Paper 045-2009

Dashboard Reports for Predictive Model Management

Jifa Wei, SAS Institute Inc., Cary, NC
Emily (Yan) Gao, SAS Institute Inc., Beijing, China
Frank (Jidong) Wang, SAS Institute Inc., Beijing, China
Robert Chu, SAS Institute Inc., Cary, NC

ABSTRACT

Dashboard applications enable users to get big pictures on very complicated sets of data and information quickly. Traditionally, dashboard tools are applied on top of rather descriptive variables. This article, however, describes how SAS® BI Dashboard 4.2 and SAS® Information Delivery Portal 4.2 are applied to a predictive model repository managed through SAS® Model Manager 2.2. Specifically, this article describes a SAS macro that creates model performance SAS data sets and SAS ODS reports based on SAS Model Manager 2.2. Using these SAS data sets and reports, this paper explains the process of creating a BI Dashboard and a WebDAV content portlet of SAS Information Delivery Portal to allow user monitor model performance in a SAS Model Manager repository. Analytical dashboards can help corporate executives quickly gain insight into their critical assets.

INTRODUCTION

The SAS Information Delivery Portal delivers information from SAS Business Intelligence products and SAS Solutions in a targeted and customized way to the Web. The SAS BI Dashboard is a special portlet that is hosted by the SAS Information Delivery Portal. It allows users to easily create a dashboard that displays key business information indicators on the Web. On the other hand, the SAS Model Manager is a SAS solution that manages and monitors analytical model performance in a central repository. This paper describes how to use SAS Information Delivery Portal and SAS BI Dashboard to display model performance from SAS Model Manager. It covers, from a high-level perspective, dashboard reports for classification model performance monitoring. This paper starts with definitions of classification model performance indicators. Then it describes a SAS macro that is used to gather data from SAS Model Manager. A detailed description is given on how to configure the SAS Information Delivery Portal and the SAS BI Dashboard to display the results from the macro. It is assumed that the reader has basic concepts of predictive modeling with minimum understanding of what model performance is. No previous knowledge of the general dashboard report concept is required to read this paper. After reading this paper, the reader should be able to 1) understand the importance of dashboard reporting for predictive model performance monitoring, and 2) know where to start to build customized dashboard reports for predictive model performance monitoring within or beyond the scope of SAS Model Manager.

DATA FEEDS TO THE DASHBOARD

PERFORMANCE INDICATOR EXAMPLES

Dashboard reports for predictive models depend on performance indicators. SAS Model Manager provides the following types of performance monitoring reports and their corresponding indicators:

- Characteristic report that detects and quantifies the shifts in the distribution of variable values in the input data over time. Two indicators, *P1* and *P25*, are used in characteristic report. *P1* and *P25* represent the count of input variables with deviation index scores exceeding 0.1 and 0.25, respectively.

- Stability report that evaluates changes in the distribution of scored output variable values as models score data over time. Indicator *outputDeviation* is used to describe stability report and it represents the deviation levels in the distribution of the model's scored output variables.

- Performance assessment report that evaluates the predicted and actual target values for a champion model at multiple points in time. Lift5Decay, lift10Decay, giniDecay, and ksDecay are key performance indicators used for the performance assessment report. They are defined as the following:

$$lift5Decay = (lift5\_b - lift5)/lift5\_b$$

$$lift10Decay = (lift10\_b - lift10)/lift10\_b$$

$$giniDecay = (gini\_b - gini)/gini\_b$$

$$ksDecay = (ks\_b - ks)/ks\_b$$

where

$lift5$ --The lift value based on the top 5% of the target population of interest of a specified time.

$lift5\_b$ --The lift value based on the top 5% of the target population of interest of training data set.

$lift10$ --The lift value based on the top 10% of the target population of interest of a specified time.

$lift10\_b$ --The lift value based on the top 10% of the target population of interest of training data set.

$gini$ --Gini index based on the target population of interest of a specified time.

