

## Decision Tree with Binary Response Interactive Training

Decision tree splits are selected on the basis of an analytic criterion automatically. Sometimes, it is necessary or desirable to select splits on the basis of a practical business criterion. For example, the best split for a particular node may be on an input that is difficult or expensive to obtain. If a competing split on an alternative input has a similar worth and is cheaper and easier to obtain, it makes sense to use the alternative input for the split at that node.

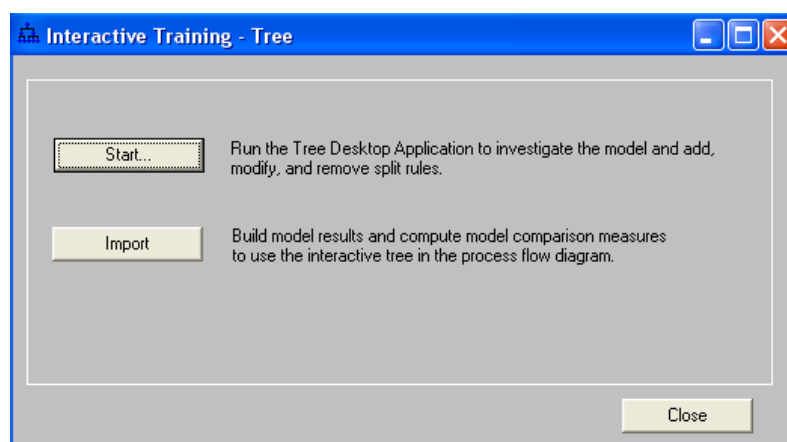
Likewise, splits may be selected that are statistically optimal but may be in conflict with an existing business practice. For example, the credit department may treat application where debt-to-income ratios are not available differently from those where this information is available. You can incorporate this type of business rule into your decision tree using interactive training in the Tree node. It might then be interesting to compare the statistical results of the original tree with the changed tree. In order to accomplish this, first make a copy of the Default Tree node.

Interactive training provides more control over how the tree is constructed. You use interactive training in order to

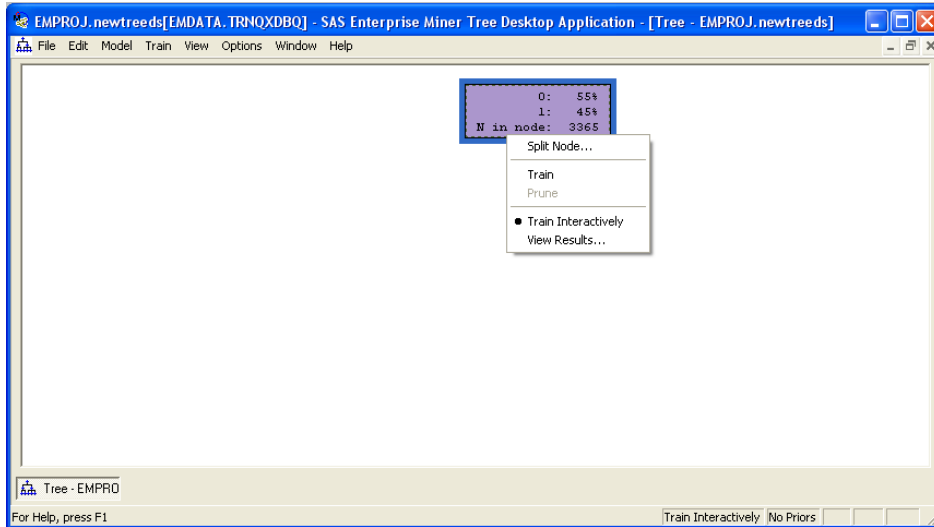
- force a particular variable(s) into the model
- force a variable to be split in a particular way
- prune a tree that has already been constructed
- continue to train the remainder of a tree automatically.

The Enterprise Miner Tree Desktop Application uses the entire training data for interactive training.

