



FOR IMMEDIATE RELEASE:

Press contact: Walt Wells, walt(at)dmg(dot)org

The Data Mining Group releases PMML v4.3

The Data Mining Group (DMG), a vendor-led consortium of companies and organizations developing standards for statistical and data mining models, announced the general availability of version 4.3 of the Predictive Model Markup Language (PMML).

Chicago, IL 8/2/2016 – The Data Mining Group is proud to announce the release of PMML v4.3. PMML is an application and system independent XML interchange format for statistical and data mining models. The goal of the PMML standard is to encapsulate a model independent of applications or systems using an XML configuration file so that two different applications (the PMML Producer and Consumer) can use it.

"The PMML standard delivers true interoperability, enabling machine learning and predictive models to be deployed across IT platforms," says Michael Zeller, CEO of Zementis, Inc. "A common standard ensures efficiency and fosters collaboration across organizational boundaries which is essential for data science to scale beyond its current use cases. With its latest release, PMML has matured to the point where it not only has extensive vendor support but also has become the backbone of many big data and streaming analytics applications."

As predictive analytics and data mining continue to gain momentum in industry and business applications across the internet of things, industrial internet, and instrumented environments, open standards like PMML provide a key platform to operationalize analytic models. Though the surge of interest in predictive analytics is relatively new, the Data Mining Group and the PMML standard have been supporting the data science community nearly 20 years. PMML v0.7 was first introduced in 1997. DMG members active in the PMML working group that developed v4.3 include IBM, SAS, Zementis, Salford Systems, and NIST.

"We at IBM are very happy to see PMML 4.3 release adding significant new model classes for Bayesian Networks," says Jean-Francois Puget, PhD, a Distinguished Engineer and Chief Architect at IBM Analytics Solutions. "Those new features will help us to further improve client experience. IBM has provided continued support for PMML since its inception. We use it both as an internal model and transformation representation, and as a way to import/export models between our products and to third party platforms."

Jonathan Wexler, Principal Product Manager of SAS Analytics Product Management says, "Time to value is critical for analytical applications. Using SAS Enterprise Miner, users can not only generate PMML, but consume PMML as part of a governed, automated system. SAS Model Manager enables users to both consume and manage PMML models, in addition to other Open Source models. We are excited about the latest release of PMML."

"NIST is pleased to have contributed to this new version of the already adopted PMML standard. We are particularly excited about the two new probabilistic models that appear in PMML v4.3: Gaussian Process Regression and Bayesian Networks. These new models provide two critical pieces of information never available before - confidence bounds and distribution for the predictive estimations. Both are needed to provide the foundation for uncertainty quantification analysis. We believe that the two models expressed in PMML will aid predictive modeling and improve decision-making in engineering, manufacturing, and other industries," says Tina Lee, Project Leader of the Data Analytics for Smart Manufacturing Systems Project in National Institute of Standards and Technology's (NIST) Smart Manufacturing Systems Design and Analysis Program.

"Salford Systems is proud to have been a contributing member of the PMML community since 2001, serving actively on the development team for many of the last 15 years," says Dr. Dan Steinberg, Founder and CEO of Salford Systems. "Much of our work has involved enhancing the representation of decision trees, and ensembles, to include the Transformations specification, which facilitates such things as linear combination splits in decision trees, and the basis functions used by MARS. We applaud the contributing members of the DMG for their superb work in the construction of this important set of standards and look forward to implementing the latest versions in our own software."

Some of the elements that are new to PMML v4.3 include:

- Improved support for post-processing, model types, and model elements
- New models for Gaussian Process and Bayesian Networks
- New built-in functions for multiple probability distributions

A full list of updates can be found at: <http://dmg.org/pmml/v4-3/Changes.html>

The DMG has actively partnered with ACM SIGKDD (<http://www.kdd.org/>) to promote awareness of available standards for predictive analytics as part . There will be a Special Session on "Standards in Predictive Analytics" on Tuesday, Aug 16th during the KDD 2016 conference in San Francisco, CA.

About the Special Session at KDD2016:

This special applied data science session on Standards for Predictive Analytics will cover available standards, discuss their adoption, and explore any needs unmet by existing standards. This event forms part of a broader SIGKDD Standards Initiative and will be of particular interest to KDD attendees focused on the operational application of data mining.

For more information, visit: <http://dmg.org/kdd2016.html>

About DMG:

The Data Mining Group (DMG) is an independent, vendor led consortium that develops data mining standards, such as the Predictive Model Markup Language (PMML) and Portable Format for Analytics (PFA).

For more information about the Data Mining Group and the PMML standard, go to:

www.dmg.org

#####

Contact:

Data Mining Group:
Walt Wells Walt(at)dmg(dot)org Data Mining Group 400 Lathrop Avenue River Forest, IL 60305