$gini\_b$ --Gini index based on the target population of interest of training data set.

ks --KS statistic based on the target population of interest of a specified time.

ks_b --KS statistic based on the target population of interest of training data set.

lift5Decay, lift10Decay, giniDecay, and ksDecay are portable performance indicators. For a more specific example, the current lift of the top 5% for model A is 4.5. Is the lift value good or bad? You do not know unless you also know the top 5% model lift value (for example, 5.0) computed based on the test partition of the training data set. Lift5Decay, in this particular case, is equal to (5.0-4.5)/5.0. Because lift5Decay is a portable performance indicator, it can be used by many models for performance monitoring purposes.

THRESHOLD DEFINITIONS

In the SAS Model Manager, users can set up production model performance measurement thresholds to benchmark and gauge the performance of a classification model. When one of the performance measurements exceeds one or more specified thresholds, warning or alert events occur. When warning or alert events occur, warning or alert notifications are automatically sent by e-mail to recipients.

The thresholds can be configured in either the Define Performance Task wizard of the SAS Model Manager or in a batch program that creates the report specifications. The user needs to configure warning and alert conditions for characteristic reports, stability reports, and performance assessment reports, respectively. The following are examples of alert and warning thresholds for each report:

- Characteristic report

```
alertCondition='p1>5 or p25>0';
warningCondition='p1>2';
```

- Stability report

```
alertCondition='outputDeviation>0.3';
warningCondition='outputDeviation>0.1';
```

- Performance Assessment report

```
alertCondition='(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or
ksDecay>0.1';
warningCondition='lift5Decay>0.05';
```

INDEX EXTRACTION

In the SAS Model Manager, there is one performance data set for each champion model, and the data set is stored under each version where the model is registered. It is not convenient for users to get model performance information of all champion models. Currently, both performance data sets and reports are mainly for the SAS Model Manager users to use. To share reports with users out of the SAS Model Manager, all models performance information should be collected.

A SAS macro %MM_getPerfIndexes is designed to help users collect this information. It creates performance summary data sets for all champion models and generates series of performance reports. These performance data sets can be used as models to define BI Dashboard indicators. These reports can be published to your Web server or integrated with the SAS Information Delivery Portal automatically.

The following paragraph describes the details of this macro. To use this macro, a user needs to define five SAS macro variables that can be used to connect to a SAS Model Manager repository. This macro generates four SAS data sets, including all model performance status table, historical performance status table, and two styles of performance indicator tables.

```
%let _MM_RepositoryId = ModelManagerDefaultRepo;
%let _MM_ServerName = <servername>;
%let _MM_PortNumber = 6411;
%let _MM_User = <userid>;
%let _MM_Password = <Password>;


%MM_getPerfIndexes
(
StatusTable,              /*Total performance status table.*/
HistoryTable,             /*Historical performance status table.*/
IndexTable,               /*Style 1 of performance indicator table.*/
IndexTable2,              /*Style 2 of performance indicator table.*/
CreateReport,             /*Yes means reports will be generated.*/
ReportStyle,              /*Any valid SAS ODS style can be used.*/
ReportFormat,             /*HTML and Excel are supported.*/
ReportDirectory,          /*Local file directory. Reports will be stored here. */
ReportLink);              /*URL link of report entrance on index.html. This is needed
                            for integration with the SAS Information Delivery Portal.*/
```

Follow these steps for core implementation:

1. Get all champion models under default versions of a model repository.

2. Loop to extract performance indicators and generate reports for each champion model of a SAS Model Manager project.

   - Extract performance indicators from a job status data set. Failed jobs are ignored here.

   - Get Warning/Alert/Normal status of current period. The performance status will be a positive number if the performance of Characteristic/Stability/Performance Assessment meets alert/warning condition.

   - Get historical performance status.

   - Parse performance indicator strings and get each performance indicators.

   - Create two styles of performance indicator tables.

   - Generate one performance monitoring report over time.

   - Generate performance indicators reports of each time period.

3. Generate summary performance report for all champion models.

The following four tables are created from executing macro %MM_getPerfIndexes.

Model	Time	Status	Model UUID	Model Path
Decision tree	2008Q3	2	a4c98b01-0a0c-0cb8-00e8-a0cd7e073b2b	/MMRoot/Demo/Hmeq/2008/Decision tree
Score card	2007Q3	0	a6119b2b-0a0c-0cb8-00e8-a0cd0d093280	/MMRoot/Demo/loan/Version 2009/Score card

Figure 1. Total Performance Status Table

- When status = 2, it indicates that the Alert condition previously defined is evaluated as true.

