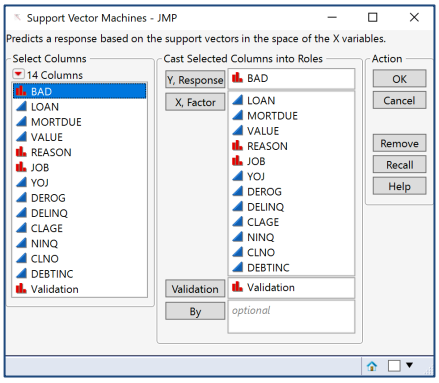
Support Vector Machines

JMP PRO Use this predictive modeling technique to predict (classify) binary and multi-class outcomes as a function of candidate categorial and/or continuous predictor variables. The technique classifies observations by optimizing a hyperplane that separates the classes.

Support Vector Machines

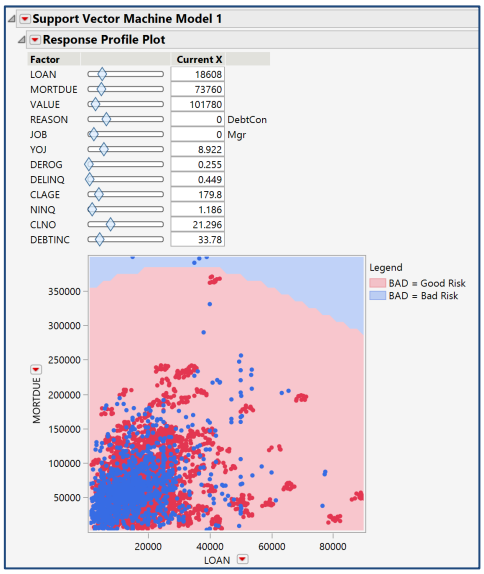
Equity.jmp (Help > Sample Data Library)



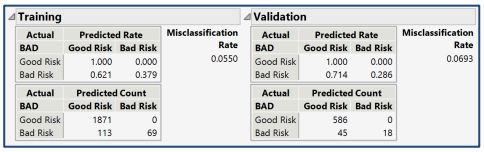
1. From an open JMP® table, select **Analyze > Predictive Modeling > Support Vector Machines.**
2. Add a nominal or ordinal response variable from **Select Columns** to the **Y, Response** role.
3. Add candidate predictor variables to the **X, Factor** role.
4. If desired, enter a validation column into the **Validation** role as shown in this example.
5. Click **OK**. The Model Launch control panel opens allowing a choice of a Kernal Function and associated options. Default settings were used for this example. Click **Go**.

   JMP displays:

   - **Response Profile Plot** displaying the classification regions and the data values for two of the predictor variables. These can be changed to other variables by selecting the red triangle next to the variable name on the axes. Levels for all the remaining predictors can be changed with the sliders above the plot.
   - **Model Summary** (not shown) and a **Confusion Matrix** detailing the classification performance.



Additional models can be made with different Kernal Function settings under Model Launch. A Model Comparison table will be generated displaying performance metrics for each fitted model.



Intepretation:

- There are 649 observations in the Vaidation Data. Of these, 45 (6.9%) where misclassified. 45/(45+18) = 71% of the Bad Risk customers were misclassified as Good Risk. 0/(0+586) = 0% of the Good Risk customers were misclassified as Bad Risk.

Note: The default rule is to classify an observation in the class with the highest estimated probability of being in that class (i.e., Prob > 0.50). It can be advantageous to evaluate different cutoff values in order to minimize a specific type of misclassificaiton rate over another. In this example it may be better to choose a cutoff level to create a lower misclassification rate for Bad Risk Customers misclassified as Good Risk while accepting a higher misclassification rate for the other.

Additional options, such as **ROC** and **Lift Curves**, **Profilers**, **Save Predicteds**, **Save Prediction Formula**, **Save Probabilities**, as well as **Publish Probability Formulas** are accessible from the **red triangle**.