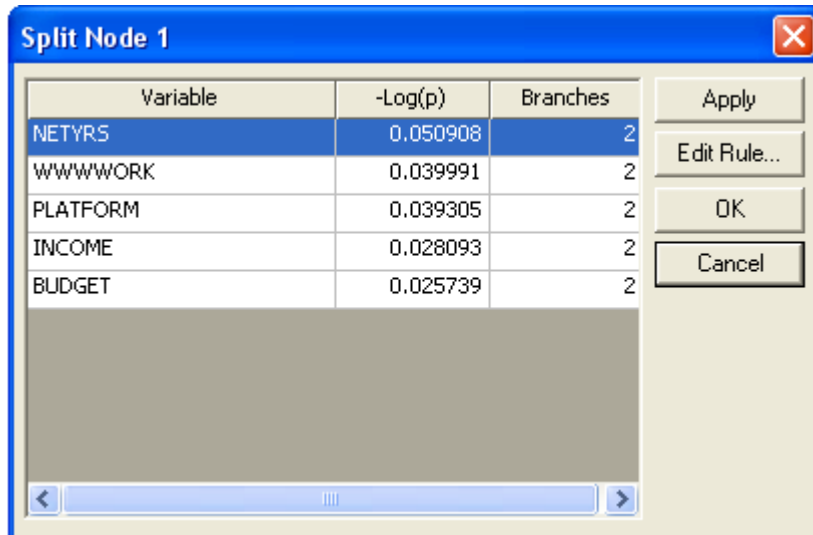
- Add a Tree node and connect it to the Data Partition node and the Assessment node. Rename this node **Interactive**.
- Right-click on the **Interactive** tree node and select **Interactive**. The Interactive Training window opens.



- Select **Start....** Then the **Tree Desktop Application** window opens.

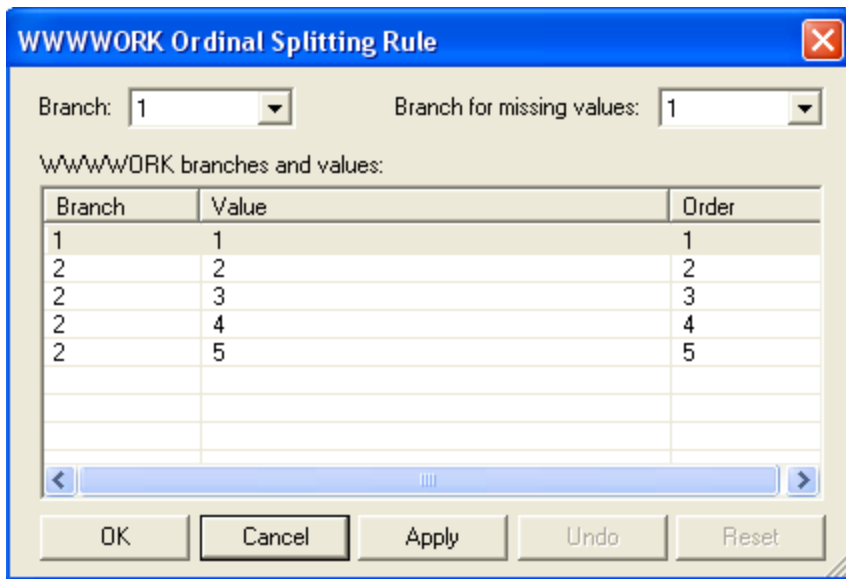


- Right-click the root node and select **Split Node....** The Create Rule window opens, listing the maximum Kass adjusted logworths for each input in the training data.

