Top 10 Tips for SAS® Enterprise Miner™ Based on 20 Years' Experience

Melodie Rush, SAS Institute Inc., Cary, NC

ABSTRACT

Over the past 20 years that I have been using SAS® Enterprise Miner™ and helping analysts with it, I have learned and developed many tips and tricks for ease of use, productivity, and just plain clever implementation. In this presentation, I cover the evolution of SAS Enterprise Miner from the original SAS/AF® software application to the current version that integrates with both open-source software and with SAS® Viya®. I share my top 10 tips for getting the most from using SAS Enterprise Miner, including sharing my favorite node that no one seems to know about and how to implement more complex modeling techniques.

INTRODUCTION

SAS® Enterprise Miner™ has been the proven data mining workbench for the past 20 years. Using it enables you to quickly create models, compare models, and create the score code for the winning model. In this paper, I cover 10 quick tips to help the novice to the expert user gain more insight about their data using SAS Enterprise Miner. These tips help with increasing productivity, learning about new nodes, and leveraging options to expand the functionality and knowledge gained.

BACKGROUND AND HISTORY

SAS Enterprise Miner was released in 1998 with the interface built in SAS/AF® (Figure 1). The first version of SAS Enterprise Miner was 2.01 released with SAS 6.12. The first release included Client Server for both Windows and UNIX, process flow diagrams with drag-and-drop capabilities based on the SEMMA (Sample, Explore, Modify, Model, and Assess) model development process, integrated model comparison, and the creation of SAS Score Code, including transformations. This first release had 15 nodes versus the 80+ that are available in the current version. This release even included nodes for decision trees, neural networks, and ensemble models.
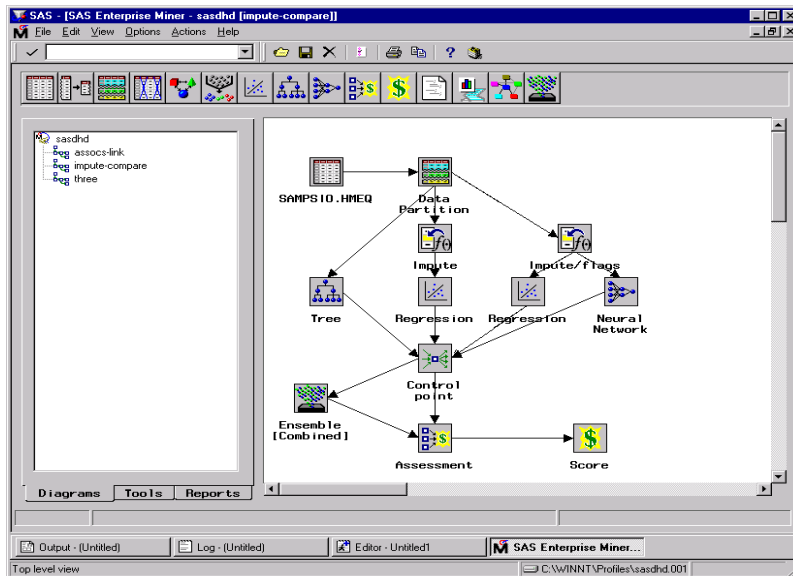

Figure 1. SAS Enterprise Miner Original Interface

Over the last 20 years, many milestones have been reached. Here are some of the highlights by year:

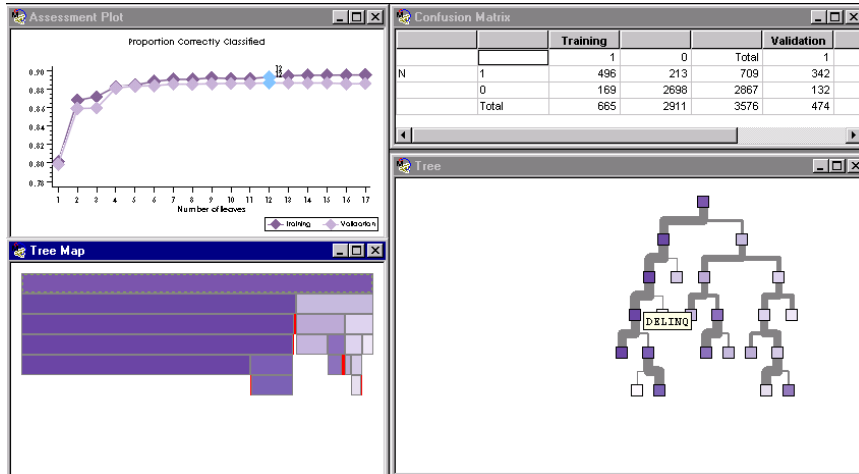- 2000: EM 4.0 – C and Java score code added, and the Tree Desktop Viewer (Figure 2).

Figure 2. Tree Desktop Viewer

- 2001: EM 4.1 – SAS 8.2 Link Analysis, Memory-Based Reasoning (MBR), and Time Series added.

- 2002: SAS® Text Miner Add-on released (Figure 3).

- 2003: EM 5.1 – SAS 9.1 interface rewritten to a rich Java client; parallel and batch processing, XML diagram exchange, model packages, graph explorer, credit scoring nodes.
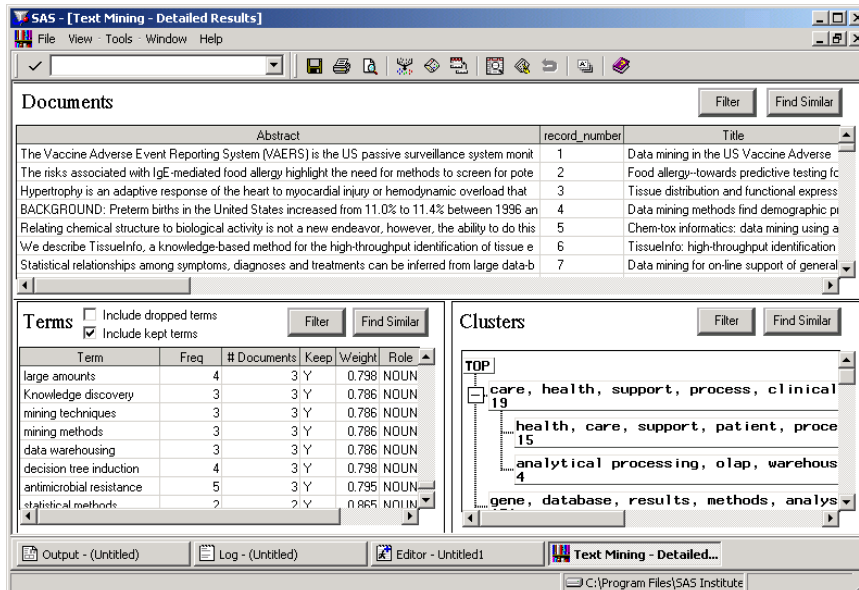

Figure 3. Text Miner Add-on

- 2005: EM 5.2 – SAS 9.1 Decision, Replacement, and SOM/Kohonen nodes, GRID processing, desktop release.

