

Paper 009-29

High-Dimensional Data Mining Techniques

David Duling, SAS Institute Inc.

ABSTRACT

A typical genomics microarray observation may contain ten to twenty thousand attributes. Applications such as computer network monitoring, and digital intelligence produce similar sized data. Without clever strategies, this data stretches the capabilities of traditional linear algebra-based models. This presentation will examine techniques for building predictive models from such data including data aggregation, variable selection and pattern recognition, and distributed processing.

NOTE

Due to the fact that this paper was a late addition to the SUGI 29 paper schedule, a final paper is not available to be included in the *Proceedings*.

CONTACT INFORMATION

David Duling
SAS World Headquarters
SAS Campus Drive
Cary, NC 27513
919-531-5267
919-677-4444
david.duling@sas.com
www.sas.com
support.sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.