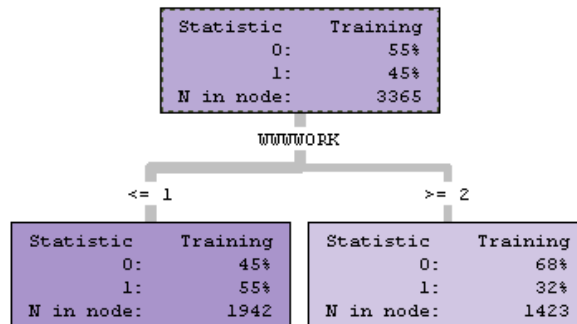


You can see that the best split is on NETYRS – years on internet (<0.5, 0.5-1, 1-3, 4-6, > 6). Suppose that we would like to see a split according to WWWORK, which has values 1-5 representing whether the respondent can access WWW from work daily, weekly, monthly, <once/month, and never. We want to separate people who have no restriction in WWW access from work from those who have. Namely, we want to see the split – WWWORK ≤ 1.

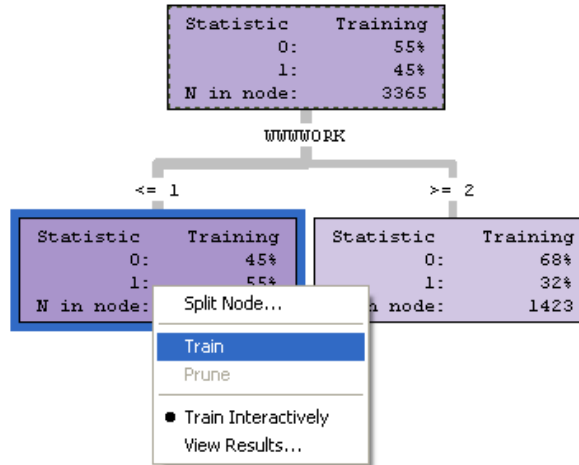
- Select the **WWWWORK** row and click on **Edit Rule...** Then the Edit Rule window opens. Since  $WWWWORK \leq 1$  is the default rule, no further modification is needed. Click on Apply.



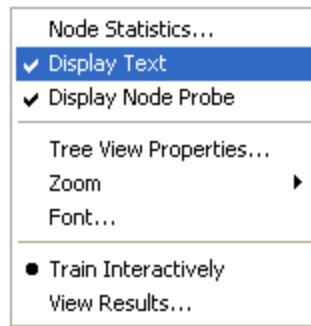
- Click on Apply. Then you can see the root node is split into two children node according to the rule that you defined.



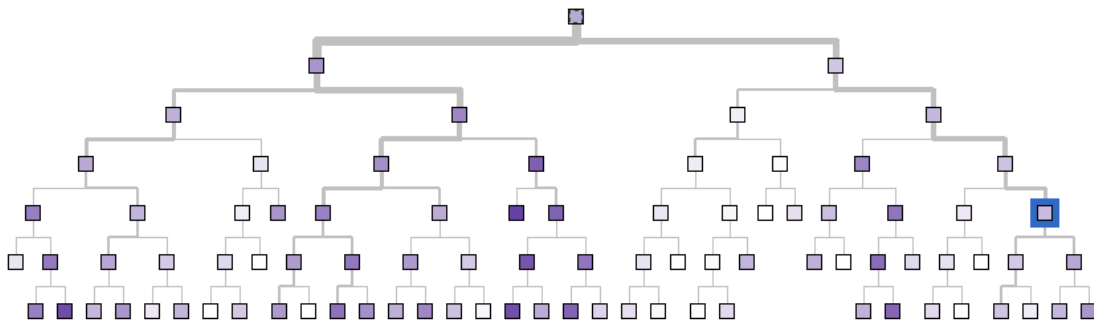
- Right Click on either child node and select **Train**. Then EM will automatically train a tree model for either child node. And you will see a massive tree structure displayed.



