Paper SAS3459-2019

Unleashing SAS® Visual Data Mining and Machine Learning Models

Shawn Pecze, Prasanth Kanakadandi, Byron Biggs, and Xin Chi

SAS Institute Inc., Cary, NC

ABSTRACT

This paper shows how you can use SAS® Visual Data Mining and Machine Learning and other SAS products to build and compare various predictive models. First, you use SAS Visual Data Mining and Machine Learning to create several models and you choose one of them as your champion model. You can publish all these models to different destination types such as Hadoop, Teradata, SAS® Cloud Analytic Services (CAS), and SAS® Micro Analytic Service. You use SAS® Embedded Process to score the data against these published models where the data reside. You can also register the models to SAS® Model Manager and compare them against other models for final champion model selection. You can then test these models to validate them for scoring. If you notice a degradation in the model, you can retrain the model. Retraining the model triggers a run of all the pipelines in the associated SAS Visual Data Mining and Machine Learning project, and the recalculated project champion is automatically registered back to SAS Model Manager. In addition, you can score streaming data by using SAS® Event Stream Processing on the models that are registered in SAS Model Manager. SAS Visual Data Mining and Machine Learning also provides a scoring API that enables you to score models directly in Model Studio by using RESTful interfaces. This paper shows how you can unleash the full power of your models by taking advantage of the model processing capabilities in all these SAS products.

INTRODUCTION

Figure 1 shows the analytics life cycle. SAS® Viya® provides all the necessary components in a seamless and intuitive manner, so managing the full analytics life cycle is easier than ever.

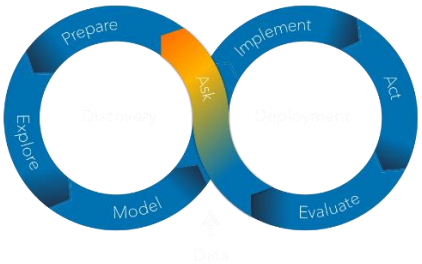


Figure 1. The Analytics Life Cycle

As with most things in life, the analytics life cycle starts with a question. Your organization probably has a multitude of data that was collected through various business processes and holds the answers to this question.

Because these data can come from any number of sources—even directly from a user—you can start by preparing your data for analysis by using SAS® Data Management. Preparation might involve joining data from multiple sources into a single source table, cleansing and augmenting the data, partitioning the data, and so on.

Once you've prepared your data, you can begin exploring and visualizing the data by using SAS® Visual Analytics. You can examine and understand patterns, trends, and the relationships in your data and apply various predictive analytics models and visualizations to begin gaining valuable insights.

You can then take the models you've created in SAS Visual Analytics and create analytics pipelines in SAS Visual Data Mining and Machine Learning. You can augment these pipelines with additional models, continually tuning and comparing them until you've built the best model for your data.

You can then deploy your model from SAS Visual Data Mining and Machine Learning by publishing the model directly to a location that contains the data you want to score. You can also register the model in SAS Model Manager to take advantage of model versioning and governance. From SAS Model Manager, you can deploy your model into your SAS Event Stream Processing environment to score streaming data at various levels, including IoT gateways and edge devices.

As your data evolve, you can use SAS Model Manager to monitor your model's performance over time so that you can make decisions about whether you should rebuild your models.

If you decide to rebuild your model, you go back to the beginning of the analytics life cycle. All these steps are contained within one unified analytics environment.

The rest of this paper presents a typical scenario that highlights this concept of multi-phase analytics: training a model using offline data in SAS Visual Data Mining and Machine Learning, deploying this model into SAS Event Stream Processing, and managing the model in SAS Model Manager. The data set for this example is the Turbofan Engine Degradation Simulation Data Set from the NASA Ames Prognostics Data Repository (Saxena and Goebel 2008). The target variable for this example is FuelRatio, which is an interval target.

This paper also describes other ways you can operationalize your model through the various supported model-publishing destinations.

BUILDING MODELS

SAS Visual Data Mining and Machine Learning enables you to build complex analytics pipelines to determine the best model for your data. You can create SAS Visual Data Mining and Machine Learning pipelines in either of the following ways:

- You can create a pipeline from a model built in SAS Visual Analytics as a starting point by clicking **Create pipeline** in the report you've created as shown in Figure 2. You can add this pipeline to a new project (SAS Visual Data Mining and Machine Learning project), or you can add the pipeline to an existing project (SAS Visual Data Mining and Machine Learning project).

Figure 2. Creating a Pipeline from SAS Visual Analytics

- You can manually create a project in Model Studio and build one or more pipelines. The remainder of this section describes this way of creating a pipeline.

To manually create your project in Model Studio, begin by creating a new SAS Visual Data Mining and Machine Learning project directly in Model Studio. You train your models using offline data that are stored in a CAS table. In the **New Project** dialog box, enter a name for the project and select the CAS table that contains the training data.

SAS Visual Data Mining and Machine Learning comes with many prebuilt pipelines that illustrate best-practice strategies for predictive modeling. Since the target variable is an interval target, you can select one of the templates specific for interval targets from the **Template** list. For this example, you select **Advanced template for interval target**. Figure 3 shows the selections for creating the project.

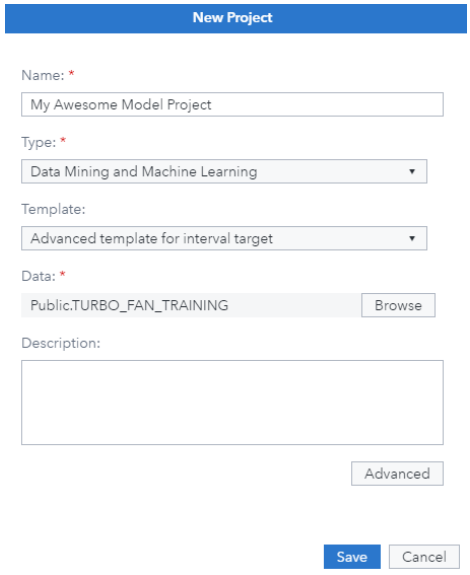


Figure 3. Creating a SAS Visual Data Mining and Machine Learning Project

When you create projects from SAS Visual Analytics, the project metadata for the variables in the selected table are automatically set for you according to the roles that you assigned in SAS Visual Analytics. However, when you create the project manually, you must ensure that the desired target variable is specified. You can modify the metadata for the other variables in the project in order to control how those variables are used to train your model.