- 2007: EM 5.3 – SAS 9.1.3 group processing, gradient boosting, variable clustering, and hierarchical associations.

- 2009: EM 6.1 – SAS 9.2 File Import node, LARS, optimized score code generation, and native interactive decision trees. Rapid Predictive Modeling task introduced as experimental.

- 2010: EM 6.2 – SAS 9.2 Rapid Predictive Modeling task, SAS® Analytics Accelerator for Teradata.

- 2011: EM 7.1 – SAS 9.3 Survival node, Incremental Response node, Support Vector Machine (SVM), and creation of PMML score code.

- 2012: EM 12.1 – SAS 9.3 Time Series Data Mining nodes (TS Similarity, TS Exponential Smoothing, and TS Data Preparation) production and redesigned Interactive Grouping node.

- 2013: EM 12.3 – SAS 9.4 High-Performance tab added with several HP nodes, including nodes for Random Forest, Neural Networks, Decision Tree, Regression (Logistic and Linear), and GLM (Generalized Linear Model).

- 2013: EM 13.1 – More high-performance nodes (SVM, Principal Components, and Clustering), three New Time Series nodes (TS Dimension Reduction, Time Series Correlation, and TS Decomposition), Open Source Integration node, Register Model node, and Save Data node.

- 2014: EM 13.2 – HP Regression creates VIF (Variance Inflation Factor), support for SAP Hana.

- 2015: EM 14.1 – HP Bayesian Network node, HP Cluster supports automatic selection for number of clusters.

- 2016: EM 14.2 – SAS Viya Code node and support of Analytic Item Store (ASTORE).

- 2017: EM 14.3 – SAS Viya Code node rewritten to support CAS (SAS® Cloud Analytic Services).

The complete list of nodes available in the current release of SAS Enterprise Miner 14.3 is in Figure 4.

Figure 4. Nodes Available in EM 14.3

TIPS FOR PRODUCTIVITY

Three quick tips for productivity include how to find the node you want, what the available properties are for that node, and how to clone a diagram. Each of these tips accelerates model development.

TIP 1: HOW TO FIND THE NODE I WANT

Inexperienced users often struggle to find the nodes they need to build their data mining flow or diagram. The nodes are organized in the proven data mining process called SEMMA, which stands for Sample, Explore, Modify, Model, and Assess. Each tab on the toolbar at the top of the diagram workspace includes the appropriate nodes (Figure 5).
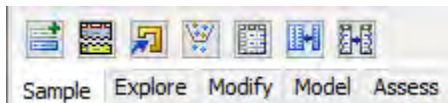

Figure 5. Sample Tab Nodes

For example, to add a decision tree to your diagram, click the **Model** tab (Figure 6).
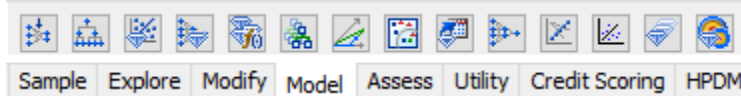
Figure 6. Model Tab Nodes

To discern which icon is for the decision tree, scroll across the nodes and position your pointer over the node to see a brief description. The first node is the AutoNeural (Figure 7).
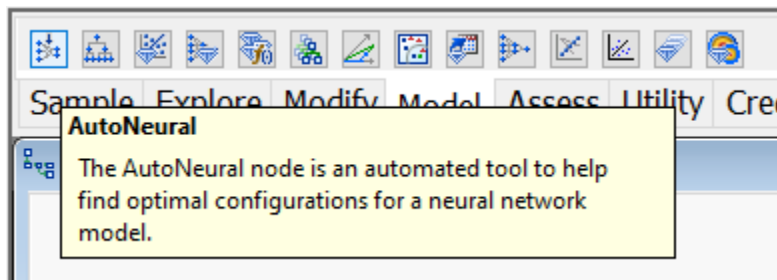
Figure 7. Tooltip for Each Node; AutoNeural Description Displayed

The second node is the Decision Tree (Figure 8).
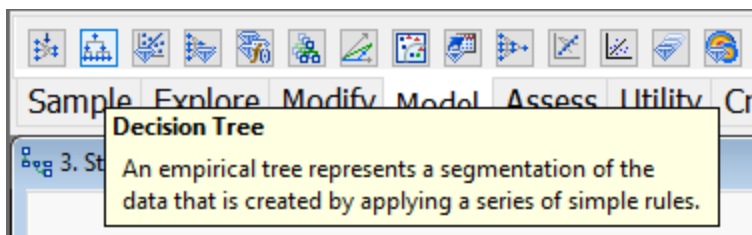
Figure 8. Tooltip for Each Node; Decision Tree Description Displayed

An additional tip: The nodes on each tab are in alphabetical order.

Another way to add a node is to right-click within the diagram you are building (Figure 9). At the top of the menu, select **Add Node**, and then select from the nodes organized by SEMMA. Note that the nodes are also in alphabetical order.

4

Figure 9. Add Node from Diagram Workspace

TIP 2: WHAT ARE THE AVAILABLE PROPERTIES FOR EACH NODE?