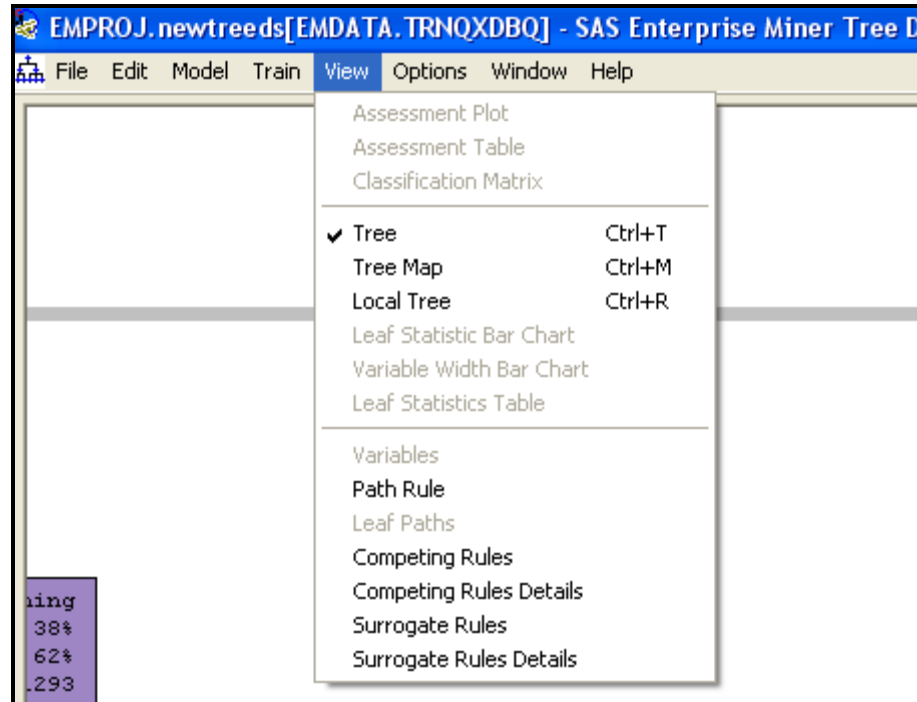
- To view the whole tree structure, right click and uncheck **the Display Text** option.



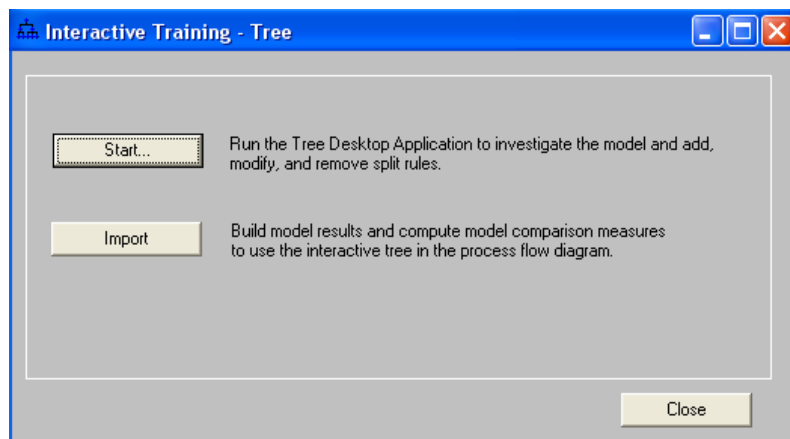
- Now the whole tree structure – **Tree Map** is displayed. EM allows you to examine each node and continue splitting, pruning, or training interactively.