To predict FuelRatio, select it as the target variable, as shown in Figure 4.

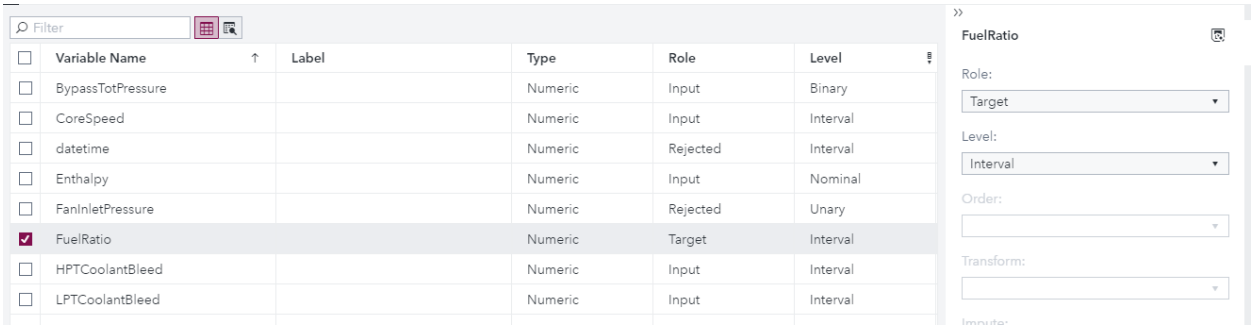


Figure 4. Defining the Target Variable

Click the **Pipelines** tab, and you will see that the pipeline that was selected during the project creation is fully built, has been automatically added to the project, and is ready for use. Figure 5 shows the pipeline.

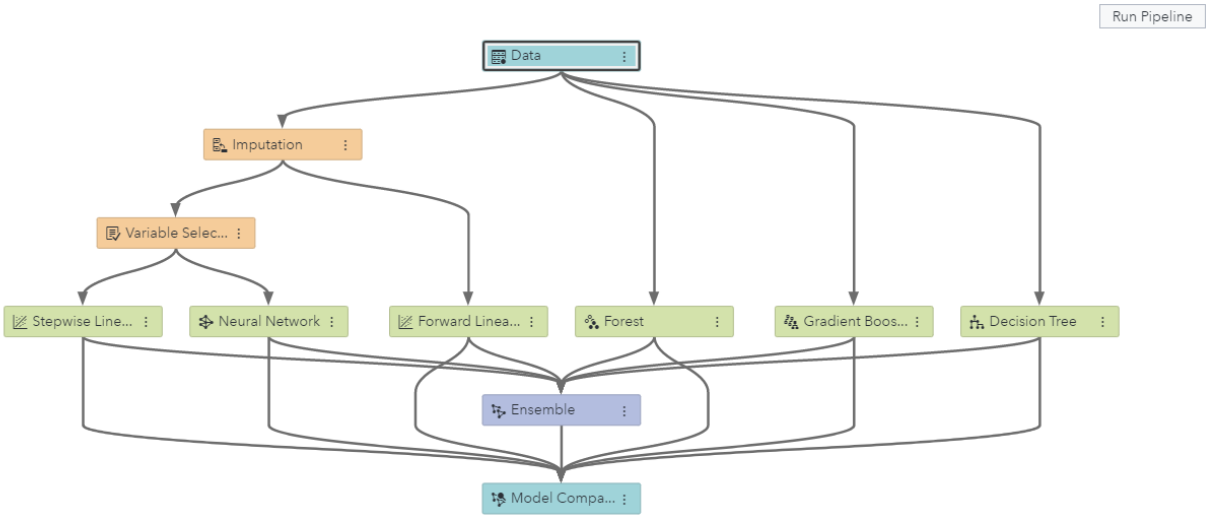


Figure 5. Initial Pipeline

You can modify this pipeline by adding or removing nodes and by tweaking the properties of the nodes. You can also add additional pipelines to the project by using any of the other available pipeline templates, or by using the blank template and building a pipeline manually. For more information about all the templates available, see the section "Available Templates" in *SAS Visual Data Mining and Machine Learning: User's Guide.*

When you run a pipeline, a champion model is automatically chosen for the pipeline according to predefined model comparison rules. You can modify the model comparison rules by selecting the Model Compare node and updating the properties to select the desired comparison options. You can also update the project settings to set the model comparison rules for all pipelines in the project. For more information, see the section "Overview of Model Comparison" in the *SAS Visual Data Mining and Machine Learning: Reference Help*.

You can see the champion model that was chosen by viewing the results for the Model Comparison node. Figure 6 shows that the gradient boosting model was chosen as the champion model for this example.

Model Comparison

Champion	Name	Algorithm Name	Average Squared Error	Root Average Squared Error
✓	Gradient Boosting	Gradient Boosting	0.0161	0.1269
	Decision Tree	Decision Tree	0.0633	0.2517
	Ensemble	Ensemble	0.0666	0.2581
	Forest	Forest	0.0824	0.2871
	Forward Linear Regression	Linear Regression	0.0909	0.3015
	Stepwise Linear Regression	Linear Regression	0.0909	0.3015
	Neural Network	Neural Network	0.2799	0.5291

Figure 6. Model Comparison Results

Although a single champion model is automatically selected for each pipeline, you can add additional models for final comparison by adding them as challenger models. To add challenger models, select the desired modeling node and select **Add challenger model** from the node menu. Figure 7 shows how you can add a decision tree as a challenger model for this example.

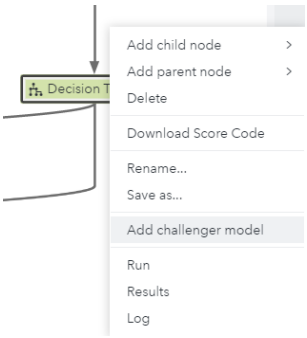


Figure 7. Adding Challenger Models

CHOOSING A CHAMPION MODEL

The **Pipeline Comparison** tab displays all the candidate models in your project. These models include the champion model for each of the pipelines in the project and any challenger models you've added. SAS Visual Data Mining and Machine Learning automatically chooses a champion model for your project. You can use this project champion, or you can override it by selecting the desired model and choosing **Set as champion** from the project pipeline menu, as shown in Figure 8.

