Goals
In this tutorial you should implement parts of the decision tree learning method and apply this algorithm to data. The goal is a deeper understanding of this learning process, the functioning of decision trees and their advantages and disadvantages.

Problem Setting
We have got information $X_{tr} = \{x_i\}$ on computed tomography images of 80 heart attack patients (see spect.mat). Each of these images is marked as conspicuous ($y_i = 1$) or inconspicuous ($y_i = 0$). The data consists of 22 binary attributes, ie 22 different investigations. The goal is to make a prediction about the label for an additional set of 187 tomography images $X_{test}$.

Task 1
Write the following MATLAB function

```matlab
function IG = infoGain(y1,y2)
```

which computes the information gain for the binary split $y_1$ and $y_2$. Here $y_1$ and $y_2$ are vectors with the binary class labels of the examples in the child node (after the split). Use the formula from the lecture slides.

Task 2
Analyze the following MATLAB function

```matlab
function Tree = learnTree(X,y,k,f_splitCriterium)
```

which learns and returns a decision tree using ID3 for a given $n \times m$ - data matrix $X$ ($n$ $m$-dimensional binary row vectors) and binary class labels $y$. The function for determining the optimum split is given by a function handle $f\_splitCriterium$. 
Task 3
Write the following MATLAB function

function y = classTree(Tree,X)

which returns for a given previously learned decision tree Tree and a $n \times m$ data matrix $X$ the $n$-dimensional class label vector $y$. Classify the test data $X_{\text{Test}}$ with the help of classTree. Compare the predicted label with the ground truth.

Task 4
Execute the MATLAB script demoTree. Examine the impact of the hyper parameter $k$ and the used split function on the prediction result.

Additional Task
Write the following MATLAB function

function IG = infoGainRatio(y1,y2)

which computes the information gain ratio for the binary split $y_1$ and $y_2$. Here $y_1$ and $y_2$ are vectors with the binary class labels of the examples in the child node (after the split). Use the formula from the lecture slides.