to the corner point in this form of the lift chart.) If you are interested in how the top quarter of the cases perform, the chart above says they provide over 45% of the return.

To build a lift chart using a decision tree, follow these steps:

1. Build the tree from the training data.
2. Order its leaf nodes by the response rate in training order\(^6\)(best first).
3. Run the model on the evaluation data, scoring each leaf node.
4. Graph evaluation response % versus solicited % using the training order of the leaf nodes from step 2.

If all nodes are well ranked, the slope of the line between successive knots (moving left to right) will decrease slightly. Note that the model represented by Figure 13.3 has some nodes whose ranking does not hold up on evaluation, as there are “dents” in the lift curve. A similar chart from a more continuous type of model, such as linear regression, will likely have many more knots, and it is possible (if the inputs are varied enough) that every predicted value will be distinct. This will make the lift curve smooth, though it can still be dented. Smoothness allows for much more variable cutoffs than can easily be had with the few knots of a tree’s lift curve.

Cross-Validation to Estimate Error Rate and Its Confidence

What if, after splitting the data into training and evaluation, we are concerned that we just got lucky (or unlucky) in how the model performed? The evaluation result is just one number. How confident in it can we be? The best way to find out is to split the data multiple times and validate the original number. If we organize our data bookkeeping in such a way as to split the data \(V\) different times so each case is in the evaluation data exactly once, we have \(V\)-fold cross-validation, as described here and shown in Figure 13.4:

1. Separate the data into \(V\) subsets of equal size (stratifying if necessary). A rule of thumb is \(V = 5\) or \(10\).
2. Train \(V\) models, leaving out a different data subset for each.\(^7\) The training data have proportionate size \((V - 1)/V\).
3. Test each model on the data held out for it (the gray blocks of Figure 13.4).
4. Accumulate these test results to get a distribution of out-of-sample accuracies.

With cross-validation (CV), an analyst gets a distribution of error results, which is a better estimate of true accuracy in two ways:

1. The mean is more accurate than a single experiment; and
2. The spread of the distribution provides an idea of the confidence in that mean value.

\(^6\) We found a serious error in a widely disseminated tool from a major vendor—which the vendor fixed immediately upon our telling it privately—where it had instead ordered the nodes of the evaluation tree by its own ranking, rather than using the ranking set from training. This look-ahead error led all users to think they had better results than they really did!

\(^7\) In the limit where \(V = N\), the number of cases, you get the famous “leave-one-out” estimation method.