- When status = 1, it indicates that the Warning condition previously defined is evaluated as true, but the Alert condition is evaluated as false.

- When status = 0, it indicates that none of the Alert or Warning condition previously defined is evaluated as true.

Model	Abnormal Count	Current	Current -1	Current -2	Abnormal Link	Abnormal Link 2
Decision tree	8	<a href="http://sbj02.na.sas target="_top"><img border="0" width="20" height="20" src="http://sbj02.na.sas.	<a href="http://sbj02.na.sas target="_top"><img border="0" width="20" height="20" src="http://sbj02.na.sas.	<a href="http://sbj02.na.sas target="_top"><img border="0" width="20" height="20" src="http://sbj02.na.sas.	http://sbj02.na.sas.com	Decision tree
Score card	0	<a href="http://sbj02.na.sas target="_top"><img border="0" width="20" height="20" src="http://sbj02.na.sas.			http://sbj02.na.sas.com	Score card

Figure 2. Historical Performance Status Table

The historical performance status table includes more columns than the screenshot above shows.

The links for Current, Current -1, Current -2 point to the Performance Indicators Report for each time period.

Abnormal Link and Abnormal Link 2 point to the Performance Monitoring Report for all time periods.

The abnormal count reflects performance degradation severity. As described before, the SAS Model Manager allows the user to define various alert and warning criteria for monitoring a model's performance. When the abnormal count is calculated, different status carries different weight. Alert is counted as 2; Warning is counted as 1; and Normal is counted as 0. The bigger the abnormal count is, the more severe the model performance degrades.

Model	Time	p1	p25	lift5	lift10	lift15	lift20	GINI Index	KS Statistic	Output Deviation	lift5 Decay	lift10 Decay	lift15 Decay	lift20 Decay	GINI Decay	KS Decay
Decision tree	2008Q3	4	3	2.477	2.078	1.828	1.624	0.253	0.18	0.003	0.383	0.355	0.292	0.266	0.455	0.478
Decision tree	2008Q2	3	2	2.842	2.327	1.982	1.762	0.32	0.245	0.001	0.293	0.278	0.232	0.204	0.31	0.289
Decision tree	2008Q1	2	0	3.191	2.533	2.115	1.862	0.356	0.267	0	0.206	0.214	0.181	0.159	0.233	0.225
Decision tree	2007Q4	0	0	3.707	2.95	2.397	2.073	0.418	0.309	0	0.077	0.085	0.071	0.064	0.099	0.102
Decision tree	2007Q3	.	.	4.018	3.224	2.581	2.214	0.464	0.344	.	0	0	0	0	0	0
Score card	2007Q3	.	.	4.592	4.382	3.979	3.474	0.814	0.644	.	0	0	0	0	0	0

Figure 3. Performance Indicators Table

The performance indicators table in Figure 3 includes more columns than the screenshot shows.

4

Model	Time	Report Name	Status	Indicator	Current	Baseline	Decay	Alert Condition	Alert Evaluation	Warning Condition	Warning Evaluation
Decision tree	2007Q3	Characteristic	0	p1	.	.	.	p1>5 or p25>0	False	p1>2	False
Decision tree	2007Q3	Characteristic	0	p25	.	.	.	p1>5 or p25>0	False	p1>2	False
Decision tree	2007Q3	Model Assessment	0	giniIndex	0.464	0.464	0	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	False	lift5Decay>0.05	False
Decision tree	2007Q3	Model Assessment	0	ksStatistic	0.344	0.344	0	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	False	lift5Decay>0.05	False
Decision tree	2007Q3	Model Assessment	0	lift10	3.224	3.224	0	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	False	lift5Decay>0.05	False
Decision tree	2007Q3	Model Assessment	0	lift15	2.581	2.581	0	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	False	lift5Decay>0.05	False
Decision tree	2007Q3	Model Assessment	0	lift20	2.214	2.214	0	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	False	lift5Decay>0.05	False
Decision tree	2007Q3	Model Assessment	0	lift5	4.018	4.018	0	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	False	lift5Decay>0.05	False
Decision tree	2007Q3	Stability	0	outputdeviation	.	.	.	outputDeviation > 0.03	False	outputDeviation > 0.01	False

Figure 4. Performance Indicators Table

The following reports are created from executing macro %MM_getPerfIndexes.

