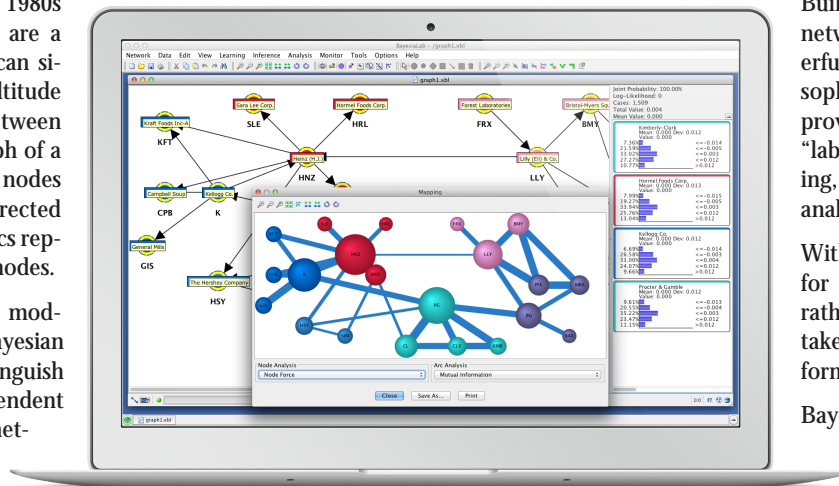# BAYESIALAB
10 PROFESSIONAL EDITION

❶ What is a Bayesian Network?

Invented by Judea Pearl in the 1980s at UCLA, Bayesian networks are a mathematical formalism that can simultaneously represent a multitude of probabilistic relationships between variables in a system. The graph of a Bayesian network contains nodes (representing variables) and directed arcs that link the nodes. The arcs represent the relationships of the nodes.

Whereas traditional statistical models are of the form $y=f(x)$, Bayesian networks do not have to distinguish between independent and dependent variables. Instead, a Bayesian network approximates the entire joint probability distribution of the system under study.

❷ What is BayesiaLab?

Built on the foundation of the Bayesian network formalism, BayesiaLab is a powerful desktop application with a highly sophisticated graphical user interface. It provides scientists with a comprehensive "lab" environment for machine learning, knowledge modeling, diagnosis, analysis, simulation, and optimization.

With BayesiaLab, it has become feasible for applied researchers in many fields, rather than just computer scientists, to take advantage of the Bayesian network formalism.

BayesiaLab builds upon the inherently graphical structure of Bayesian networks and provides highly advanced visualization techniques to explore and explain complex problems. As a result, a broad range of stakeholders, regardless of their quantitative skills, can engage with a Bayesian network model and contribute their expertise.
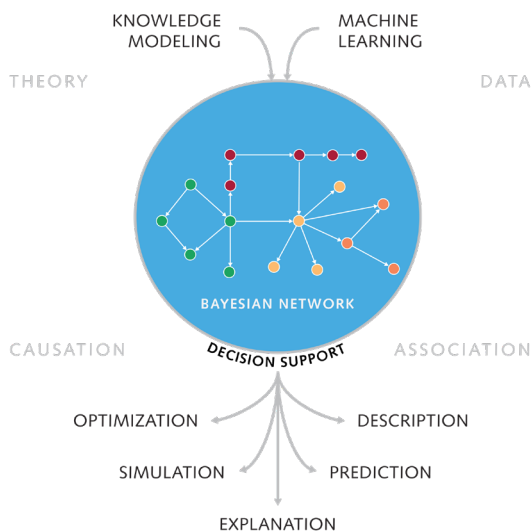
❸ The BayesiaLab Workflow in Practice

Researchers can use BayesiaLab to encode their domain knowledge into a Bayesian network. Alternatively, BayesiaLab can machine-learn a network structure purely from data collected from the problem domain.

Irrespective of the source, a Bayesian network becomes a representation of the underlying, often high-dimensional problem domain.