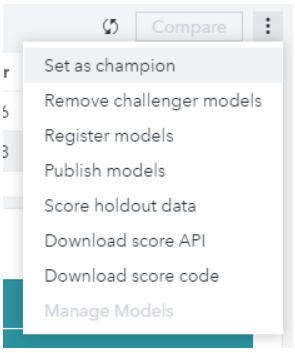


Figure 8. The Project Pipeline Menu

If you want to further validate the project models, you can select **Score holdout data** from the menu. You will be prompted to select a CAS table that contains additional test data that are separate from the data that were used to train the models. New assessment statistics are generated for each of the models and you can use that information to help in determining the champion model.

For this example, you override the champion model by setting the challenger decision tree model as the champion model.

DEPLOYING THE MODEL TO SAS EVENT STREAM PROCESSING

You can deploy the champion model to your SAS Event Stream Processing environment so that you can use the model to score streaming data in real time. The four steps in deploying the model to SAS Event Stream Processing (registering the model, preparing the model, building the project, and deploying the project) are described in the following subsections.

REGISTERING THE MODEL IN SAS MODEL MANAGER

Registering your SAS Visual Data Mining and Machine Learning models to SAS Model Manager enables you to store them in a common model repository alongside your other analytical models. Storing the model enables you to compare heterogeneous models, run tests against the models, publish models, monitor model performance, and create custom workflows for your business processes.

You can use the common model repository to separate your project and model content, and to set permissions for objects within a repository. For example, you can have different repositories for test and production environments, or for different organizations. The default model repository for SAS Visual Data Mining and Machine Learning is DMRepository.

In the example, register the champion model to SAS Model Manager by first selecting the model in the **Pipeline Comparison** tab. From the project pipeline menu (see Figure 8), select **Register models**. Figure 9 shows the result of registering the champion decision tree model.

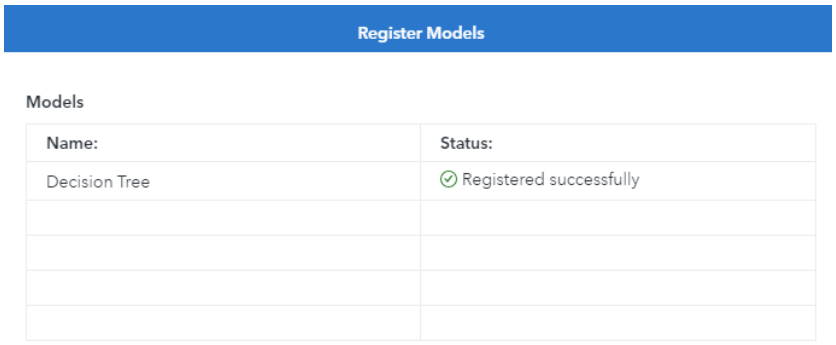


Figure 9. Model Registration Status

Once you have registered the model, a new SAS Model Manager project is created, and the model artifacts (score code, analytics stores, and so on) are associated with the new project. You can navigate to the SAS Model Manager project from SAS Visual Data Mining and Machine Learning by selecting **Manage Models** from the project pipeline menu in the **Pipeline Comparison** tab (see Figure 8).

PREPARING THE MODEL FOR SAS EVENT STREAM PROCESSING

Now that the champion model is registered with SAS Model Manager, there are a few steps you need to take to prepare the model for deployment to SAS Event Stream Processing.

1. Navigate to the **Model** View for the registered model in SAS Model Manager by selecting **Manage Models** from the project pipeline menu (see Figure 8). Figure 10 shows the registered content for your champion model.

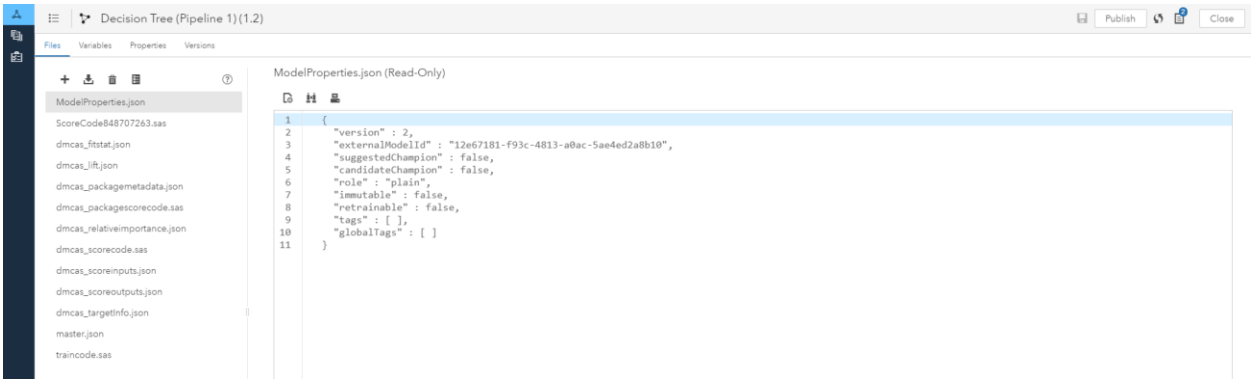


Figure 10. Model View

2. Select the **Properties** tab.
3. From the **Target level** list, select **Interval** because the level for the target variable is interval.
4. From the **Output prediction variable** list, select EM_PREDICTION.
5. Save the changes by clicking the Save icon ( ) on the toolbar.
6. Navigate to the **Project View** by clicking the project name link.
7. Set the champion model for the SAS Model Manager project by selecting the desired model and choosing **Set as champion** from the action menu ( ). Figure 11 shows how you set the decision tree model as the champion model.

