

Considerations for ID3

Relevant Readings: Section 3.7 in Mitchell

CS495 - Machine Learning, Fall 2009

Pruning

- ▶ One way deal with overfitting is to prune the tree produced by ID3

Pruning

- ▶ One way deal with overfitting is to prune the tree produced by ID3
- ▶ To *prune* a node:

Pruning

- ▶ One way deal with overfitting is to prune the tree produced by ID3
- ▶ To *prune* a node:
 - ▶ Remove the subtree rooted at that node

Pruning

- ▶ One way deal with overfitting is to prune the tree produced by ID3
- ▶ To *prune* a node:
 - ▶ Remove the subtree rooted at that node
 - ▶ Replacing it with a leaf node that takes the majority vote of training instances associated with it

Pruning

- ▶ One way deal with overfitting is to prune the tree produced by ID3
- ▶ To *prune* a node:
 - ▶ Remove the subtree rooted at that node
 - ▶ Replacing it with a leaf node that takes the majority vote of training instances associated with it

Reduced error pruning

- ▶ First, grow the tree completely (as in ID3)

Reduced error pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then, keep pruning the node which causes the most improvement on the tuning set when pruned

Reduced error pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then, keep pruning the node which causes the most improvement on the tuning set when pruned
- ▶ If any possible further pruning causes accuracy on tuning set to decrease, stop.

Reduced error pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then, keep pruning the node which causes the most improvement on the tuning set when pruned
- ▶ If any possible further pruning causes accuracy on tuning set to decrease, stop.

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:
 - ▶ Make a copy of the original tree

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:
 - ▶ Make a copy of the original tree
 - ▶ Choose a random number of nodes to prune

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:
 - ▶ Make a copy of the original tree
 - ▶ Choose a random number of nodes to prune
 - ▶ Prune that many number of nodes, uniformly at random

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:
 - ▶ Make a copy of the original tree
 - ▶ Choose a random number of nodes to prune
 - ▶ Prune that many number of nodes, uniformly at random
 - ▶ If this gets the best accuracy so far on the tuning set, remember it

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:
 - ▶ Make a copy of the original tree
 - ▶ Choose a random number of nodes to prune
 - ▶ Prune that many number of nodes, uniformly at random
 - ▶ If this gets the best accuracy so far on the tuning set, remember it
- ▶ Return the best tree found (as measured on the tuning set)

Stochastic pruning

- ▶ First, grow the tree completely (as in ID3)
- ▶ Then for some large number of trials:
 - ▶ Make a copy of the original tree
 - ▶ Choose a random number of nodes to prune
 - ▶ Prune that many number of nodes, uniformly at random
 - ▶ If this gets the best accuracy so far on the tuning set, remember it
- ▶ Return the best tree found (as measured on the tuning set)

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?
- ▶ One easy way:

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?
- ▶ One easy way:
 - ▶ Split the attribute by the mean or the median

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?
- ▶ One easy way:
 - ▶ Split the attribute by the mean or the median
- ▶ A better way:

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?
- ▶ One easy way:
 - ▶ Split the attribute by the mean or the median
- ▶ A better way:
 - ▶ Sort the instances by the attribute we are trying out

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?
- ▶ One easy way:
 - ▶ Split the attribute by the mean or the median
- ▶ A better way:
 - ▶ Sort the instances by the attribute we are trying out
 - ▶ Try out every midpoint between these consecutive values, and take the best (in terms of information gain)

Dealing with continuous attributes

- ▶ How do we deal with continuous attributes?
- ▶ One easy way:
 - ▶ Split the attribute by the mean or the median
- ▶ A better way:
 - ▶ Sort the instances by the attribute we are trying out
 - ▶ Try out every midpoint between these consecutive values, and take the best (in terms of information gain)

Example

- ▶ Do a complete example of ID3 with continuous features

Example

- ▶ Do a complete example of ID3 with continuous features