Now that you know how to find a node, you might want to know which properties are available for each node. Simply double-click a node on the toolbar. For example, double-click **Data Partition** on the **Sample** tab (Figure 10).
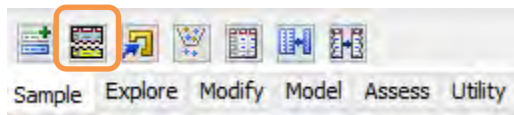
Figure 10. Data Partition Node

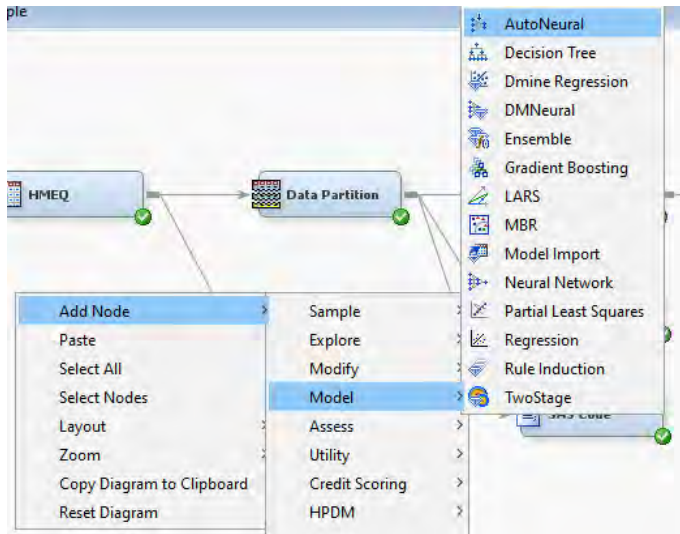The properties for the Data Partition node open in a separate window (Figure 11). This window enables you to see all the current property values and whether the property can be edited.

View	Property	Batch Name	Description	Type	Editable	Valid Values	Initial Value
General	Node ID	NODEID	Node Identifier	String	No		
	Imported Data	ImportSet	Set of tables imported by this node.	String	Yes		
	Exported Data	ExportSet	Set of tables exported by this node.	String	Yes		
	Notes	NotesFile	Enter notes for this node.	String	Yes		
	Variables	VariableSet	Variable Properties	String	Yes		
	Output Type	OutputType	Indicates if the node should create data set(s) or DATA step view(s).	String	Yes	Data, View	Data

Data Partition
Description: Partitions data into separate tables.
Component: Partition
Group: Sample

Figure 11. Properties of the Data Partition Node

An additional tip: More details are outlined in the SAS Data Mining and Machine Learning Community: https://communities.sas.com/t5/SAS-Communities-Library/SAS-Enterprise-Miner-shortcut-How-to-quickly-see-node-properties/ta-p/375805.

TIP 3: CLONE A PROCESS FLOW

Do you have a process flow that you want to reuse within your project? It doesn't have to be perfect to make a copy; sometimes we make copies because we are conducting trial and error or another team member would like to use a copy to build a new model faster based on what's already been defined and vetted. It's very easy to clone your process flow and replicate it in the same diagram workspace or a new diagram workspace in three easy steps:

1. Highlight the process flow by dragging your mouse across the process flow (Figure 12).
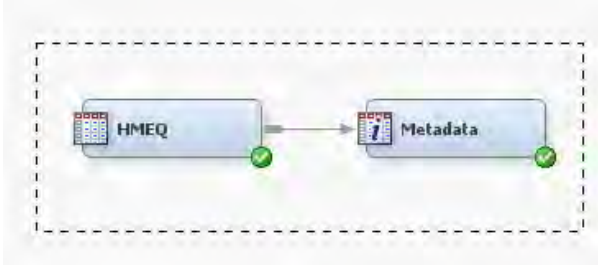
Figure 12. Simple Process Flow Selected

2. Right-click and select **Copy** or select CTRL+C to copy (Figure 13).
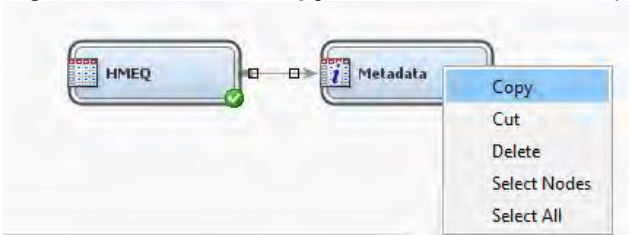
Figure 13. Right-Click to Copy

3. Click where you want to insert the process flow, and then right-click and select **Paste** or use CTRL+V to paste (Figure 14).
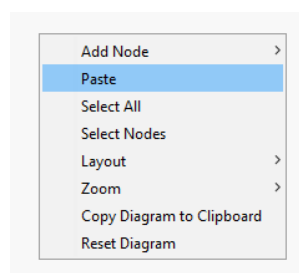
Figure 14. Right-Click to Paste

TIPS ON NODES

SAS Enterprise Miner currently has 80 nodes in the standard installation. If you also have the SAS Text Miner add-on and the Credit Scoring Add-on for SAS® Enterprise Miner™, these two together add an additional 11 nodes. The next three tips cover some of the new nodes, my favorite node that no one knows about, and the node that changes everything.

TIP 4: WHAT'S NEW

You might be like me and stick with what you know, so sometimes I miss new features and functionality when a new release becomes available. With each new release of SAS Enterprise Miner, new nodes are added. This tip is about the new nodes. The current version of SAS Enterprise Miner is 14.3 on SAS 9.4M5.

HPDM Tab and Nodes

Starting with SAS Enterprise Miner 12.3, there is a brand-new tab, HPDM (which stands for High-Performance Data Mining) with several new nodes (Figure 15). These nodes are optimized to run in a distributed environment, meaning the processing can be split among many processors to help minimize processing time. Nodes cover both data mining and machine learning algorithms.