1. Summary Report

The Summary Report is created using proc print with model history table.

Performance Status of All Champion Models

Model	Current	Current -1	Current -2	Current -3	Current -4	Current -5	Current -6	Current -7	Current -8	Current -9
Decision tree	2008Q3	2008Q2	2008Q1	2007Q4	2007Q3					
Score card	2007Q3									

Figure 5. Models Performance Summary Report

2. Performance Monitoring Report

If you click on entries in "Model" column of the Summary Report, it will lead you to the Performance Monitoring Report. Part of the report is captured in the following figure.
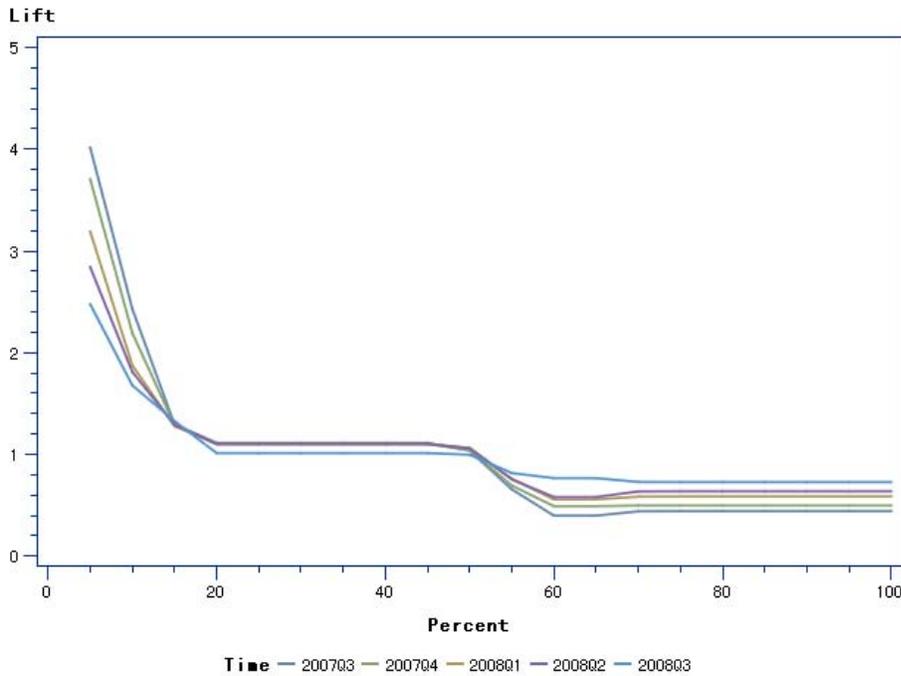
Figure 6. Models Historical Performance Report

3.   Performance Indicators Report

If you click on entries in "Current", "Current -1", or "Current -2" of the Summary Report, it will lead you to
the Performance Indicators Report of each time period, which is created using proc print with
performance indicators table 2.

Time = 2008Q3

Report Name	Status	Indicator	Current	Baseline	Decay	Alert Condition	Alert Evaluation	Warning Condition	Warning Evaluation
Characteristic	🔴	p1	4.000	.	.	p1>5 or p25>0	True	p1>2	True
		p25	3.000	.	.	p1>5 or p25>0	True	p1>2	True
Model Assessment	🔴	giniIndex	0.253	0.464	0.455	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	True	lift5Decay>0.05	True
		ksStatistic	0.180	0.344	0.478	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	True	lift5Decay>0.05	True
		lift10	2.078	3.224	0.355	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	True	lift5Decay>0.05	True
		lift15	1.828	2.581	0.292	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	True	lift5Decay>0.05	True
		lift20	1.624	2.214	0.266	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	True	lift5Decay>0.05	True
		lift5	2.477	4.018	0.383	(lift5Decay>0.15 and lift10Decay>0.12) or giniDecay>0.1 or ksDecay>0.1	True	lift5Decay>0.05	True
Stability	🟢	outputdeviation	0.003	.	.	outputDeviation > 0.03	False	outputDeviation > 0.01	False

