Note that $V$-fold cross-validation is a time-consuming procedure that will validate each tree in the tree sequence several times; however, a significant benefit is that the program will now pick for you the best tree, i.e., the one with the best cross-validation cost and node complexity trade-off.

After all nodes are updated, double-click on the (now single) results node to review all results in the workbook. Note that both the Descriptive Statistics as well as the results of the Standard Classification Trees analysis are displayed in the same workbook.

7. **Deploy solution (models) for new data.**

**STATISTICA** Data Miner includes a complete deployment engine for Data Miner solutions that comprises various tools. For example,

- You can create Visual Basic or C/C++/C# program code in most interactive analysis modules that will compute predictions, predicted classifications, and clusters assignments (such as General Regression Models, Generalized Linear Models, General Discriminant Function Analysis, General Classification and Regression Trees (GC&RT), Generalized EM & $k$-Means Cluster Analysis, etc.).

- You can create XML-syntax-based PMML files with deployment information in most interactive modules that will compute predictions, predicted classifications, or cluster assignments (i.e., the same modules mentioned in the preceding paragraph). One or more PMML files with deployment information based on trained models can be loaded by the Rapid Deployment of Predictive Models modules to compute predictions or predicted classifications (and related summary statistics) in a single pass through the data; hence, this method is extremely fast and efficient for scoring (predicting or classifying) large numbers of new observations.

- General Classification and Regression Trees and General CHAID modules can be used to create SQL query code to retrieve observations classified to particular nodes or to assign observations to a node (i.e., to write the node assignments back into the database).

- Complex neural networks and neural network ensembles (sets of different neural network architectures producing an average or weighted predicted response or classification) can also be saved in binary form and later applied to new data.

In addition, **STATISTICA** Data Miner contains various designated procedures in the (Node Browser) folders titled Classification and Discrimination, Regression Modeling and Multivariate Exploration, and General Forecaster and Time Series, to perform complex analyses with automatic deployment and cooperative and competitive evaluation of models (see Figure 10.26).

For example, the Classification and Discrimination folder contains nodes for stepwise and best-subset linear discriminant function analysis, various tree classification methods, generalized linear models procedures, and different neural network architectures (see Figure 10.27).

The analysis nodes with automatic deployment are generally named `TypeOfAnalysis with Deployment`. Simply connect these nodes to an input data source, update (train) the project, and you are ready for deployment. To the node, connect a data source marked for deployment (i.e., select the Data for Deployed Project check box in the dialog specifying the variables for