model 1 and 2 do not.

Ensemble methods

Pure prediction

In pure prediction problems, nearly always correct to use an ensemble method (if computational power available).

Ensemble methods

Pure prediction

In pure prediction problems, nearly always correct to use an ensemble method (if computational power available).

As ensemble methods fairly universal, usually not mentionned.