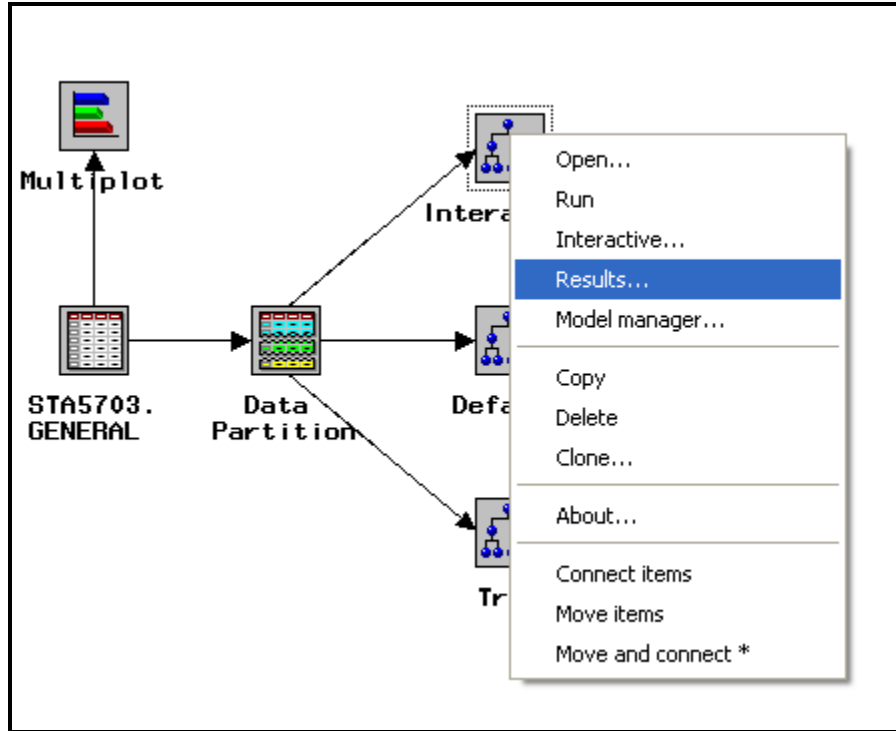
- You may also select the **View Tree Properties ...** and explore the options and features. Besides, you may select **View** option and check out Path Rule, Competing Rules, and Surrogate Rules.



- When you are done, close the **Tree Desktop Application** window. And then select **Import** in the **Interactive Training - Tree** window and close it by selecting **Close**.



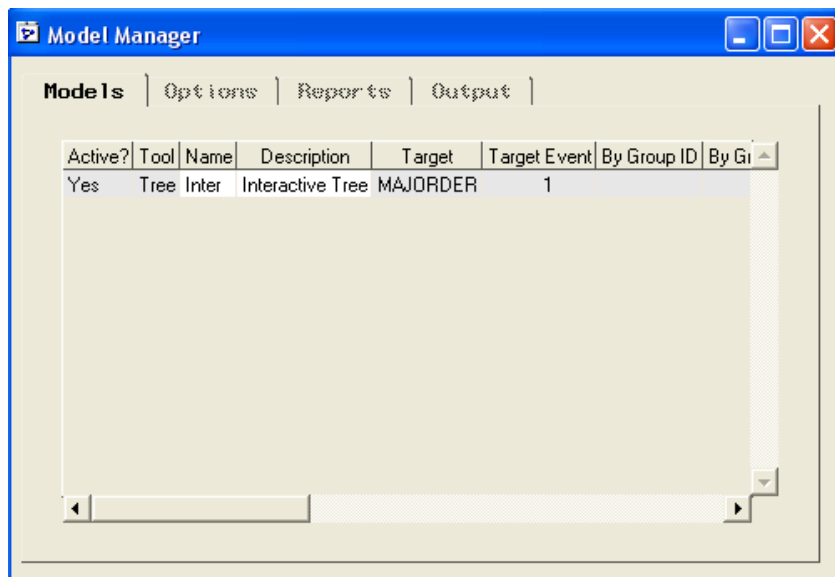
- Right click on the Tree node and Select **Results....** You may view the tree results from the interactively trained tree model.