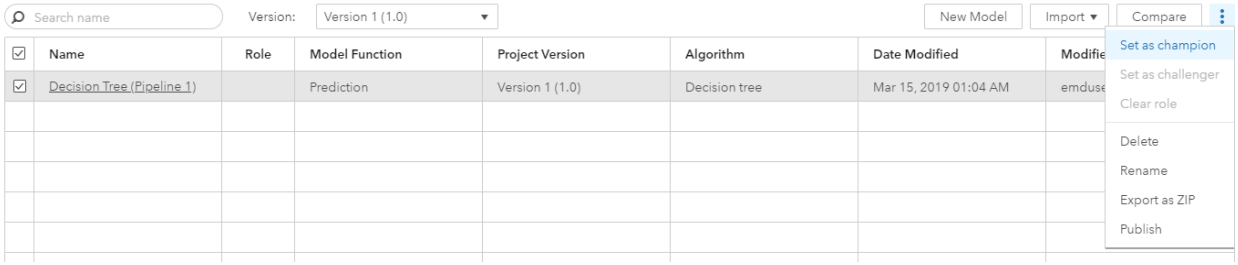


Figure 11. Setting the Project Champion

BUILDING THE SAS EVENT STREAM PROCESSING PROJECT

To deploy your model to your SAS Event Stream Processing environment, you can import the model directly from SAS Model Manager into SAS Event Stream Processing Studio project.

From SAS Event Stream Processing Studio, navigate to the **Projects** view and click **New** to create a new project. For this example, name your project FuelRatio.

A SAS Event Stream Processing Studio is comprised of one or more continuous queries. A default continuous query is created in your project. For more information, see *SAS Event Stream Processing: Using SAS Event Stream Processing Studio*.

Defining an Event Source

Once you've created the project, you need to define an event source. From the **Windows** view, expand **Input Streams** and drag a **Source** window into the continuous query. Figure 12 shows the **Windows** view.

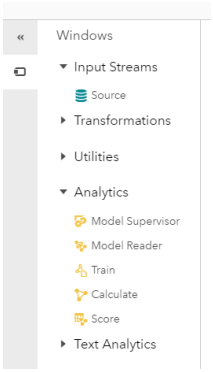


Figure 12. SAS Event Stream Processing Studio Windows View

For this example, you simulate the streaming data by configuring the **Source** window to read data from a CSV file. In production, however, your model would get its data from an actual data source. For more information about the supported event sources, see *SAS Event Stream Processing: Connectors and Adapters*.

To configure the source to read the CSV file, from the **Source** window properties, expand the **Input Data (Publisher) Connectors** section and add a new entry. The entry will have a **Connector type** of **File/Socket Connector**, and you will need to specify the path to the CSV file that contains the data. Figure 13 shows the sample source configuration for the example.

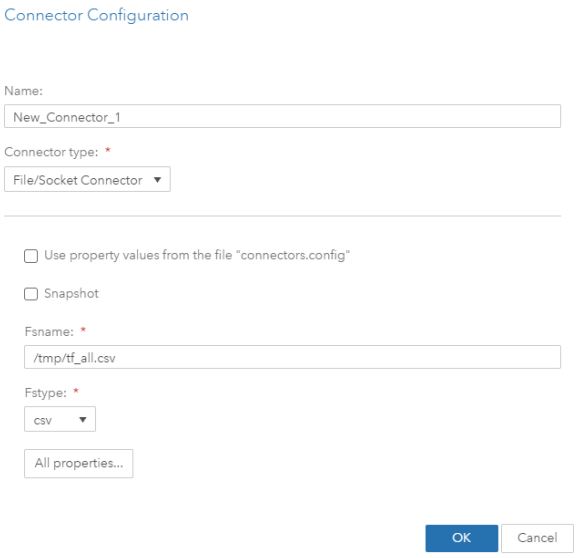


Figure 13. Defining a Source Connector

You also need to define the schema for the data that the **Source** window will provide. Figure 14 shows the schema that is defined for the example data.

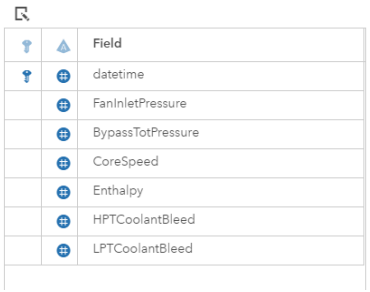


Figure 14. Defining the Source Window Schema

Importing the Model from SAS Model Manager

Once you've defined an event source, you can now import the model. You do this by adding a **Calculate** window to the continuous query and connecting it to the **Source** window.

From the **Windows** view (see Figure 12), expand the **Analytics** section and drag a **Calculate** window into the continuous query. Connect the **Source** window to the **Calculate** window by selecting the **Source** window and drawing a connection to the **Calculate** window. Figure 15 shows the continuous query with the Source window connected to the Calculate window.

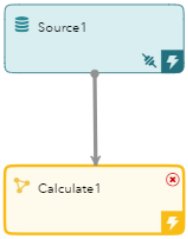


Figure 15. Continuous Query

To import the model:

1. Click in the **Calculate** window to display the window properties panel.
2. In the properties pane, expand the **Settings** section, and select **User-specified** from the **Calculation** list. An undefined handler is automatically added to the handlers list.
3. Edit this handler by selecting it and clicking the Edit icon ( ).
4. Select **Import a module from SAS Model Manager** from the **Handler type** list, as shown in Figure 16.

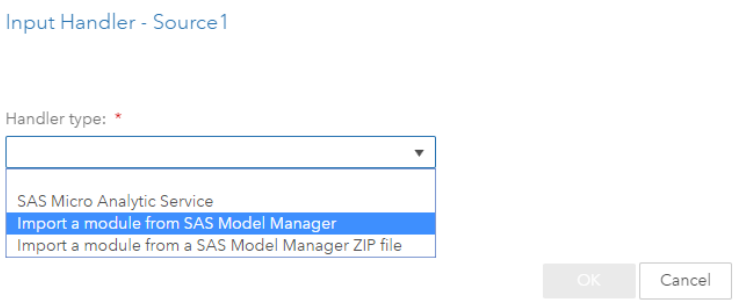


Figure 16. Selecting a Handler Type

5. The **Import from SAS Model Manager** dialog box appears, as shown in Figure 17. Select the decision tree model that you registered to SAS Model Manager.