Figure 15. High-Performance Data Mining Tab

In Version 14.3, these nodes are included:

- HP Bayesian Network Classifier
- HP Cluster
- HP Data Partition
- HP Explore
- HP Forest
- HP GLM
- HP Impute
- HP Neural
- HP Principal Components
- HP Regression
- HP SVM
- HP Text Miner
- HP Transform
- HP Tree
- HP Variable Selection

Programming Code Nodes

Two new nodes appear on the **Utility** tab to help incorporate programming code from both Open Source (R) and SAS® Viya® (Figure 16). These two programming nodes join the SAS Code node, which has been available in SAS Enterprise Miner for several releases.
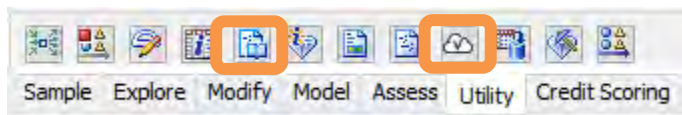

Figure 16. New Nodes for Open Source and SAS Viya Code

Open Source Integration Node

The Open Source Integration node highlighted first in Figure 16 enables you to use code from the R language inside SAS Enterprise Miner diagrams. This node allows for both supervised and unsupervised algorithms and PMML (Predictive Model Markup Language) and non-PMML R packages. You can compare SAS Enterprise Miner models with R models (Figure 17), ensemble SAS Enterprise models with R models, and create the corresponding SAS DATA step scoring code if the R model comes from a PMML-supported package. This node transfers data, metadata, and results automatically between SAS Enterprise Miner and R.

An additional tip: For more information, watch this video: [Using R in SAS Enterprise Miner.](#)
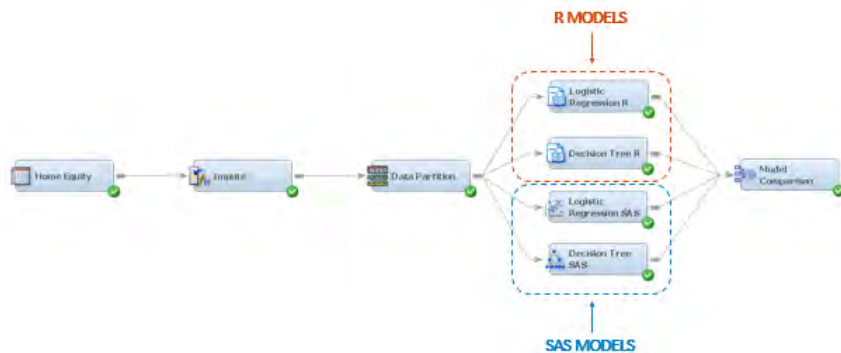
Figure 17. Comparing SAS and R Models Diagram

SAS Viya Code Node

The SAS Viya Code node (the second highlighted node in Figure16) is created to incorporate code that will be executed in SAS Viya and CAS (SAS Cloud Analytic Services). It allows you to include the new data mining and machine algorithms available in SAS® Visual Data Mining and Machine Learning as part of your SAS Enterprise Miner diagrams (Figure 18).

Figure 18. SAS Viya Code Nodes in Diagram

Saving and Sharing Results

Two more nodes recently added are the Register Model node and the Save Data node. Both nodes are located on the **Utility** tab (Figure 19). These nodes enable you to save and share your output and results.

Figure 19. Register Model and Save Data Nodes

Register Model Node

The Register Model node is highlighted first in Figure 19. The node enables you to register segmentation, classification, or prediction models to the SAS Metadata Server. Why register your models? Register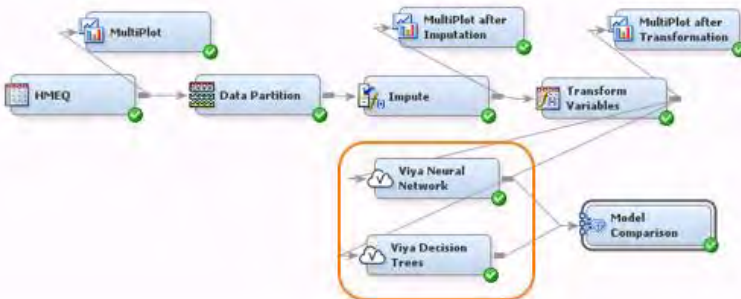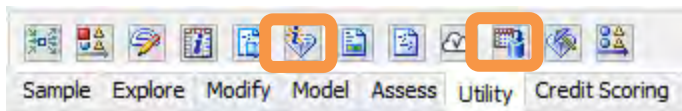ed models can be used and monitored by SAS® Decision Manager and SAS® Model Manager; they can easily score data in SAS® Enterprise Guide®; and they can score or compare models in SAS Enterprise Miner. Using the Register Model node extends and expands your models' intelligence. In previous versions of SAS Enterprise Miner, registering models took several steps; now registering can be done within the diagram using the Register Model node.

An additional tip: The Register Model node provides a model registration mechanism that can run in batch code.

The Register Model node enables users to select the path to register, the name of the model, a model description, and the data mining function of segmentation, classification, or prediction (Figure 20).

Train	
Repository Path	...
Model Name	Propensity Model
Model Description	...
Mining Function	Classification

Figure 20. Properties of the Register Model Node

Save Data Node

The Save Data node is highlighted second in Figure 19. This node can be used after any node in the process flow diagram to save the training, validation, test, score, or transaction data. The data can be saved as a SAS data set, a JMP data set, and an Excel, CSV, or tab-delimited file. You can also opt to replace existing files, include all or a subset of observations, and all or selected data sets (Figure 21).

.. Property	Value
Train	
Output Options	
Variables	...
Filename Prefix	
Replace Existing Files	Yes
All Observations	Yes
Number of Observations	1000
Output Format	
File Format	SAS (.sas7bdat)
SAS Library Name	DM ...
Directory	...
Output Data	
All Roles	No
Select Roles	...

Figure 21. Properties of the Save Data Node

TIP 5: MY FAVORITE NODE THAT NO ONE KNOWS ABOUT

Over the years, I have asked many SAS Enterprise Miner users if they use this one node and usually the response is no. Do you like to document your SAS processes? For some of you, the answer is yes, but for most the answer is no. Either way, this node helps you easily document your SAS Enterprise Miner process flow diagrams. Which node is it? It's the Reporter node located on the **Utility** tab (Figure 22).
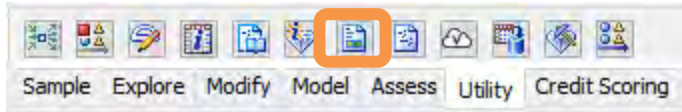
Figure 22. Reporter Node

The Reporter node creates a .pdf or .rtf file to document the entire process flow. It includes an image of the diagram (Figure 23), detailed information about each node included in the diagram (Figure 24), and output from each node.