Figure 7. Model Performance Report for One Time Period

INTEGRATION WITH THE SAS INFORMATION DELIVERY PORTAL

The reports generated by %MM_getPerfIndexes can be easily incorporated into the SAS Information
Delivery Portal using a WebDAV Content Portlet provided by the SAS Information Delivery Portal.
Detailed instructions on how to create and set up a WebDav Content Portlet can be found from the

SAS Information Delivery Portal online Help. The following are the major steps to integrate these reports into the portal:

1. Determine which SAS Folder to use to store these reports. For example, you could use the following *SAS Folders/sasdav/users/mgrmodels* folder.

2. Run %MM_getPerfIndexes, and generate the three reports. The ReportLink should have the URL link where the reports will be stored in the SAS Folder. For example, the ReportLink could have the following value, depending on a DAV server configuration:

    http://myserver.com:8080/SASContentServer/repository/default/sasdav/Users/mgrmodels

3. Move the generated reports to the folder under sasdav on the SAS Content Server using Microsoft WebFolder application.

4. Create a WebDAV Content Portlet using the contents in the folder you populated in previous steps.

Once the portlet is configured, you can easily create a script to automate Step 2 and 3. Consequently, the content of this portlet will be automatically updated.

INTEGRATION WITH THE SAS BI DASHBOARD

Dashboards are information delivery tools that are used to graphically display key metrics within an organization. SAS BI dashboard is an application that represents the latest evolution of dashboard using SAS. SAS BI Dashboard is surfaced as a portlet in the SAS Information Delivery Portal.

KEY GOALS IN THE VISUAL DESIGN PROCESS

Edward R. Tufte introduced a concept in his 1983 classic *The Visual Display of Quantitative Information* that he calls the "data-ink ratio." When quantitative data is displayed in printed form, some of the ink that appears on the page presents data and some presents visual content that is not data (also known as non-data).

Tufte defines the data-ink ratio in the following way:

A large share of ink on a graphic should present data-information, the ink changing as the data change. Data-ink is the non-erasable core of a graphic, the non-redundant ink arranged in response to variation in the numbers represented. Then,

Data-ink ratio

= data-ink / total ink used to print the graphic

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= 1.0 - proportion of a graphic that can be erased without loss of data-information.

He then applies it as a principle of design: "Maximize the data-ink ratio, within reason. Every bit of ink on a graphic requires a reason. And nearly always that reason should be that the ink presents new information."

EFFECTIVE VISUALIZATION OF SAS BI DASHBOARD

The key goals in the visual design process are applied for dashboard. Additionally, dashboards must be able to condense a lot of information onto single screen and present it at a glance without sacrificing anything important or compromising clarity.

The important information can be divided into two categories:
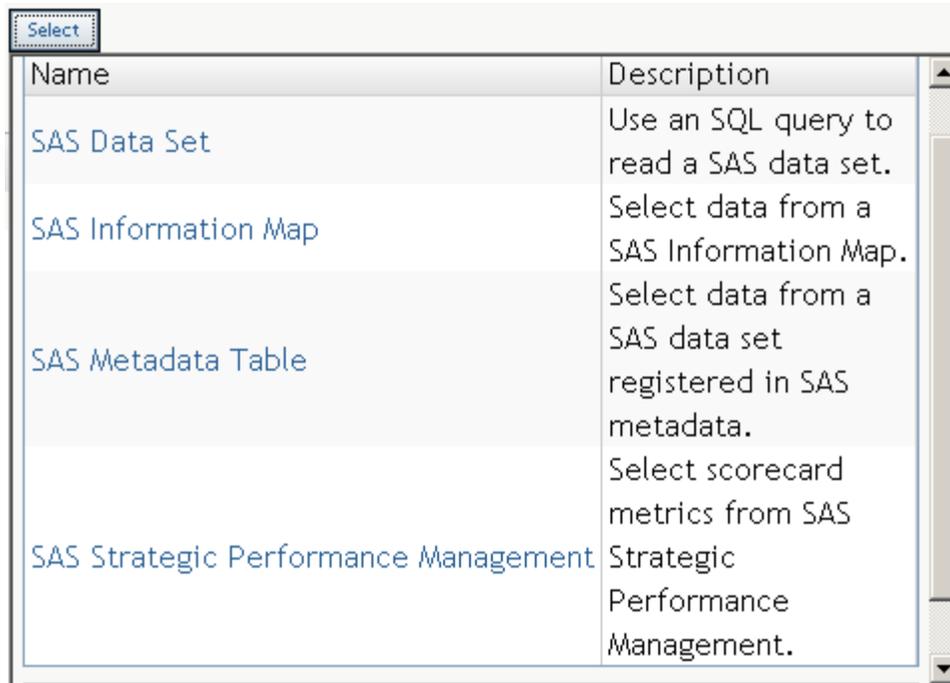
1. information that is always important

