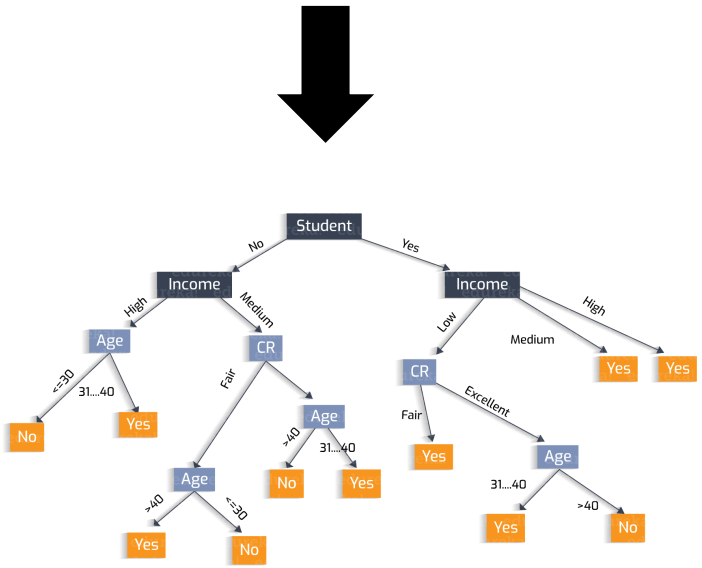
Decision Trees: Learning

- **Learning Algorithm:** MakeSubtree(set of training instances D)

$\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})\}$



MakeSubtree(set of training instances D)

C = **DetermineCandidateSplits**(D)

if **stopping criteria** is met

make a leaf node N

determine class label for N

else

make an internal node N

S = **FindBestSplit**(D, C)

for each group k of S

D_k = subset of training data in group k

k^{th} child of N = MakeSubtree(D_k)

return subtree rooted at N