

## Random Forest Iris

- Iris finns i R. Det är bara att kalla på iris så får man datasetet
- Fit ett Classification Tree till iris datan och spara det i en variabel *decision\_tree*. Du kan använda `tree()`
- Skriva ut *decision\_tree*. Vad tror du detta berättar?
- Använd `plot(decision_tree)` efterföljd av `text(decision_tree)`. Blir det lite lättare att förstå vad som skrevs ut i föregående punkt?
- Använd `ggpairs()` på iris datan. Vilka slutsatser drar du?
- Dela datan in i träning och test data och gör en RandomForest på iris och spara det i en variabel *iris\_classifier*.
- Skriva ut *iris\_classifier*. Vad tror du detta berättar?
- Gör en `plot(iris_classifier)`. Tolka resultatet. Hur många träd tror du är tillräckligt att använda?
- Ta reda på betydelsen till de olika variablerna genom att använda `importance()` och `varImpPlot()`. Vilka variabler borde vara viktigast att inkludera?
- Gör en scatter plot av petal width mot petal length och en annan scatter plot av sepal width mot sepal length. Färja datapunkterna efter vilken art det är. Görs enkelt genom att ange `color=Species`. Ser du några klustre?
- Prediktera med att använda `predict()`. Gör en confusion matrix. Verkar Random Forest fungerat bra? Hur är den jämfört med `tree()`?