SAS Enterprise Miner Report
Process Flow Diagram

Figure 23. Process Flow Diagram in a Report

Role	Level	Frequency Count	Name
TARGET	BINARY	1	BAD
INPUT	BINARY	1	REASON
INPUT	INTERVAL	7	CLAGE CLNO DEBTINC LOAN MORTDUE VALUE YOJ
INPUT	NOMINAL	4	DELINQ DEROG JOB NINQ

Figure 24. Data Source Node Settings in a Report

An additional tip: If you have included notes in your nodes or used the SAS Code node to create output in report or graphic format, those notes are included in the report as well. The reports created can range from 30 to 100+ pages depending on the complexity of your process flow diagram. I recommend that you end each flow with a Reporter node so that automatic documentation is created.

TIP 6: THE NODE THAT CHANGES EVERYTHING

One of the most valuable nodes in SAS Enterprise Miner is the Metadata node. This node is on the **Utility** tab (Figure 25).
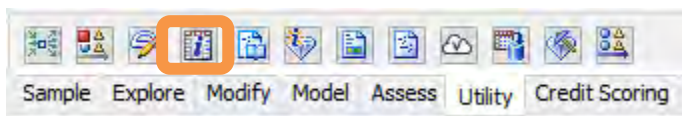
Figure 25. Metadata Node

This node enables you to change the metadata information in your process flow diagram. You can modify any attribute such as variable roles, measurement levels, and so on. You can also use it to merge predecessor nodes. An example of this is in Tip 9.

Have you ever wanted to use your settings from one data set in another data set? I discovered one of the best tips for the Metadata node in the SAS Data Mining and Machine Learning Community. The tip is to always use a Metadata node after a data set (Figure 26). Doing this enables you to capture the settings

for your variables so that you can apply them to new data or to data in a different diagram. This allows for repeatability and consistency when you set up and use your data. It's also a great time saver.

Figure 26. Metadata Node Example

To use the Metadata node in your flow:

1. Create a new diagram.

2. Add your data source using basic settings in the Data Source Wizard.

3. Add a Metadata node:

    • Set up all your roles and levels.

4. Copy and paste the Metadata node to another data set.

TIPS FOR USING OPTIONS

One of the most powerful capabilities of SAS Enterprise Miner is its ability to change options and properties for the process flow diagram and nodes. All the nodes come with what I like to call "smart properties" so that they will run without making any changes. The next four tips are about changing the options or properties to gain even more insight from your data and models.

TIP 7: HOW TO GENERATE A SCORECARD

Did you know that you can create scorecards for your models? With just a couple of modifications to the Reporter node, you can generate a scorecard that emphasizes which variables and values are important (and which are not).

First, what is a scorecard? A scorecard displays your model in such a way that quickly reveals which variables are important and which values are important (Figure 27). The summary ranges from 0 to 1,000. The closer to 1,000 the more likely your event will happen. The closer to 0 the less likely. In the following scorecard, if the customer purchased two or more blankets we would assign them 61 points; 2 domestic products 18 points; 4 or more Heat products 106, and so on. Add all the highlighted numbers and you get 61 + 18 + 106 + 32 + 74 + 77 + 113 + 296 = 777. In this case, we would say that the customer was likely to purchase from our new campaign (the event we are predicting in this model).

		Scorecard Points
Blankets Purch.	1: LOW - 0.5	0.00
	2: 0.5 - 1.5	13.00
	3: 1.5 - HIGH	61.00
Domestic Prod.	1: LOW - 0.5	0.00
	2: 0.5 - 1.5	18.00
	3: 1.5 - 4.5	46.00
	4: 4.5 - HIGH	195.00
HEAT	1: LOW - 1.5	0.00
	2: 1.5 - 2.5	36.00
	3: 2.5 - 3.5	77.00
	4: 3.5 - HIGH	106.00
Kitchen Prod.	1: LOW - 0.5	0.00
	2: 0.5 - 1.5	32.00
	3: 1.5 - HIGH	62.00
Outdoor Prod.	1: LOW - 0.5	0.00
	2: 0.5 - 3.5	19.00
	3: 3.5 - HIGH	74.00
Promo: 1-7 mon.	1: LOW - 6.5	93.00
	2: 6.5 - 10.5	77.00
	3: 10.5 - 67.5	46.00
	4: 67.5 - HIGH	0.00
Recency	1: LOW - 71.5	113.00
	2: 71.5 - 142	84.00
	3: 142 - 393.5	45.00
	4: 393.5 - HIGH	0.00
Telemarket Ind.	NO	296.00
	YES	0.00

Figure 27. Scorecard

Also, the model tells us the more Blankets, Domestic, Heat, Kitchen, and Outdoor products purchased, the more likely the customer will purchase from our next campaign. It also indicates the more recently the customer received a promotion and the more recently they purchased, the more likely they are to purchase from our next campaign. The model also indicates that if the customer has not received a telemarketing call, they are more likely to purchase from the next campaign.

How do you produce a scorecard in SAS Enterprise Miner? Simply change the properties on the Reporter node. First, the Reporter node needs to follow a Score node. Second, change the Nodes property to Summary (Figure 28) in the Reporter node properties.

Train	
Document Format	PDF
Nodes	Summary
Font size	
Summary Report Options	
Basic Reports	Yes
Summarization	Yes
Variable Ranking	Yes
Classification Matrix	Yes
Cross Tabs	Yes
Lift Chart	Yes
Fit Statistics	Yes
Model Comparison	Yes

Figure 28. Reporter Node Properties for Scorecard

TIP 8: HELP, I HAVE MORE THAN 512 LEVELS

One of the more common error messages in SAS Enterprise Miner is "Maximum target levels of 512 exceeded" (Figure 29). Here are the two questions that it generates:

1. What does this error message mean?
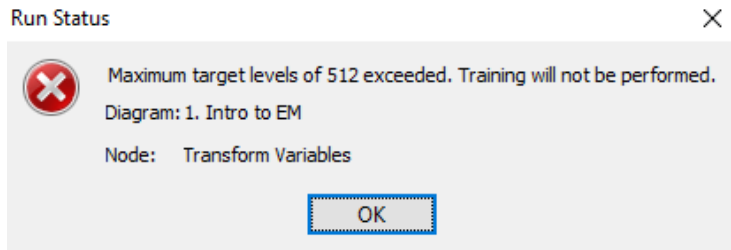2. How can I override or overcome it?

Figure 29. Error Message for Maximum Levels Exceeded

This error occurs when you have a categorical input variable (nominal or ordinal) that has 512 or more distinct values (called cardinality). An example might be a ZIP code. SAS Enterprise Miner set this default for a couple of reasons. It prevents novice users from accidentally using a variable with a bunch of levels because this takes additional processing time and is often unintentional by the user (that is, using a unique ID variable as in input). For example, a ZIP code might have as many as 40,000 levels. If a ZIP code is used as a categorical input into our regression model, the model would create 39,999 parameters to represent the 40,000 levels. Using a neural network model, the number of parameters increases quickly depending on the architecture and number of hidden layers. Having this many parameters to estimate also causes additional issues with sparsity and convergence.

Sometimes it might make sense to use these high cardinality variables in our models. The default can be overridden by changing the EM_TRAIN_MAXLEVELS macro variable to a higher value. There are two ways to do this:

Change macro variable in properties

1. Click your project name in the project window.
2. Scroll down to the project properties.
3. Click the Project Macro Variables ellipsis (Figure 30).

Figure 30. Project Properties

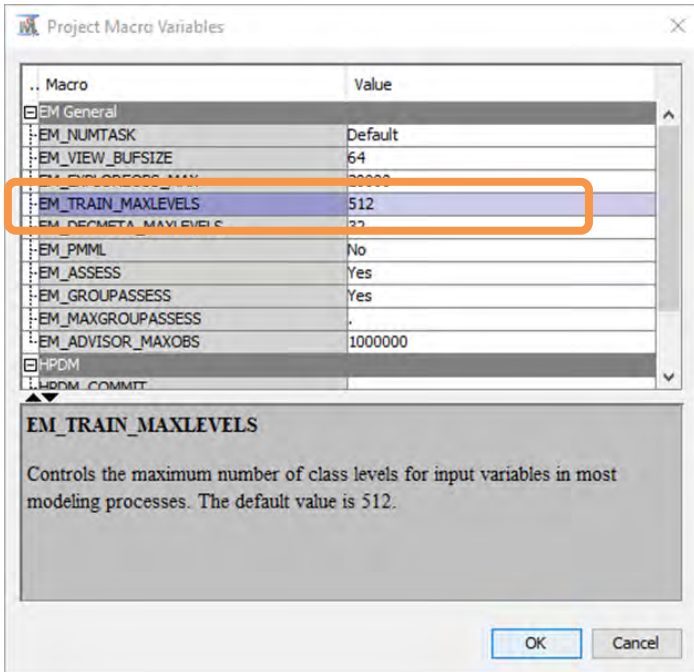4. Change the value for EM_TRAIN_MAXLEVELS (Figure 31).

Figure 31. Project Macro Variable Values

Change macro variable in project start code

1. Click your project name in the project window.

2. Scroll down to the project properties.

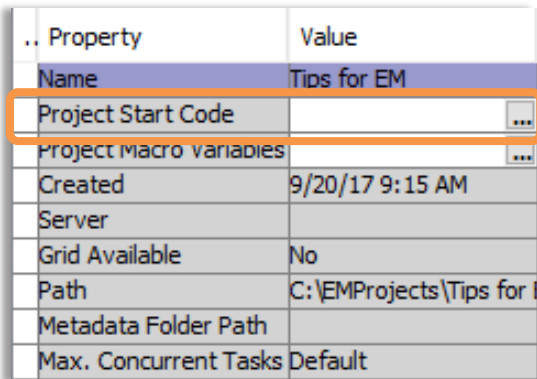3. Click the Project Start Code (Figure 32).

Figure 32. Project Properties

4. Add the statement %let EM_TRAIN_MAXLEVELS = *MYVALUE;* (Figure 33).
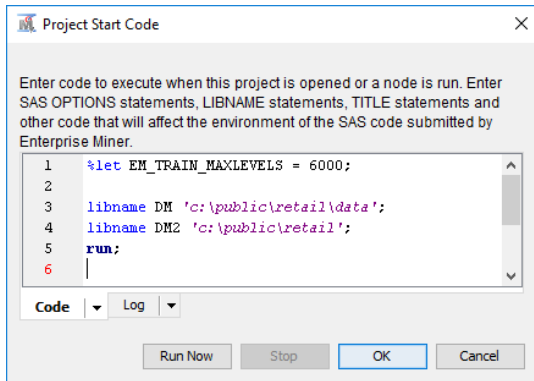
5. Click **Run Now**.

14

Figure 33. Project Start Code Window

TIP 9: WHICH VARIABLE SELECTION METHOD SHOULD I USE?

SAS Enterprise Miner has several variable selection methods such as Stepwise, Forward, Backward, Decision Trees, R^2, Chi-square, Random Forest, and more. The question becomes which one should be used. The good news is you don't have to choose just one. You can use multiple methods and combine the results using the Metadata node from Tip 6 (Figure 34).

Figure 34. Example of Metadata Node Variable Selection

The preceding example shows using the LARS, Variable Selection, Variable Clustering, HP Variable Selection, and Decision Tree for variable selection. Connect all the nodes to the Metadata node and navigate to the properties to specify how you want to combine the results (Figure 35). Here are some of the choices:

- None – keeps the original metadata and makes no changes based on the variable selection methods of the previous nodes.

- Any – a variable is set to rejected if any of the previous variable selection nodes rejected it.

- All – a variable is set to rejected if all of the previous variable selection nodes rejected it.

- Majority – a variable is set to rejected if the majority of the previous variable selection nodes rejected it.

Figure 35. Metadata Node Properties

An additional tip: More details are outlined in the SAS Data Mining and Machine Learning Community. There is also a SAS Ask the Expert Session on Variable Selection Using SAS Enterprise Guide and SAS Enterprise Miner.

TIP 10: HOW DO I INTERPRET MY NEURAL NETWORK?

