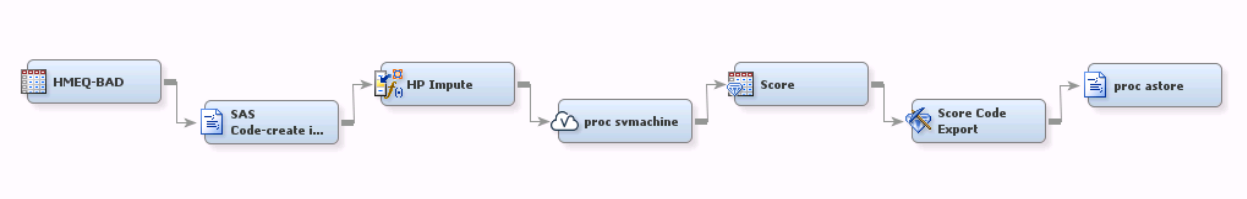
Scoring in SAS® Viya™ Using SAS® Enterprise Miner™

Process Flow Diagram:



Data:
The Home Equity (HMEQ) data set in the SAMPSIO SAS library is used to create the data source. You can generate the data in SAS Enterprise Miner 14.3 by selecting **Help > Generate Sample Data Sources** to open the Generate Sample Data Sources window and then selecting **Home Equity.** The target is the binary variable **BAD**, which indicates whether an applicant defaulted (or was seriously delinquent) on a home equity loan, or paid the loan on time.

Goal:
The goal is to use a SAS Viya Code node to run the SVMACHINE procedure in SAS® Visual Data Mining and Machine Learning to train a support vector machine (SVM) algorithm and score data.

Flow:
The Input Data node (titled "HMEQ-BAD" in the flow diagram) represents the data source and provides details about the variables in the data source.

The Input Data node is connected to a SAS Code node (titled "SAS Code-create id variable" in the flow diagram), which enables you to incorporate SAS code into the process flow diagram. You can use the **Code Editor** to enter your SAS code. The SAS code in the **Training code** pane creates an **ID** variable that is used to join the input records with their corresponding scores.

The HP Impute node executes high-performance numeric variable imputation. The **Default Input Method** property is set to **Count** for classification variables. This property requests that missing classification values be replaced with the most frequently occurring classification variable value. The **Default Input Method** property is set to **Mean**. This property requests that missing values of interval variables be replaced with the arithmetic average of the variables. The **Random Seed** property specifies the initial seed value for generating random numbers. The default value is 12345.

The SAS Viya Code node (titled "proc svmachine" in the flow diagram) executes the support vector machine algorithm in SAS Viya using the SVMACHINE procedure in SAS Visual Data Mining and Machine Learning. INPUT statements are included to define the input variables as categorical (nominal) and continuous (interval) variables. The SAVESTATE statement is included to save the state of the SVMACHINE procedure in the analytic store, which is stored in the table named **&em_data_rstore**.

The Score node evaluates and saves the scoring code that is generated from the flow, including the SVM algorithm implemented by the SAS Viya Code node. After you run the Score node, the **SAS Code** table in the **Results** shows the score code that is generated by the process flow. The **Output Variables** table displays the output variables that are created by the score code.

The Score Code Export node exports the SAS scoring model program **score.sas,** the analytic store **score.sasast**, and the DS2 scoring model program **epscore.sas**. These files are necessary for score code deployment.

The second SAS Code node (titled “proc astore” in the flow diagram) includes SAS code that scores the input data in two different ways. One way uses the **score.sas** file directly and stores the scored results in the data table **score**. The second way executes the ASTORE procedure, which uses the analytic store file **score.sasast**. In the OUT= option in the SCORE statement, you specify a new data table (**score2**) to contain the scored results. The COMPARE procedure compares the contents of the two SAS data tables (**score** and **score2**) and shows that the scoring methods yield the same results.