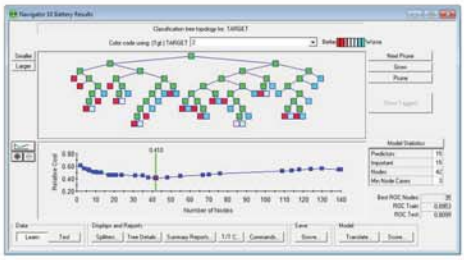
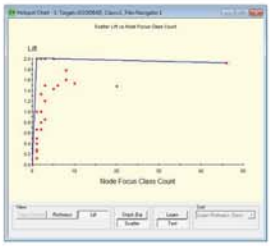
CART Classification and Regression Trees

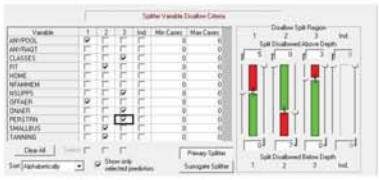
CART is the ultimate classification tree that has revolutionized the entire field of advanced analytics and inaugurated the current era of data mining. CART, which is continually being improved, is the most important tool in modern data mining methods. Others have tried to copy CART but no one has succeeded as evidenced by accuracy, performance, feature set automation and ease of use.



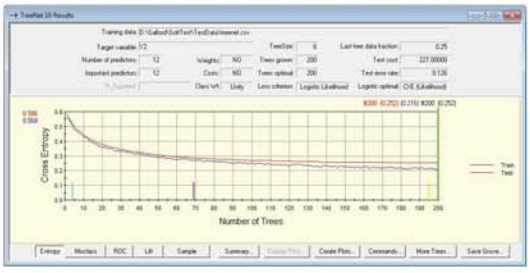
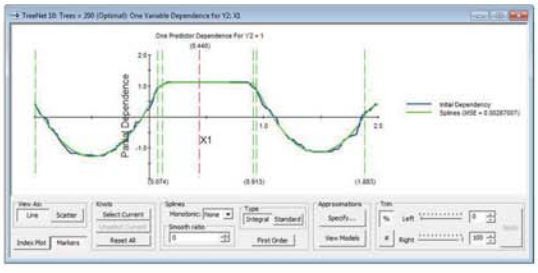


CART's Hotspot Detector ™ is specifically designed to search many trees to find nodes of ultra-high response. For example, in a fraud detection problem, you might be interested in identifying a set of rules that lead to a high ratio of fraudulent transactions, allowing you to flag future records that fit those rules as almost certainly fraudulent.

CART's Train/Test Consistency Report provides graphical and tabular reports to summarize train/test agreement: how well do train and test match up across all nodes. Classic CART trees are evaluated on the basis of overall tree performance. However, many users of CART are more interested in the performance of specific nodes and the degree to which terminal nodes exhibit strongly consistent results across the train and test samples.



Advanced Tree Controls: CART users can dictate the splitting variable to be used in the root node or in either of the two child nodes of the root, allowing the user to impose structure on a tree. More specific controls allow the user to specify the split values for both continuous and categorical variables.

TREENET Stochastic Gradient Boosting

TreeNet is Salford's most flexible and powerful data mining tool, capable of consistently generating extremely accurate models. If you need accuracy and automation, you cannot do better than TreeNet. TreeNet has been responsible for at least a dozen prizes in major data mining competitions since its introduction in 2002. TreeNet demonstrates remarkable performance for both regression and classification and can work with varying sizes of data sets, from small to huge, while readily managing a large number of columns. The algorithm typically generates thousands of small decision trees built in a sequential error-correcting process to converge to an accurate model. TreeNet models are usually complex and thus the software generates a number of special reports designed to extract the meaning of the model, such as a ranking of the variables in order of importance and graphs illustrating the relationship between inputs and outputs.