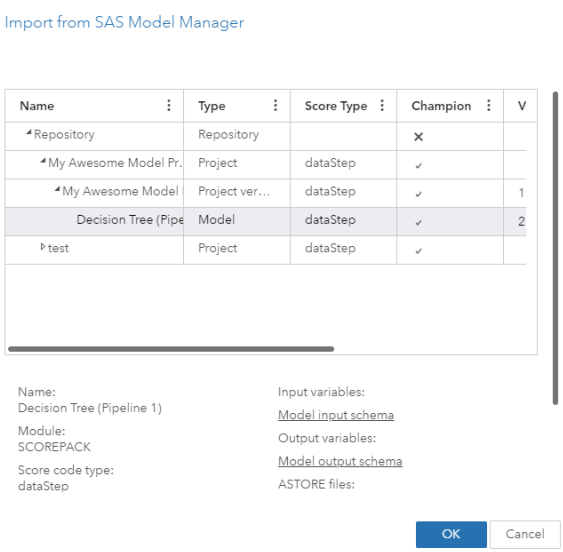


Figure 17. Selecting the Model to Import

6. Click **OK** to close the dialog box, and then click **OK** again to close the **Input Handler** dialog box.
7. You may optionally configure the **Calculate** window to write the score outputs to a new file. To do this, expand the **Subscriber Connectors** section and define a new **File/Socket Connector**. See Figure 13 for the options available for the connector.

Testing the Project

You can now test the project in SAS Event Stream Processing Studio. You must have a SAS Event Stream Processing server running and registered in SAS Event Stream Processing Studio. For more information on how to do this, see the section "Managing ESP Servers in SAS Event Stream Processing Studio" in *SAS Event Stream Processing: Using SAS Event Stream Processing Studio*.

You enter test mode by clicking **Enter Test Mode** in the project toolbar, as shown in Figure 18.

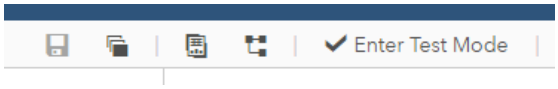


Figure 18. Project Toolbar

In the **Test** view, click **Run Test** to start the test. Figure 19 shows the buttons used to start and stop the test.

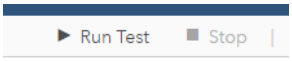
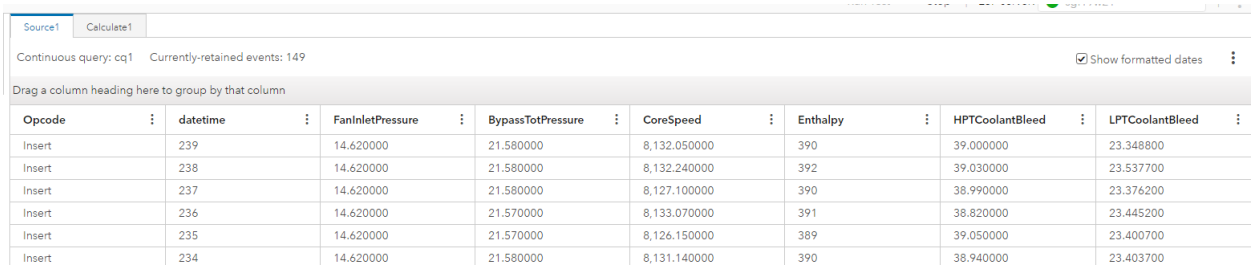


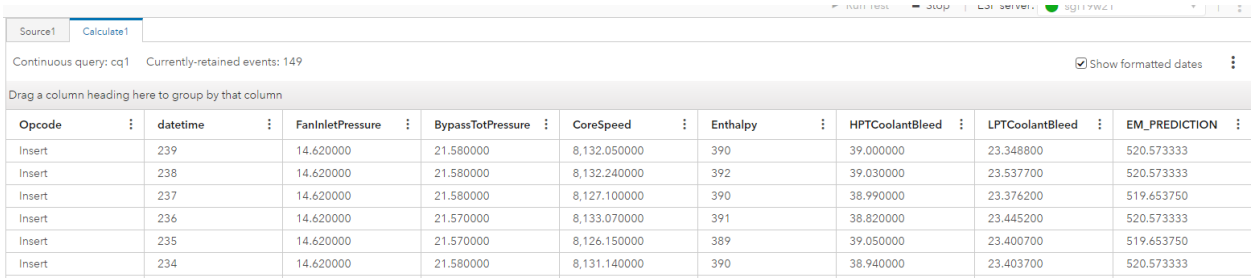
Figure 19. Test Control Buttons

As the test runs, you will see the output for each of the windows that are defined in the project. The **Source** window (see Figure 20) shows the events that were processed, and the **Calculate** window (see Figure 21) shows the results of scoring each event against the model.



Opcode	datetime	FanInletPressure	BypassTotPressure	CoreSpeed	Enthalpy	HPTCoolantBleed	LPTCoolantBleed
Insert	239	14.620000	21.580000	8,132.050000	390	39.000000	23.348800
Insert	238	14.620000	21.580000	8,132.240000	392	39.030000	23.537700
Insert	237	14.620000	21.580000	8,127.100000	390	38.990000	23.376200
Insert	236	14.620000	21.570000	8,133.070000	391	38.820000	23.445200
Insert	235	14.620000	21.570000	8,126.150000	389	39.050000	23.400700
Insert	234	14.620000	21.580000	8,131.140000	390	38.940000	23.403700

Figure 20. Source Window Output



Opcode	datetime	FanInletPressure	BypassTotPressure	CoreSpeed	Enthalpy	HPTCoolantBleed	LPTCoolantBleed	EM_PREDICTION
Insert	239	14.620000	21.580000	8,132.050000	390	39.000000	23.348800	520.573333
Insert	238	14.620000	21.580000	8,132.240000	392	39.030000	23.537700	520.573333
Insert	237	14.620000	21.580000	8,127.100000	390	38.990000	23.376200	519.653750
Insert	236	14.620000	21.570000	8,133.070000	391	38.820000	23.445200	520.573333
Insert	235	14.620000	21.570000	8,126.150000	389	39.050000	23.400700	519.653750
Insert	234	14.620000	21.580000	8,131.140000	390	38.940000	23.403700	520.573333

Figure 21. Calculate Window Output

Click **Stop** to end the test (see Figure 19).

