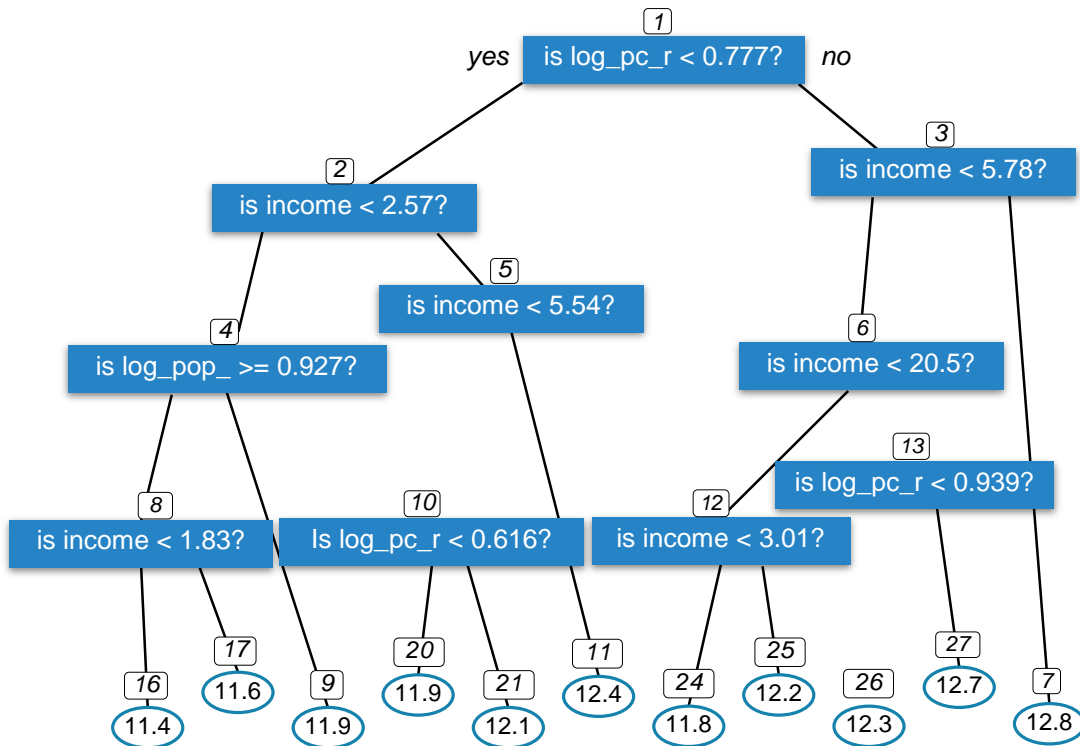




Trees Forests

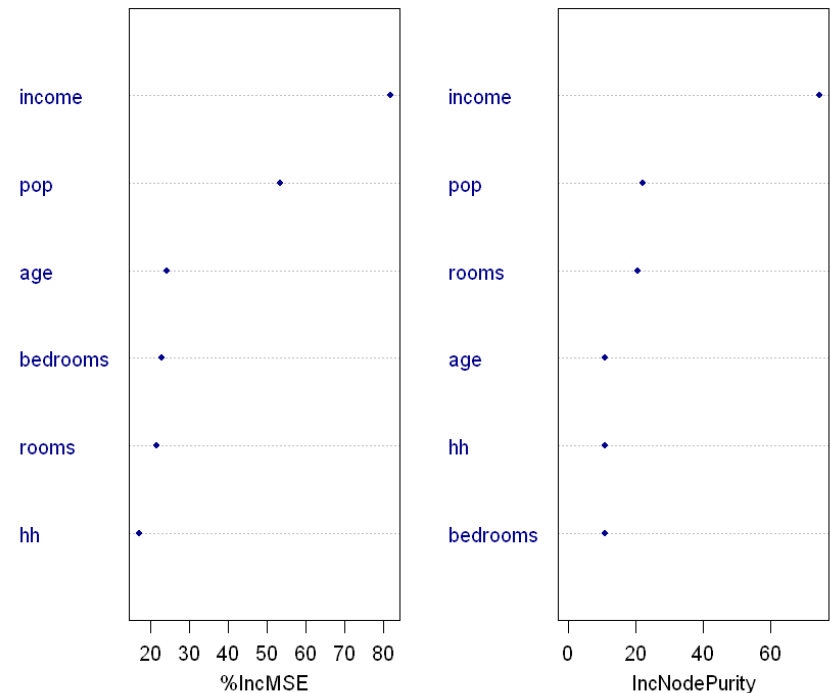
School of Information Studies
Syracuse University

Picking a Tree in the Forest



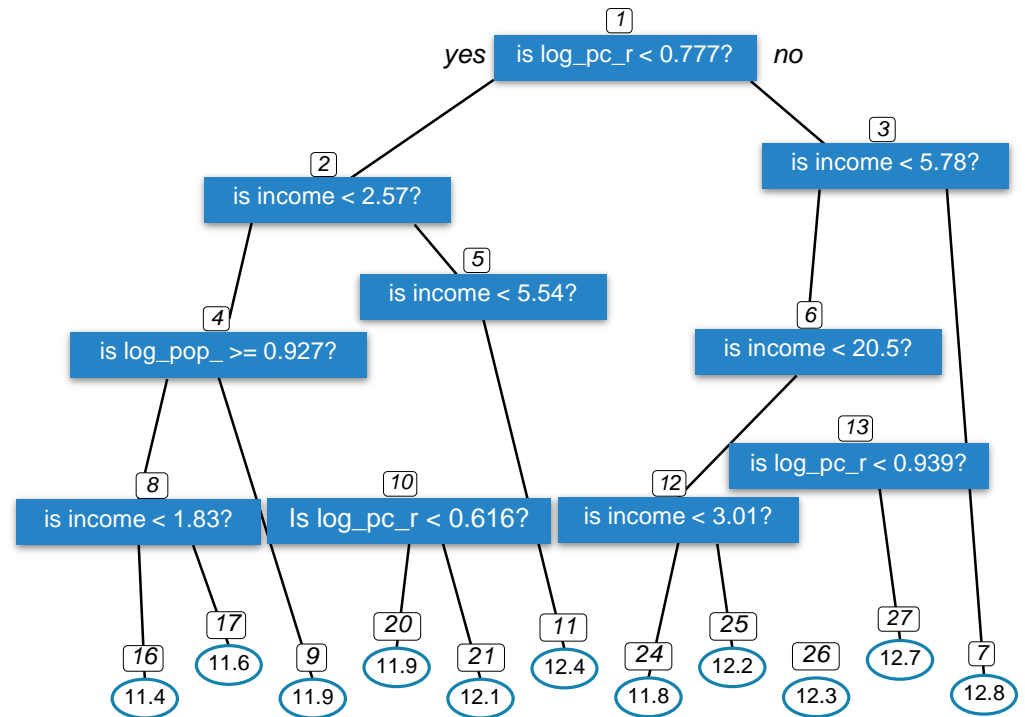
Decision Trees

- Key advantage is interpretability
- Partition the space into simple regions to achieve best fit
- Pruning methods control the size of the tree



Random Forests

- Ensemble method using multiple decision trees
- Recursive partitioning on the training set
- Effective with large number of explanatory variables



Random Forests (cont.)

- Provides interpretability through use of one tree from set
- Significant difference in performance between train and test indicates overfitting
- Individual explanatory variables can still be inferred

