

BIOVIA PIPELINE PILOT ANALYTICS AND MACHINE LEARNING COLLECTION

Datasheet

Analytics is the key to unlocking meaningful patterns in data. In science based industries it requires awareness of complex data types to process large chemical, genomic, proteomic, textual and image data streams. BIOVIA Pipeline Pilot's capabilities harness multiple disciplines to help organizations understand complex systems and decipher everything from molecular behavior to optimal production processes.

Simplify your data science workflow. The Analytics and Machine Learning Collection for Pipeline Pilot gives you the tools for everything from data ingestion, cleaning and exploration, to model building, validation, deployment, optimization, and design of future experiments – all in a single environment.



BENEFITS OF BIOVIA PIPELINE PILOT ANALYTICS AND MACHINE LEARNING

FAST, EASY, AND ROBUST

- Build on a general and reliable framework for model development, validation, and deployment
- Utilize a comprehensive set of pre-designed prediction algorithms or those customized by in-house experts
- Data pipelining architecture scales to enormous data sets
- Use code-optional drag and drop, or write your own scripts in R or Python

PURPOSE-BUILT FOR SCIENCE

- Our algorithms are designed to work with various numerical, chemical, biological, textual, and image data types
- Employ custom, domain-specific analyses with a library of curated scientific component collections

VALIDATED MODELS

- Develop superior models with our built-in validation techniques
- Easily interpret model results with interactive visualizations
- Increase your odds of success with tools that ensure appropriate models are applied for various prediction tasks

MODEL DEPLOYMENT OPTIONS

- All models are automatically saved in Pipeline Pilot, easily accessible for inference, retraining, sharing, and deployment
- Deploy models across BIOVIA and 3rd party applications or as web services to achieve faster, more efficient processes to drive innovation

KEY CAPABILITIES

- Apply any of 15+ machine learning (ML) methods to your scientific and engineering data
- Merge, join, characterize, and clean your data sets
- Perform exploratory analysis, including PCA, clustering, and multi-dimensional data visualization
- Build fast, scalable Bayesian classification models
- Use the GFA method's genetic algorithm for variable selection and building regression ensemble models
- Build accurate, easy-to-use RP Forest regression and classification models
- Use R-based ML methods such as support vector machines, neural networks, and XGBoost without writing R scripts
- Use Python ML libraries including scikit-learn and TensorFlow
- Employ the ML framework for cross-validation, hyperparameter tuning, and variable importance assessment for any type of model

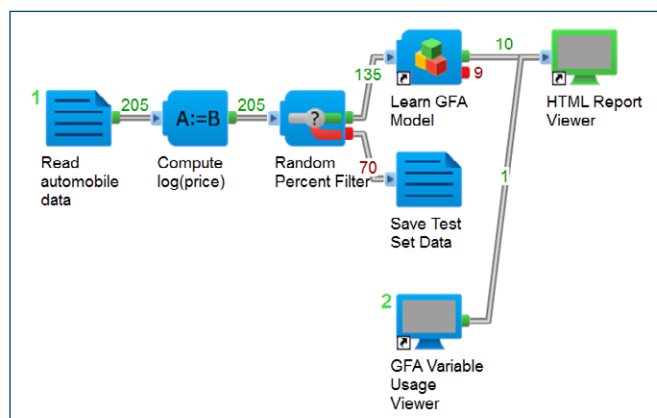


Fig 1. An example machine learning protocol

- Use regression and classification model evaluation viewers to assess and compare model test set performance
- Use built-in applicability domain measures and error models to assess sample-specific prediction confidence
- Apply Pareto optimization to multi-objective optimization problems