Publishing a New Version of the SAS Event Stream Processing Studio Project

When you are satisfied with your project, click the Versioning icon ( ) on the project toolbar (see Figure 18).

From the **Versioning** view, select the Publish a new version icon ( ), as shown in Figure 22.

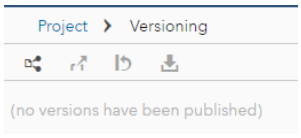


Figure 22. Versioning Toolbar

You can add optional notes, and then click **OK** to create the new version. The newly created version will appear in the **Versioning** view. Figure 23 shows the newly created version of the project.

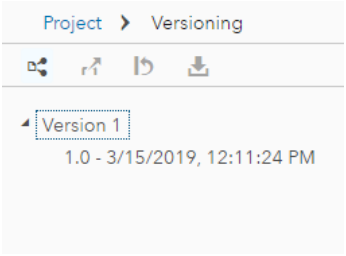


Figure 23. New Version Created

DEPLOYING THE SAS EVENT STREAM PROCESSING PROJECT

To deploy the SAS Event Stream Processing Project to SAS Event Stream Processing, you use SAS® Event Stream Manager. When you publish a version of a SAS Event Stream Processing Studio project, the published version becomes available for deployment in SAS Event Stream Manager.

If you navigate to SAS Event Stream Manager and select the **Projects** view, you will see your project in the list. Figure 24 shows the project that you published in the list.

Name	Production	Tags	Latest Version	Published
FuelRatio	Yes		1.0	3/15/2019, 12:11:24 PM

Figure 24. SAS Event Stream Manager Projects View

You can choose whether the project is a production project by right-clicking the project and selecting **Toggle production** from the context menu.

To deploy the project, you need to create a job template. For more information about how to define job templates, see the chapter "Working with Job Templates" in *SAS Event Stream Manager: User's Guide*.

Once you have a job template defined for deploying the project, navigate to the **Job Templates** view. Select the desired job template and click the Run a job using the template icon ( ). Figure 25 shows the toolbar options.

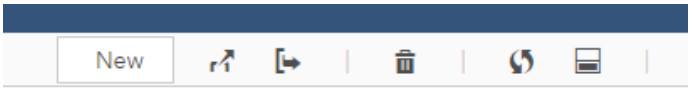


Figure 25. Job Templates View Toolbar

In this example, a job template called Load Project is defined and is used to deploy the project to the production SAS Event Stream Processing server. Figure 26 shows the typical options that are needed to deploy a project. The job requires you to select the desired deployment, project, version, and server.

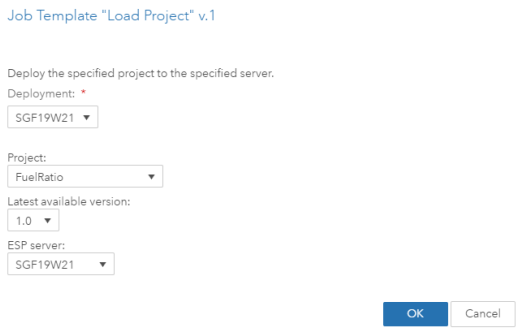


Figure 26. Sample Job Template for Loading Projects

Click **OK** to run your job. When the job completes, you will see the job completion status for each of the servers to which you deployed the project. Figure 27 shows the completion status for the sample job.

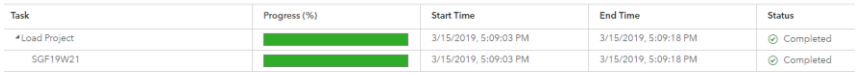


Figure 27. Load Project Status

Once your project is deployed, you can navigate to the **Deployments** view to see which projects are running on which servers. When you click on a deployment, you will see information about the servers and projects available in that deployment. Figure 28 shows the status of a sample deployment with a single server. You can see that the FuelRatio project is running on the server that was specified in the preceding job.

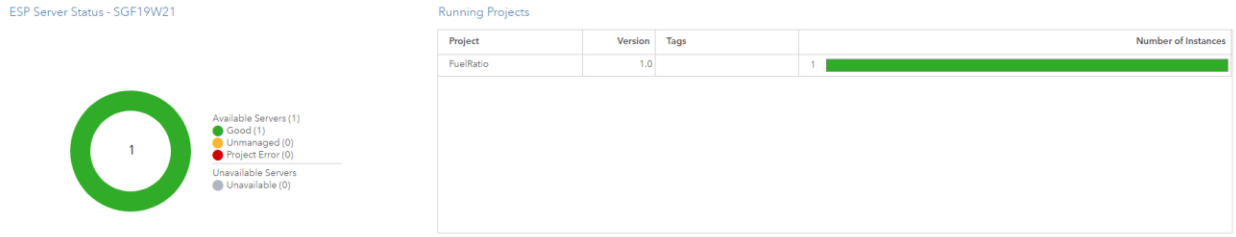


Figure 28. Sample Deployment Status

From within a deployment, you can open a running project to view the data that are associated with each window in the project (**Source** window, **Calculate** window, and so on).

MONITORING MODEL PERFORMANCE

Now that you have deployed the model into your environment, you can use SAS Model Manager to view the performance of the model.

You can collect performance data that have been created by the model at intervals that are determined by your organization. A performance data set is used to assess model prediction accuracy. It includes all the required variables in addition to one or more actual target variables. For example, you might want to create performance data sets monthly or quarterly and then use SAS Model Manager to create a performance definition that includes each time interval.

You can allow SAS Model Manager to score a data set against the model, or you can provide a data set that already contains the predicted values.

SAS Model Manager generates plots—such as variable distribution, characteristic, stability, lift, Gini, ROC, Kolmogorov-Smirnov (KS), and average squared error (ASE) charts—so that you can visualize how the model is performing. Figure 29 shows the charts for the example project.

Figure 29. Model Performance Charts in Model Manager

For more information about model performance, see the section "Monitoring Performance" in *SAS Model Manager: User's Guide*.

REBUILDING MODELS

If you notice that the performance of your SAS Visual Data Mining and Machine Learning model has degraded, you can retrain the registered model from SAS Model Manager.

To retrain a model from SAS Model Manager, navigate to the **Project** view for the project that contains the model and select **Retrain** from the action menu (⋮) at the top of the **Project** view. The **Retrain Project** dialog box appears, as shown in Figure 30.