The researcher can then use BayesiaLab to carry out "omni-directional inference," i.e., reason from cause to effect (simulation) or from effect to cause (diagnosis) within the Bayesian network model.

On this basis, BayesiaLab offers an extensive analytics, simulation, and optimization toolset, providing comprehensive support for policy development and decision-making.

In this context, BayesiaLab is unique in its ability to distinguish between observational and causal inference. Thus, decision-makers can correctly simulate the consequences of actions not yet taken.

Recent Clients and Customers

- Accenture
- Acxiom
- Allstate Insurance
- Amazon
- American Family Insurance
- ArcelorMittal
- A*STAR
- Astellas Pharma
- Australian Defence Science and Technology Group
- Bentley University
- Booz Allen Hamilton
- BP
- Burke, Inc.
- Cap Gemini
- Cancer Care Ontario
- Cargill
- Caterpillar
- CDC
- CFI Group
- Chadwick Martin Bailey

- Charles Schwab
- Cisco
- Cognizant
- Comcast
- Crédit Agricole
- Dassault Aviation
- Dell
- Deloitte Consulting
- Dextro Analytics
- DGA
- Duke University
- EADS
- Électricité de France
- ENI
- Facebook
- FedEx
- Firmenich
- Fractal Analytics
- France Télécom
- Geisinger Health System
- Georgia-Pacific

- Georgetown University
- George Washington University
- GfK
- GlaxoSmithKline
- GroupM
- Hewlett-Packard
- Hilton
- Hyatt
- Indian Institute of Management Bangalore
- Indian Institute of Technology Roorkee
- Intel Corporation
- InterContinental Hotels Group
- Ipsos
- Johns Hopkins Bloomberg School of Public Health
- KPMG
- L'Oréal
- La Poste
- Lowe's

- LRW
- Lilly
- Lockheed Martin
- Mead Johnson
- McGill University
- McKinsey & Company
- Microsoft
- Millward Brown
- Multimedia University
- Mu Sigma
- Nanyang Technological University
- Nanyang Polytechnic
- NASA
- National Analysts
- National Aerospace Laboratories
- National University of Singapore
- Nestlé
- Neiman Marcus

- Nielsen
- Northwestern Mutual
- Northwestern University
- NTT
- NYU School of Medicine
- Oak Ridge National Laboratory
- Procter & Gamble
- PSA Peugeot Citroën
- PwC
- Rainman Consulting
- Renault
- Rhodia
- RTI
- Rutgers University
- Saint-Gobain
- Samsung
- SingTel
- Smuckers
- Société Générale
- Sony
- Starcom Mediavest

- Synovate
- PennState
- Texas State University
- The Pert Group
- TNS
- Total
- Turbomeca
- UCLA
- Unilever
- University of Maryland
- University of North Carolina
- University of Toronto
- University of Virginia
- U.S. Army
- U.S. EPA
- U.S. FAA
- U.S. FDA
- U.S. Veterans Administration
- Vanderbilt University
- Virginia Tech
- World Bank

❹ Key Features of BayesiaLab

Expert Knowledge Modeling. Despite the importance of "Big Data," we often face "no data" problems. Thus, human knowledge, especially causal knowledge, remains irreplaceable in many domains. Bayesian networks mirror the commonly-used representation of causal effects as directed arcs. In BayesiaLab, nodes (representing variables) can be added and positioned with a mouse click; arcs (representing relationships) can be drawn between nodes. The arc direction can encode the causal direction. The quantitative part of these relationships can be defined nonparametrically with BayesiaLab's Node Editor.

The Bayesia Expert Knowledge Elicitation Environment (BEKEE) is available as an extension to BayesiaLab. It provides an intuitive web interface to elicit the explicit and tacit knowledge of domain experts.