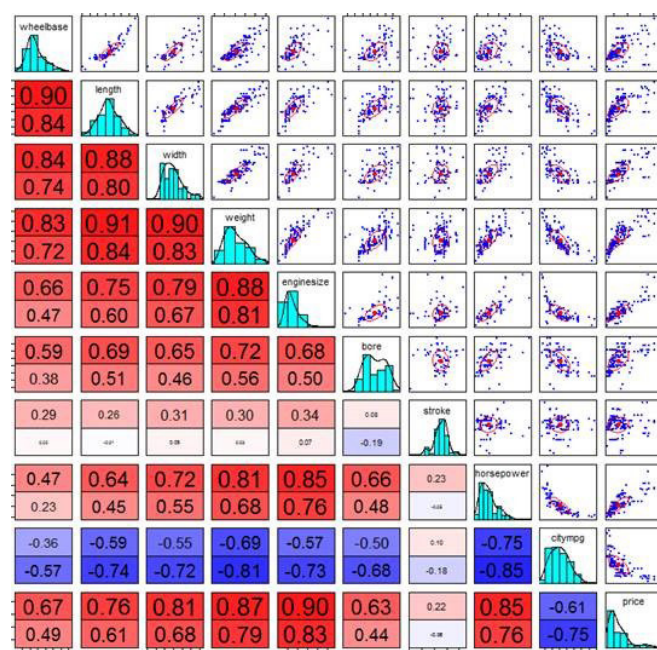
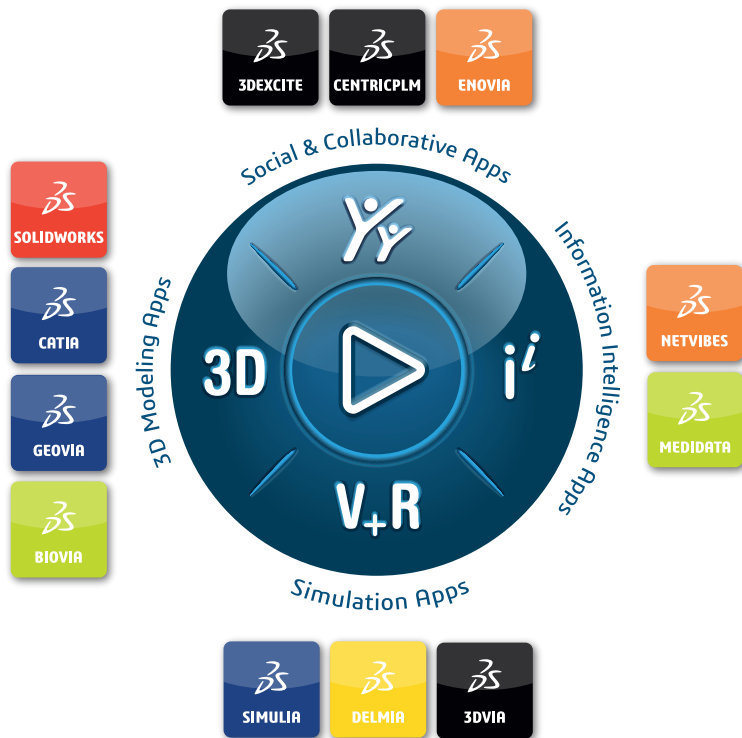


Fig 2. Even simple characterization analysis can yield valuable insight. Our R Pairs Plot displays pairwise relationships within multivariate data and helps you visualize strongly or weakly correlated variables.



Our 3DEXPERIENCE® platform powers our brand applications, serving 11 industries, and provides a rich portfolio of industry solution experiences.

Dassault Systèmes, the 3DEXPERIENCE Company, is a catalyst for human progress. We provide business and people with collaborative virtual environments to imagine sustainable innovations. By creating 'virtual experience twins' of the real world with our 3DEXPERIENCE platform and applications, our customers push the boundaries of innovation, learning and production.

Dassault Systèmes' 20,000 employees are bringing value to more than 270,000 customers of all sizes, in all industries, in more than 140 countries. For more information, visit www.3ds.com.

Europe/Middle East/Africa
Dassault Systèmes
10, rue Marcel Dassault
CS 40501
78946 Vélizy-Villacoublay Cedex
France

Asia-Pacific
Dassault Systèmes K.K.
ThinkPark Tower
2-1-1 Osaki, Shinagawa-ku,
Tokyo 141-6020
Japan

Americas
Dassault Systèmes
175 Wyman Street
Waltham, Massachusetts
02451-1223
USA

DS DASSAULT SYSTEMES | The **3DEXPERIENCE®** Company