2. information that is only important at the moment

For model performance key indicators (Gini index, lift value, KS statistics etc.), the values in current period are more important than previous periods. A well designed dashboard can balance the indicators for different periods and tell you the model performance at a glance.

A GENERAL WAY TO CREATE SAS BI DASHBOARD FOR MODEL MANAGEMENT

1.   Understand model performance data, and create data models.

Everything you want to display in a dashboard must be defined in a data model. The data used in a data model for the SAS BI Dashboard can come from different data sources such as SAS Information Maps or SAS tables. A data model must be available before creating a dashboard.

Figure 8. SAS BI Dashboard Data Model Sources

SAS Data Set data source is a flexible and powerful way to get data. You can use an SQL query to get the data value you want to display in an indicator.

SAS Information Map data source is designed for the business users. It maps the physical data structure with business information.

SAS Metadata Table data source is SAS data set registered in SAS metadata. You can use them directly.

SAS Strategic Performance Management data sources make the scorecard metrics from SAS Performance Management available for you.

2.   Define ranges using the model performance indicators.

The range defines the criteria against which KPIs will be measured. The gauge shows how the significance of the values is interpreted.

Figure 9. SAS BI Dashboard Ranges Definition

The ranges map the interval data with different colors. For example, in the above picture, we map the values above 10 with green, the values above 0 and below 10 with yellow, and the values below 0 with red.

3.   Design the indicators, and associate model data, ranges, and gauges together.
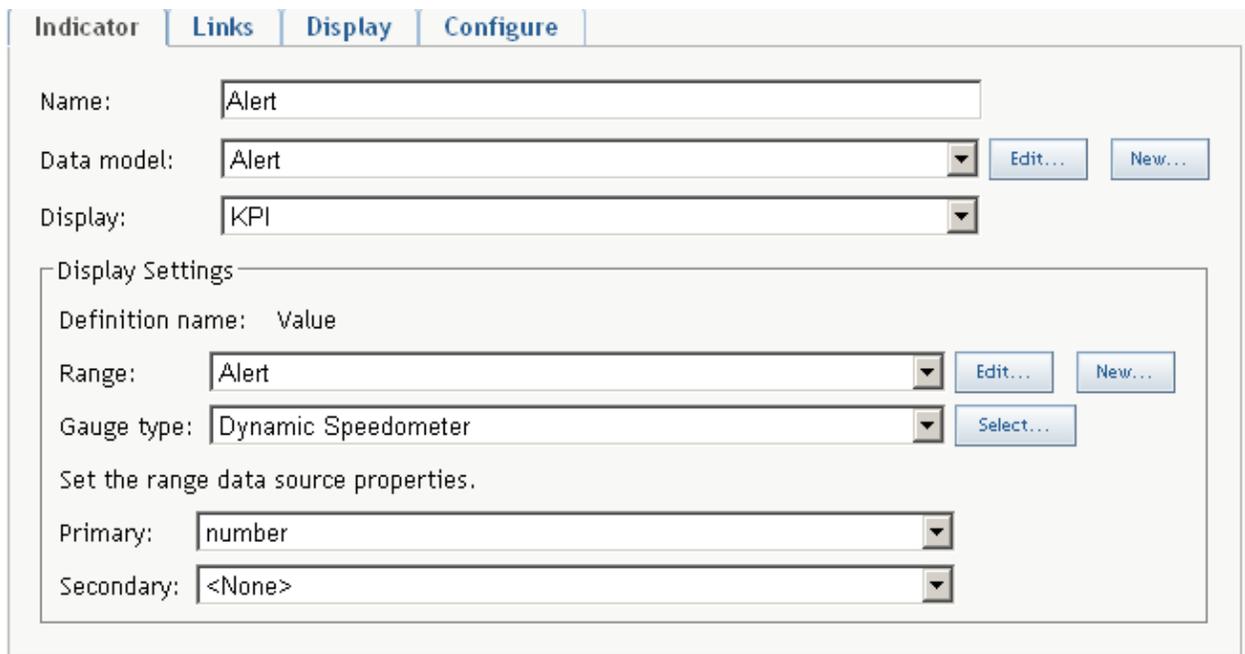
Figure 10. SAS BI Dashboard Indicator Definition

