Master BeNeFri in Computer Science

Course: Statistical Learning Methods with R
Spring 2022

Exercise #9. Classification and Regression Trees

Download from the ILIAS website the `LowWeight` dataset (filename: LowWeight.txt). This dataset describes the weight of a baby given different attributes like the age of the mother, race, smoking habits etc.
Further information is provided in the file LowWeight.pdf.

1. Create a regression tree using the variable `birth_weight` as a target. Plot the resulting model.

2. Calculate the train and test MSE. Describe the results you obtained.

3. Should your regression tree be pruned? If yes, which strategy would you use? Compare the previous test MSE with the one obtained with the pruned tree. Plot the new model.

Download from the ILIAS website the `Heart` dataset (filename: Heart.txt). This dataset describes the attributes influencing the insurgence of heart diseases. Further information is provided in the file Heart.pdf.

4. Create a classification tree using the variable `disease` as a target. Plot the resulting model.

5. Compute the confusion matrix for your model and calculate the accuracy, sensitivity and specificity. Describe the results you obtained.

6. Should your classification tree be pruned? If yes, which strategy would you use? Compare the previous results with the one obtained with the pruned tree. Plot the new model.