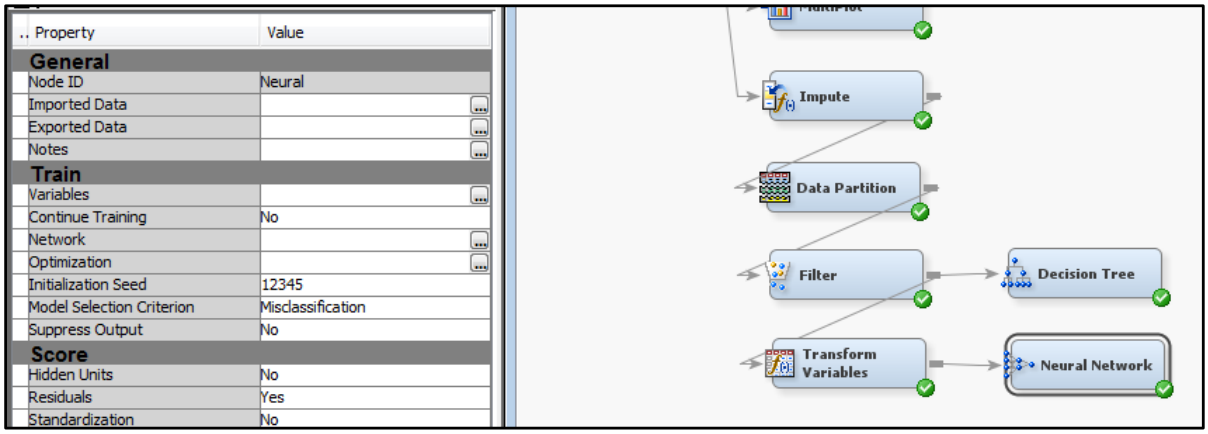
In the final assessment step, you will employ the Model Comparison node to compare all three models, using both gains tables and a cumulative-lift graph.

Neural Network

As described in Chapter 3, neural networks are models that capture many of the nonlinear relationships in data. So, although neural network models sometimes do a better job of fitting the model than the decision tree or regression models, they might not perform as well on the validation data or maintain their ability to predict as accurately over time. Another disadvantage of neural network models is that they are more difficult for marketers to interpret than the decision tree models.

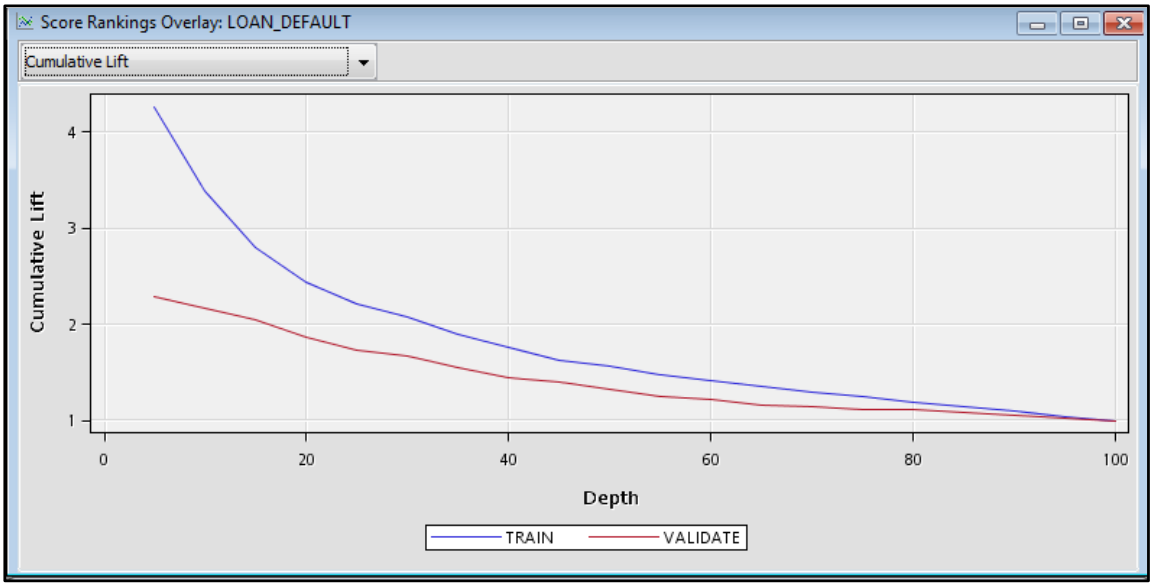
Under the **Model** tab, drag the **Neural Network** icon (fifth from the right) onto the diagram and connect it to the **Transform Variables** node. Figure 9.14 shows the diagram and the **Neural Network Property** window. In the **Train** section, next to **Model Selection Criterion**, change **Value** to **Misclassification.** Right-click the model node, and click **Run ▶ Yes.** When the window appears, click **Results.**

Figure 9.14: Neural Network Property Window



As shown in Output 9.11, the cumulative lift is much higher for the neural network train data than for the validation data.

Output 9.11: Neural Network Cumulative Lift



The **Train** data appears to capture more than 45% of the defaults in the top 10% of the file. The **Validation** data captures only a little more than 20% of the defaults in the top 10% of the file. As with the Decision Tree mode, this difference in lift between the train and validation data implies that the model is not robust. In other words, it is overfitting the train data and probably won't predict as well as expected when the model is implemented. You will compare the results of the decision tree model with the results of the neural network and regression models in an upcoming section.

Regression

As described in Chapter 3, regression models use a series of independent, or predictive, variables to predict an outcome. In predictive modeling, you can use logistic regression, which builds a model with a binary outcome. The resulting equation assigns a probability of the event's occurring. In the current case, the binary outcome or event is **Loan Default**.

Under the **Model** tab, drag the **Regression** icon (third from the right) onto the diagram, and connect it to the **Transform Variables** node. Figure 9.15 shows the diagram and the **Regression Property** window. Because the target variable, LOAN_DEFAULT, is binary, the **Regression Type** defaults to **Logistic**.

As mentioned in Chapter 3, each selection option has pros and cons. For this model, change **Selection Model** to **Stepwise**. Change **Use Selection Defaults** to **No** (Figure 9.15).