In this example, we create an indicator Alert. It associates the alert data model, alert range, and a Dynamic speedometer gauge together. This indicator will surface the number of models in alert status to customers.

4. Organize the indicators as dashboard.
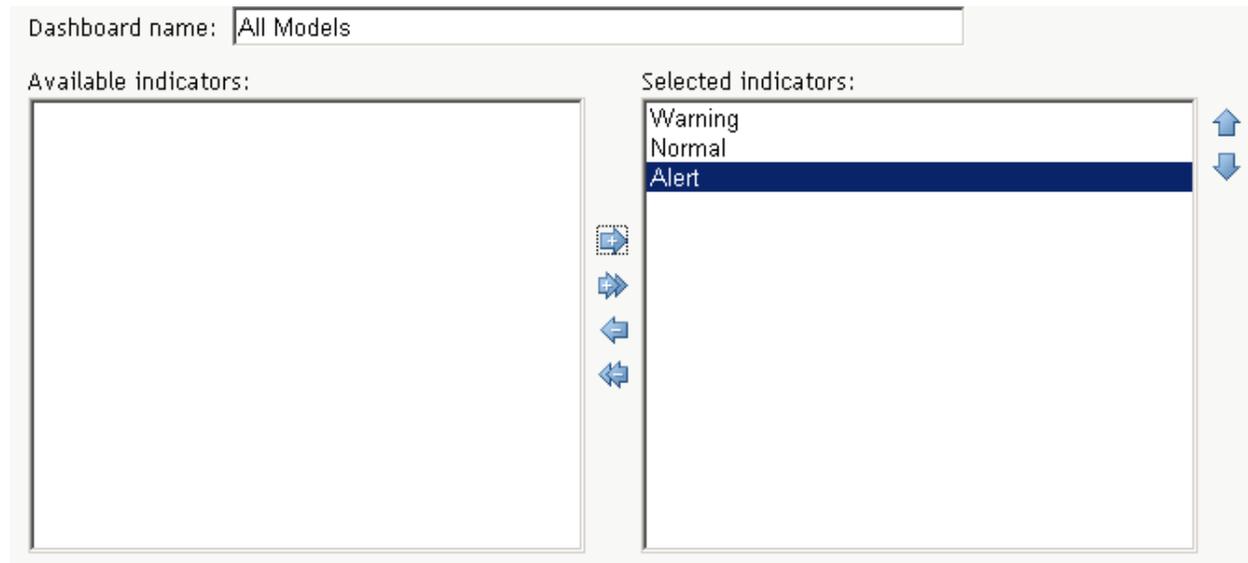
Figure 11. SAS BI Dashboard Creation

In this example, we create a dashboard "All Models." It has three indicators: Warning, Normal, and Alert.

The Warning indicator tells the user the number of models in a SAS Model Manager Repository that are in warning status.

The Normal indicator tells the user the number of models in a SAS Model Manager Repository that are normal.

The Alert indicator tells the user the number of models in a SAS Model Manager Repository that are in alert status.

We used the total performance status table created by the macro as data models for all three indicators. An SQL query is used to count the number of models in each category.