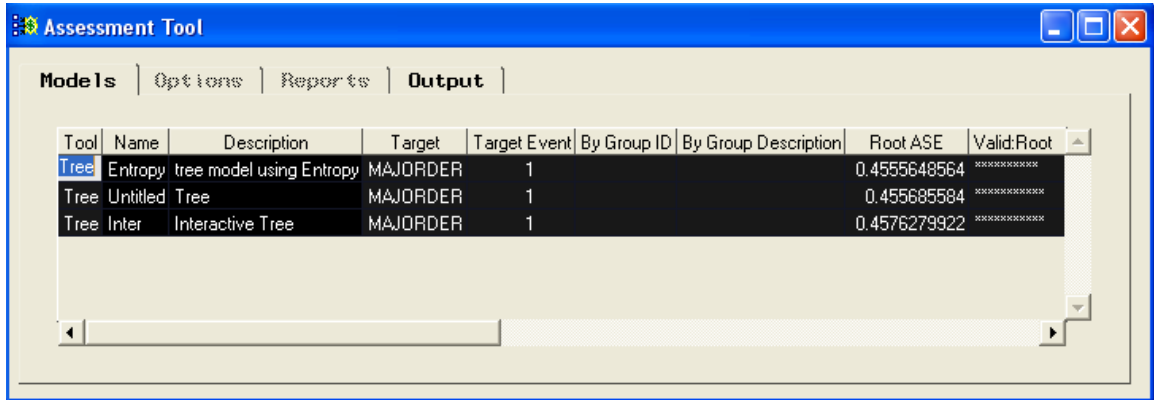
- You may further modify the tree structure interactively by selecting **Interactive** and then **Start** in the **Interactive Training – Tree** window.

Next, we want to compare the tree models that we have built. Before we do so, we want to describe each model more clearly.

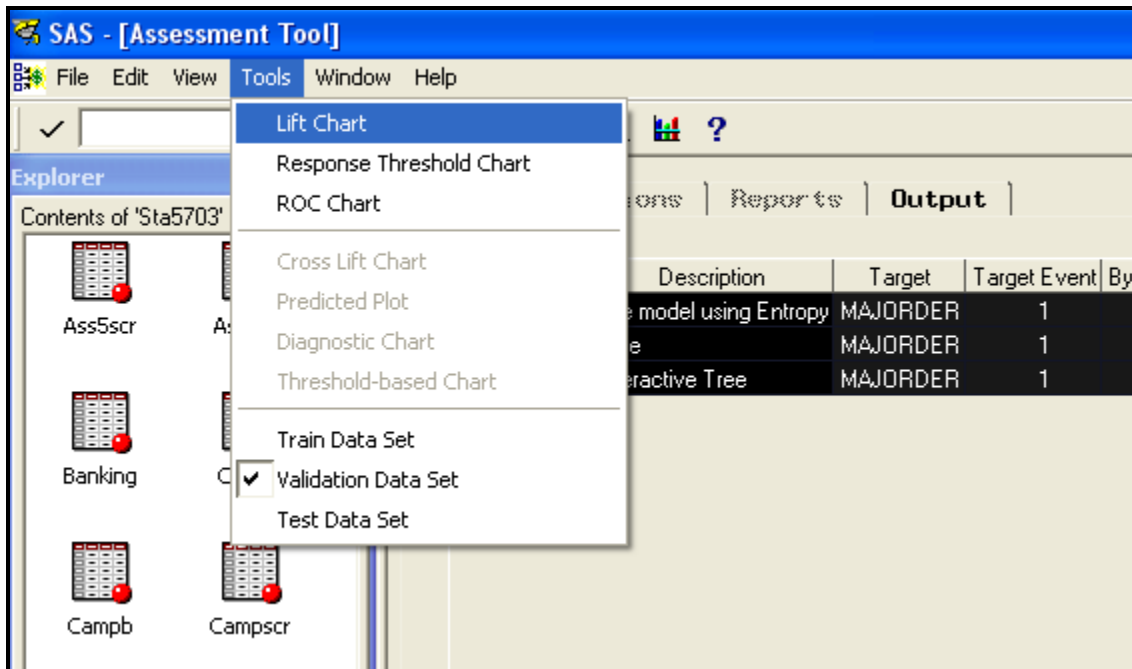
- Right click and then select **Model Manager....** You may rename the model and add Descriptions.



- Add an **Assessment** node and connect it to every tree node. Right click and select **Run**. Then select **Yes** when a window is prompted to ask whether to View Results.



- Highlight all tree models. Click on **Tools** and select **Lift Chart**.



- Click on **Format** and select **Model Name**. Then you may compare the tree models.