Neural networks are notoriously hard to interpret. This tip shows how to use a decision tree to create an alternate or proxy interpretation.

First, run you Neural Network, connect a Metadata node, and then connect a Decision Tree node (Figure 36).

Figure 36. Example of Metadata Node Neural Network

Click the Metadata node, and then click the ellipsis next to **Variables→Train** (Figure 37).

Figure 37. Metadata Node Properties

Change the Prediction variable to be your Target and the original Target variable to be rejected (Figure 38). By doing this, the decision tree is using the predicted values from the neural network as the Y or Target variable (what it is predicting).

P_BAD1	N	Default	Prediction	Target
BAD	N	Default	Target	Rejected
DEBTINC	Y	Default	Rejected	Default
CLAGE	Y	Default	Rejected	Default
CLNO	Y	Default	Rejected	Default
DELINQ	Y	Default	Rejected	Default
DEROG	Y	Default	Rejected	Default
IMP_CLNO	N	Default	Input	Default
F_BAD	N	Default	Classification	Default
IMP_CLAGE	N	Default	Input	Default
IMP_DEBTINC	N	Default	Input	Default

Figure 38. Settings for Metadata Node Variable

The resulting decision tree (Figure 39) shows variables that are important to the predictive value of the neural network. A simplified tree is shown at the bottom of Figure 39. Credit Line Age (CLAGE) and Debt to Income Ratio (DEBTINC) are the two most important variables.
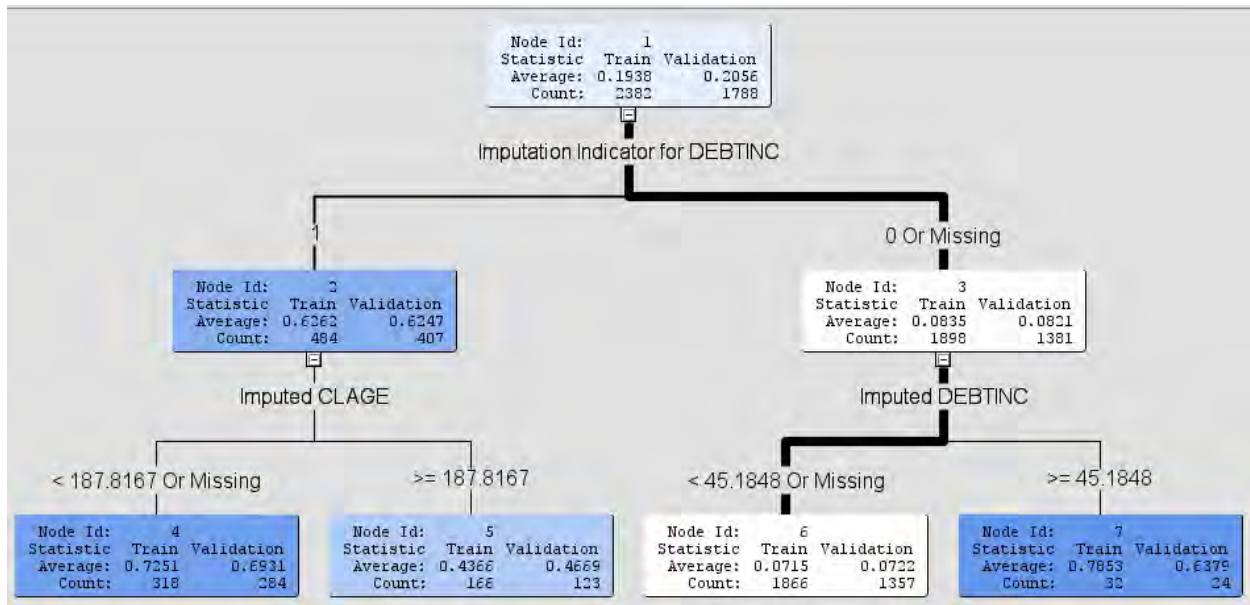
Figure 39. Decision Tree Based on Neural Network Predictors

BONUS TIP

Because SAS Enterprise Miner is loaded with functionality, there is a wealth of resources to help you learn and exploit it. Here are some of my favorites:

- The SAS Data Mining and Machine Learning Community available at https://communities.sas.com/t5/SAS-Data-Mining-and-Machine/bd-p/data_mining has new tips added each month. Plus, it's a great place to ask any questions you have and see what others are asking and solving.

- The SAS Enterprise Miner Learn page available at https://www.sas.com/en_us/learn/software/enterprise-miner.html has resources for new users and advanced tips for more experienced users, including videos, documentation, and examples.

- The Ask the Expert series available at http://support.sas.com/training/askexpert.html includes live session and recorded videos where you can witness SAS Enterprise Miner in action. In these one-hour sessions, attendees can ask SAS analysts questions. Past recorded sessions are available on demand. Here are the current sessions (with new ones added often):
  - SAS Enterprise Miner: Getting Started
  - Ensemble Models and Partitioning Algorithms in SAS Enterprise Miner
  - Model Selection Techniques in SAS Enterprise Guide & SAS Enterprise Miner
  - Variable Selection Using SAS Enterprise Guide and SAS Enterprise Miner
  - Data Mining Tasks with SAS Enterprise Guide
  - SAS Text Miner: Getting Started