We can get an overview for all models status from this dashboard. This is the important information for us now.

5. Surface the dashboard as a portlet in SAS Information Delivery Portal to customers.

Add a SAS BI dashboard portlet to SAS Information Delivery Portal.
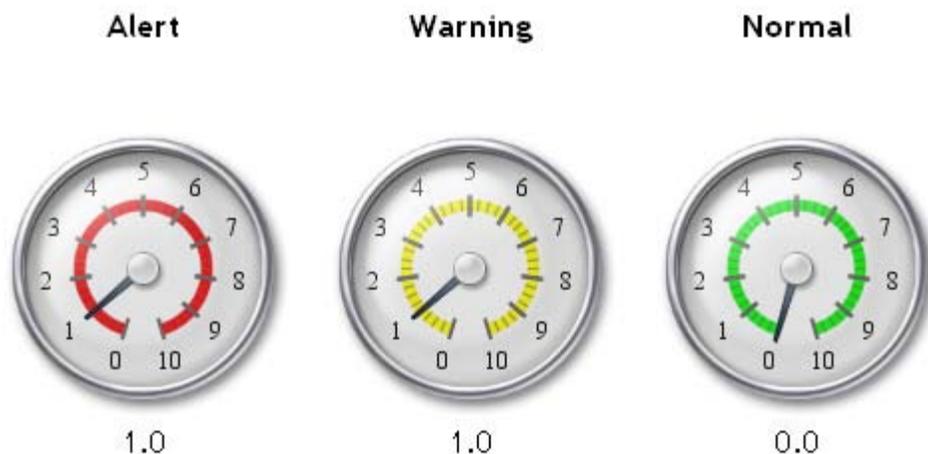
Figure 12. SAS BI Dashboard Configuration

Figure 13 shows a Model Monitoring Dashboard. It shows that one model is in alert status; one model is in warning status; and zero model is in normal status. You can click on any of the gauges and view a more detailed report for models in those categories.

Figure 13. Sample of Model Monitoring Dashboard

Based on the same data set, users can also customize data providers, link handlers, data visualization, and gauges to create their own dashboards. Please refer to BI Dashboard Architecture, Customization, and Development Guide for details.

CONCLUSION

This paper introduces the concepts of dashboard reports for predictive model performance monitoring. It then demonstrates a process of creating a SAS Information Delivery Portal portlet and a SAS BI Dashboard using a SAS macro that runs against a SAS Model Manager repository. It also describes rather portable techniques without being based on any application domains, such as credit scoring, marketing campaigns, insurance fraud detection, and so on. With application-specific data models, more

specific dashboard reports can be built in order to meet specific business needs. We believe that predictive model dashboard reporting tools will certainly evolve, as will the predictive model dashboard reporting processes. The macro described in this paper can be expanded to publish the generated reports to a SAS Folder automatically. Furthermore, the automation of predictive model dashboard reporting processes will be nicely integrated with processes of automated predictive model re-training, which is an important subject for future SAS Global Forums.

REFERENCES

Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*, 2nd Edition. Graphics Press.

ACKNOWLEDGMENTS

We would like to thank Jane (Jianping) Xue for helping us configure the SAS Information Delivery Portal and the SAS BI Dashboard, Elizabeth Maldonado for reviewing this article, and Sallie Fiore for providing invaluable editing service.

CONTACT INFORMATION

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

Jifa Wei, Principal Software Developer
SAS Institute Inc.
S6112 SAS Campus Dr.
Cary, North Carolina 27513
Work Phone: 919-531-7022
E-mail: jifa.wei@sas.com

Emily (Yan) Gao, Manager of ADG
SAS Institute Inc.
Beijing, China
Work Phone: +86 10 8319-3720
E-mail: yan.gao@sas.com

Frank (Jidong) Wang, Globalization Engineer
SAS Institute Inc.
Beijing, China
Work Phone: +86 10 8319-3802
E-mail: jidong.wang@sas.com

Robert Chu, Senior Manager, Software Development
SAS Institute Inc.
S6120 SAS Campus Dr.
Cary, North Carolina 27513
Work Phone: 919-531-7696
E-mail: robert.chu@sas.com