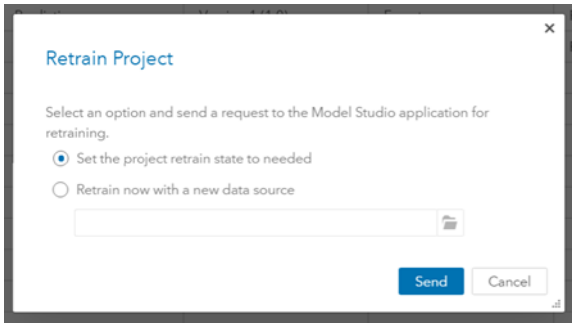


Figure 30. Retraining a Model from Model Manager

If you select **Set the project retrain state to needed**, you can open the SAS Visual Data Mining and Machine Learning project and retrain your models using new data. Before you run the SAS Visual Data Mining and Machine Learning pipelines, you have the option of tweaking the settings of existing modeling components and adding new ones. You can then run the pipelines in the project to recalculate a champion model for the project. You can then choose which models will need to be registered or published (or both) after the retraining operation completes.

If you select **Retrain now with a new data source**, you are prompted to select a CAS table that contains the new data. The SAS Visual Data Mining and Machine Learning project

is automatically retrained with the selected data. The project pipelines are run automatically, and the champion model for the project is registered in SAS Model Manager.

For this example, select **Set the project retrain state to needed**. Then navigate back to your SAS Visual Data Mining and Machine Learning project in Model Studio.

Once in your SAS Visual Data Mining and Machine Learning project, from the **Data** tab, expand the **Data sources** view and click the Replace data source icon ( ). Figure 31 shows the project data sources view.

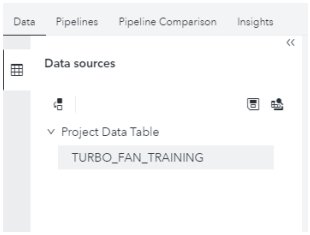


Figure 31. Project Data Sources View

Select a new CAS table that contains the updated data you want to use to retrain your models. Once you've selected a new table, you need to rerun any pipelines in your project. However, before you run the pipelines, you can modify them as needed.

When you run your pipelines, champion models are recalculated for each pipeline. All the champion models and any previously defined challenger models will be shown in the **Pipeline Comparison** tab.

In the example, you previously had overridden the calculated champion model to select the decision tree as the project champion. This setting was remembered, and the decision tree is once again set as the project champion.

However, as you look at the results, you might determine that a different model performed better against the new data. In this example, you add the forward linear regression model as a challenger model and select it as the new project champion.

You register this new model to SAS Model Manager and navigate to the model project in SAS Model Manager. Select the newly registered linear regression model and mark it as the new project champion. Figure 32 shows the new champion model selection.

Figure 32. Selecting a New Project Champion in SAS Model Manager

DEPLOYING THE UPDATED MODEL

When you navigate back to SAS Event Stream Manager, you will notice a notification that an update is available for your FuelRatio project. Figure 33 shows the update notification.

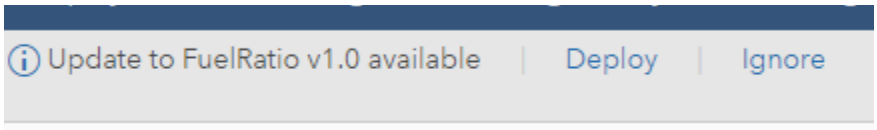


Figure 33. SAS Event Stream Manager Notification

Select **Deploy,** and you have the option to update all projects or update only selected projects, as shown in Figure 34.

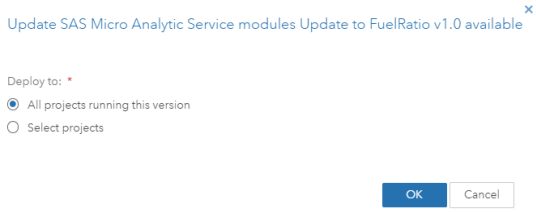


Figure 34. Deploying Updates

For this example, select **All projects running this version**, and click **OK**.

When you navigate to the **Projects** view, you will see that your project version has been incremented, and the **Deployments** view will show that the new version is now running in your environment.

Now you can repeat the process of monitoring the newly deployed model and retraining as needed.

DEPLOYING THE MODEL TO OTHER ENVIRONMENTS

SAS Visual Data Mining and Machine Learning models can also be deployed to other environments.

PUBLISHING THE MODEL

Publishing a model deploys the model directly to the environment that contains the data you want to score. These environments are referred to as publishing destination types.

You can publish your models from SAS Model Manager after you have registered them from SAS Visual Data Mining and Machine Learning. For more information, see the section "Publish Models" in *SAS Model Manager: User's Guide*. You can also publish directly from SAS Visual Data Mining and Machine Learning, which is particularly useful if your environment does not include SAS Model Manager. You can find instructions for publishing directly from SAS Visual Data Mining and Machine Learning in the section "Publish Models" in *SAS Visual Data Mining and Machine Learning: User's Guide*.

Supported Publish Destinations

You can publish a model to CAS, Hadoop, Teradata, or SAS Micro Analytic Service.

Before you can publish models, you must define one or more publishing destinations for your desired destination types. Users with administrative privileges can create the publishing destinations.

Note that publishing to SAS Micro Analytic Service is available only if SAS Model Manager is installed. When SAS Model Manger is installed, a SAS Micro Analytic Service destination named "maslocal" is automatically created during the installation.

For all other publishing destinations, a global CAS library is required for each unique destination:

- To create a publishing destination for publishing models to CAS, the CAS library to which you publish must be in global scope and you must have the required authorizations to the library.

- To publish to Teradata or Hadoop, you must create a global CAS library the contains the database connection information.

For more information about creating publishing destinations, see *SAS Viya Administration Guide: Publishing Destinations: How To*.

Validating Published Models

After you publish a model, you can access the **Publishing Validation** tab in the SAS Model Manager project. You can edit the publishing validation test to select a test data source and output library. The validation test will score the specified test data source against the published model and generate scoring outputs in the output library. Figure 35 shows the result of a publishing validation task that uses a model that was published in SAS Micro Analytic Service.