- Help within SAS Enterprise Miner is available by clicking the book with a ? icon or selecting **Help→Contents** from the menu. This Help includes a node reference guide that gives detailed information about all the nodes, including examples (Figure 40).
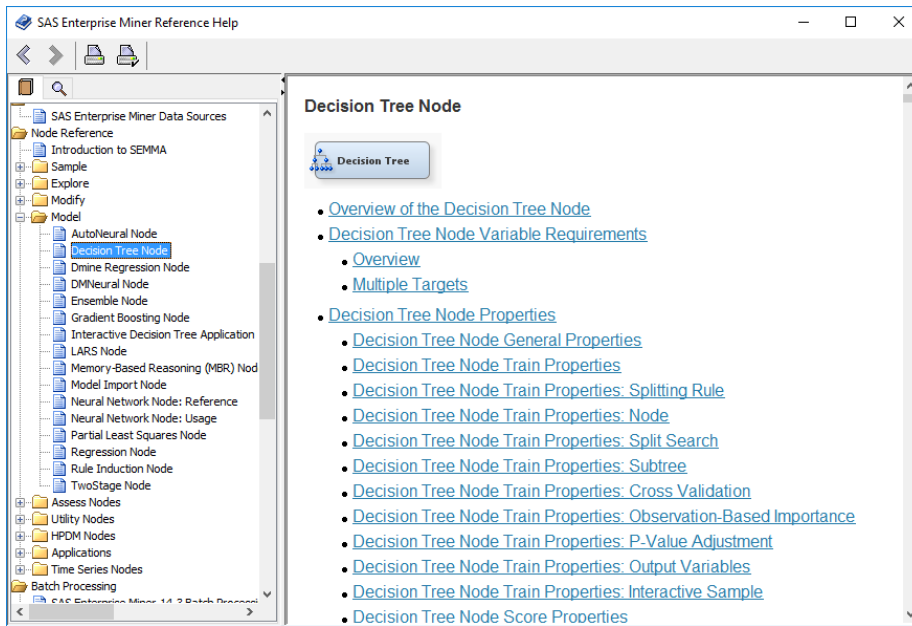
Figure 40. In-Product Node Reference Help

- The SAS Enterprise Miner documentation includes What's New, Getting Started, and administration information available at http://go.documentation.sas.com/?docsetId=emref&docsetTarget=titlepage.htm&docsetVersion=14.3&locale=en.

- Github.com is a wonderful place to find and share process flow diagrams. SAS has a library of process flow diagrams to help you learn by example available at https://github.com/sassoftware/dm-flow. Here is a video with instructions for using these process flow diagrams: https://www.youtube.com/watch?v=oSLrkvQH7iU.

CONCLUSION

SAS Enterprise Miner is a powerful tool for conducting data mining and machine learning projects. The tips shared in this paper enable users to gain more insight quicker by being more productive, to use new nodes, and to modify options and properties to leverage even more efficiency and knowledge.

The 10 tips shared in this paper are just the tip of the iceberg. You can find more tips by referencing the links provided in the Bonus Tip section. Becoming active on communities.sas.com yields even more tips, and you can share your tips too.

REFERENCES

- Ask the Expert series available at http://support.sas.com/training/askexpert.html

- Github.com available at https://github.com/sassoftware/dm-flow

- SAS Data Mining and Machine Learning Community available at https://communities.sas.com/t5/SAS-Data-Mining-and-Machine/bd-p/data_mining

- SAS Enterprise Miner documentation available at http://go.documentation.sas.com/?docsetId=emref&docsetTarget=titlepage.htm&docsetVersion=14.3&locale=en

- SAS Enterprise Miner Learn page available at https://www.sas.com/en_us/learn/software/enterprise-miner.html

VIDEOS

- Deep Learning in SAS Enterprise Miner
    - https://www.youtube.com/watch?v=HOEqvyyuPrk
- Getting Started with SAS Enterprise Miner Tutorial Videos
    - https://www.youtube.com/playlist?list=PLVBcK_IpFVi-xzvJiOlf33UvVbRoLRu0z
- How to Execute a Python Script in SAS Enterprise Miner
    - https://www.youtube.com/watch?v=GROwni8nw64
- Learn by Example with SAS Enterprise Miner Templates
    - https://www.youtube.com/watch?v=oSLrkvQH7iU
- The New HP GLM Node
    - https://www.youtube.com/watch?v=88qWDc1pGUU
- Random Forest and Support Vector Machines
    - https://www.youtube.com/watch?v=EOxwpnbFqIU
- Using R in SAS Enterprise Miner
    - https://www.youtube.com/watch?v=TbXo0xQCqDw

RECOMMENDED READING

- Chu, Robert, David Duling, and Wayne Thompson. 2007. *"*Best Practices for Managing Predictive Models in a Production Environment." *Proceedings of the 2007 SAS Global Forum Conference*. Cary, NC: SAS Institute Inc. Available http://www2.sas.com/proceedings/forum2007/076-2007.pdf.

- Collica, Randall S. 2017. Customer Segmentation and Clustering Using SAS Enterprise Miner, Third Edition. Cary, NC: SAS Institute Inc.

- Dean, Jared. 2014. Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners. New York: Wiley.

- de Ville, Barry, and Padraic Neville. 2013. Decision Trees for Analytics Using SAS Enterprise Miner. Cary, NC: SAS Institute Inc.

- Duling, David, Howard Plemmons, and Nancy Rausch. 2008. *"*From Soup to Nuts: Practices in Data Management for Analytical Performance. *Proceedings of the 2008 SAS Global Forum Conference*. Cary, NC: SAS Institute Inc. Available http://www2.sas.com/proceedings/forum2008/129-2008.pdf.

- Hall, Patrick. "The Open Source Integration node installation cheat sheet." https://communities.sas.com/docs/DOC-9988. 2014.

- Huang, Shunping. "Spectral Clustering in SAS Enterprise Miner Using Open Source Integration Node." Available https://communities.sas.com/docs/DOC-8011. 2014.

- Linoff, Gordon S., and Michael J. A. Berry. 2011. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, Third Edition. New York: Wiley.

- Wielenga, Doug. 2007. *"*Identifying and Overcoming Common Data Mining Mistakes." *Proceedings of the 2007 SAS Global Forum Conference.* Cary, NC: SAS Institute Inc. Available http://www2.sas.com/proceedings/forum2007/073-2007.pdf.

ACKNOWLEDGMENTS

reviewing this paper and providing constructive comments. Both of you have made this paper better. Wayne, thank you for your reference material on the history of SAS Enterprise Miner and for all your support over the years in answering my crazy and sometimes frantic questions.

I would also like to thank SAS Enterprise Miner R&D, Data Mining Technical Support staff, and all the coworkers and customers who have helped, supported, and walked beside me on this journey over the past 20 years. Thank you!

CONTACT INFORMATION

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

Melodie Rush
100 SAS Campus Drive
Cary, NC 27513
SAS Institute Inc.
Melodie.Rush@sas.com
http://www.sas.com

Twitter: https://twitter.com/Melodie_Rush
LinkedIn: https://www.linkedin.com/in/melodierush