3. CART can use the same variables more than once in different parts of the tree. This capability can uncover complex interdependencies between sets of variables.
4. CART can be used in conjunction with other prediction methods to select the input set of variables.
5. CART can be incorporated into hybrid models, where CART feeds inputs to a neural network model (which itself cannot select variables).

**Uses of CART**

1. **CART is simple!**
2. **Data preparation.** Classical statistical models require that the analyst has a clear understanding of the nature of the function inherent in the data to be modeled. CART requires very little input for the beginning data miner.
3. **Variable selection.** CART can be used to create the short-list of predictor variables to submit to the modeling algorithm. There is no guarantee that the variables most useful for a tree will also prove most useful for a neural network or other function, but in practice this is a useful technique.
4. The use of predictors multiple times in the tree helps to detect complex interactions in the data.
5. CART can handle missing values by identifying surrogate (alternate) splitting rules. During training, after the best split is found for a node, new splits using other variables are scored according to their similarity in distributing the data to the left and right child nodes. The best five or so are then stored as backup or surrogate questions to ask should the main variable not be available.

**GENERAL CHAID MODELS**

CHAID is an acronym for Chi-Square Automatic Interaction Detector. CHAID differs from CART in that it allows multiple splits on a variable. For classification problems, it relies on the Chi-square test to determine the best split at each step. For regression problems (with a continuous target variable), it uses the F-test.

Key elements of the CHAID process are

1. **Preparing the predictor variables:** Continuous variables are “binned” to create a set of categories, where each category is a subrange along the entire range of the variable. This binning operation permits CHAID to accept both categorical and continuous inputs, although it internally works only with categorical variables.
2. **Merging categories:** The categories of each variable are analyzed to determine which ones can be merged safely to reduce the number of categories.
3. **Selecting the best split:** The algorithm searches for the split point with the smallest adjusted $p$-value (probability value, which can be related to significance).