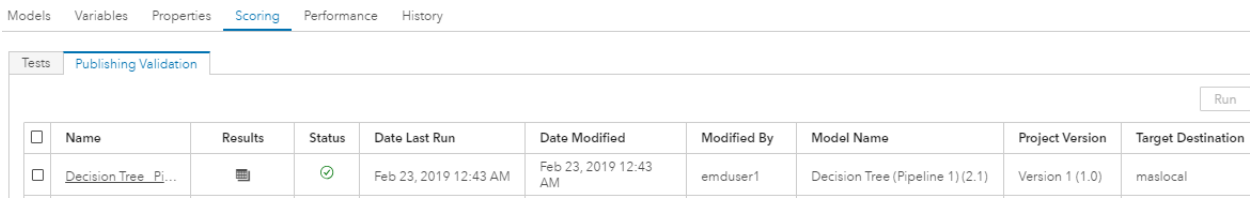


Figure 35. Publishing Validation in Model Manager

Scoring Data in CAS, Teradata, and Hadoop

When you publish a model to CAS, Teradata, or Hadoop, you can use CAS actions to score the data in that environment. You can run these CAS actions by interacting with CAS using any of the supported clients: CAS language (CASL), Python, Lua, and R. For more information about using the clients, see the following documentation:

- *Getting Started with SAS Viya for Python*
- *Getting Started with SAS Viya for Lua*
- *Getting Started with SAS Viya for R*
- *Getting Started with CASL Programming*

To score data that are present on a CAS server, use the runModelLocal action.

To score data in Teradata or Hadoop, use the runModelExternal action.

For more information about the options available for each of the actions, see the chapter "Model Publishing and Scoring Action Set" in *SAS Visual Analytics: Programming Guide*.

Running the Models in SAS Micro Analytic Service

Publishing your model to SAS Micro Analytic Service is useful if you want to score streaming data but don't have SAS Event Stream Processing.

When the model is published to SAS Micro Analytic Service, the ability to score data against the model is surfaced through a set of RESTful interfaces.

The published model is available as a module in the Micro Analytic Score service in your SAS Viya environment. These modules contain a single scoring step.

Using the previous example, if you had published your champion linear regression model to SAS Micro Analytic Service, it would have a URI like the following:

```
/microanalyticScore/modules/forwardlinearregression
```

To score the model, you would issue an HTTP post command to the URI as follows:

```
/microanalyticScore/modules/forwardlinearregression/steps/score
```

The request body for the POST operation would be a JSON string that contains the data to score. The response body would be a JSON string that contains the output of the scoring operation.

For more information about using the SAS Micro Analytic Score REST API, see SAS Institute Inc. (2019).

CREATING A SAS EVENT STREAM PROCESSING STUDIO PROJECT WITHOUT SAS MODEL MANAGER

If you don't have SAS Model Manager in your environment, you can still deploy your SAS Visual Data Mining and Machine Learning models to your SAS Event Stream Processing environment, as follows:

1. Download the score code for the desired model by navigating to the **Pipeline Comparison** tab, selecting the model, and selecting **Download score code** from the project pipeline menu (see Figure 8).
2. When you download the score code for models containing an analytics store, the analytic store is copied to a CAS table in the Models CAS library. You can use the DOWNLOAD statement in the ASTORE procedure to retrieve the binary analytic store file and save it to the local file system. For more information about the DOWNLOAD statement, see the chapter "The ASTORE Procedure" in *SAS Visual Data Mining and Machine Learning: Procedures*.
3. Copy the score code or analytic store (or both) to a disk location that is accessible from your SAS Event Stream Processing environment.
4. When you build your SAS Event Stream Processing Studio project, instead of using a **Calculate** window as in the previous example, use a **Model Reader** window to specify the location of the model score code files. Figure 12 shows the **Model Reader** menu item in the **Analytics** section of the **Windows** view.

DEPLOYING A SAS EVENT STREAM PROCESSING STUDIO PROJECT WITHOUT SAS EVENT STREAM MANAGER

You can deploy your SAS Event Stream Processing Studio project to an SAS Event Stream Processing server without using SAS Event Stream Manager.

To do so, configure and start a SAS Event Stream Processing Server. For more information about running your project in a SAS Event Steam Processing Server, see *SAS Event Stream Processing: Using the ESP Server*.

SCORING THE MODEL IN SAS VISUAL DATA MINING AND MACHINE LEARNING BY USING THE SCORE API

SAS Visual Data Mining and Machine Learning provides a RESTful API that enables you to score data against your model directly in Model Studio. You can download sample code that shows you how to construct and submit the REST request. The sample code is available in SAS and Python. In addition, you can simply download a file that contains the REST request information (endpoint, accept types, and so on) that is used for the scoring operation.

The request body includes the URI of a CAS table that contains the data you want to score. You also provide a CAS library and table to which the score outputs will be written.

To download the sample code for a model, from the **Pipeline Comparison** tab, select the desired model and select **Download score API** from the project pipeline menu (see **Error! Reference source not found.**).

CONCLUSION

As you can see, the integration of SAS products in SAS Viya makes it tremendously easy to manage your analytics life cycle. You can quickly build models in SAS Visual Data Mining and Machine Learning and deploy them to any number of locations in your environment. Although this paper focuses on specific scenarios for operationalizing your SAS Visual Data Mining and Machine Learning models, several alternative methods to achieve the full analytics life cycle are available. The variety of methods available enables you to select the method most appropriate for your situation or environment.

REFERENCES

1. Saxena, A., and Goebel, K. (2008). "Turbofan Engine Degradation Simulation Data Set." Accessed January 17, 2017. NASA Ames Prognostics Data Repository. NASA Ames Research Center, Moffett Field, CA. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/.
2. SAS Institute Inc. (2019). REST APIs for SAS Viya and CAS. Accessed March 22, 2019. https://developer.sas.com/guides/rest.html.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Shawn Pecze
SAS
Shawn.Pecze@sas.com

Prasanth Kanakadandi
SAS
Prasanth.Kanakadandi@sas.com

Byron Biggs
SAS
Byron.Biggs@sas.com

Xin Chi
SAS
Xin.Chi@sas.com