Unsupervised Structural Learning. "Unsupervised learning" is typically understood to be a classification or clustering task. As the name implies, Unsupervised Structural Learning expands this concept. BayesiaLab can discover structure, i.e., it can learn probabilistic relationships between a large number of variables—*without* the researcher having to define inputs or outputs. This is the quintessential form of knowledge discovery, as no assumptions whatsoever are required to learn from unknown datasets.

Supervised Learning. Supervised Learning in BayesiaLab has the same purpose as conventional predictive modeling techniques, i.e., developing a model for predicting a dependent variable. Within BayesiaLab, a wide range of algorithms is available for searching for many types of Bayesian networks to characterize the target variable while simultaneously considering the complexity of the resulting network. In particular, the **Markov Blanket** algorithm is helpful when dealing with hundreds or thousands of potential predictors. As such, it can serve as an exceptionally powerful variable selection tool.

Observational Inference. Bayesian networks are "omni-directional inference engines." Given an observation of any of the network's nodes (or a subset of nodes), BayesiaLab instantly computes the conditional probabilities of all other nodes in the network. Both *exact* and *approximate* inference algorithms are implemented in BayesiaLab.

Causal Inference. Bayesian networks have unique properties when it comes to causality. With specific theory-driven assumptions, Bayesian networks can support *causal* inference. Thus, Bayesian network models can cover the entire range of reasoning from *association* to *causation*. For instance, BayesiaLab can perform *causal* inference computations to determine the impact of *intervening* on variables—instead of merely *observing* their states. In BayesiaLab, **Pearl's Do-Operator** and **Likelihood Matching** are available for this purpose. The latter can be used even without having formally defined a *causal* Bayesian Network. **Likelihood Matching** also serves as the basis for all **Direct Effect** computations available in BayesiaLab's analytics functions.

Diagnosis, Prediction, and Simulation. In the Bayesian network framework, *diagnosis*, *prediction*, and *simulation* are identical computations. They all consist of inference conditional upon observing (or setting) evidence. A distinction between them only exists from the perspective of the researcher, who would, for instance, understand the symptom of a disease is an *effect* and the disease itself being the *cause*. Hence, carrying out inference based on observed symptoms is interpreted as a diagnosis.

BayesiaLab offers numerous functions related to inference. For ad-hoc analysis, inference can be performed by setting evidence on any node, and results are returned instantly for all the other nodes. **Batch Inference** is available for scoring a large number of records in a database. The optional **Bayesia Engine API** provides programmatic access to these inference functions; the optional **Export Module** translates predictive models into code for SAS, R, VBA, Python, PHP, and JavaScript.

Reasoning Under Uncertainty. Even with uncertain or conflicting evidence, BayesiaLab performs mathematically correct inference. The inherent ability of Bayesian networks to facilitate computations under uncertainty makes them ideally suited for real-world applications, which typically contain some degree of uncertainty. This approach entirely avoids presenting potentially misleading point estimates.

Optimization. Performing inference across all possible states of all nodes in a network allows searching for optimum values concerning a target variable. BayesiaLab's **Target Optimization** algorithms were developed for this purpose. Using these functions in combination with **Direct Effects** is particularly powerful when searching for the optimum combination of values of variables that have nonlinear causal relationships with the target, plus a multitude of interaction effects between them. A typical application is marketing mix optimization. BayesiaLab can search, within a given resource constraint, e.g., a marketing budget, for those scenarios that optimize the target criterion.

www.BayesiaLab.com

For North America

Bayesia USA, LLC
4235 Hillsboro Pike
Suite 300-688
Nashville, TN 37215 USA

+1 888-386-8383
info@bayesia.us

For Asia Pacific

Bayesia Singapore Pte. Ltd.
1 Fusionopolis Place
#03-20 Galaxis
Singapore 138522

+65 6809 1210
info@bayesia.com.sg

Head Office

Bayesia S.A.S.
Parc Cérès, Bâtiment N
21, rue Ferdinand Buisson
53810 Changé, France

+33 (0)2 43 49 75 69
info@bayesia.com