



BayoLinkS 新バージョンのご紹介

NTT DATA
Trusted Global Innovator

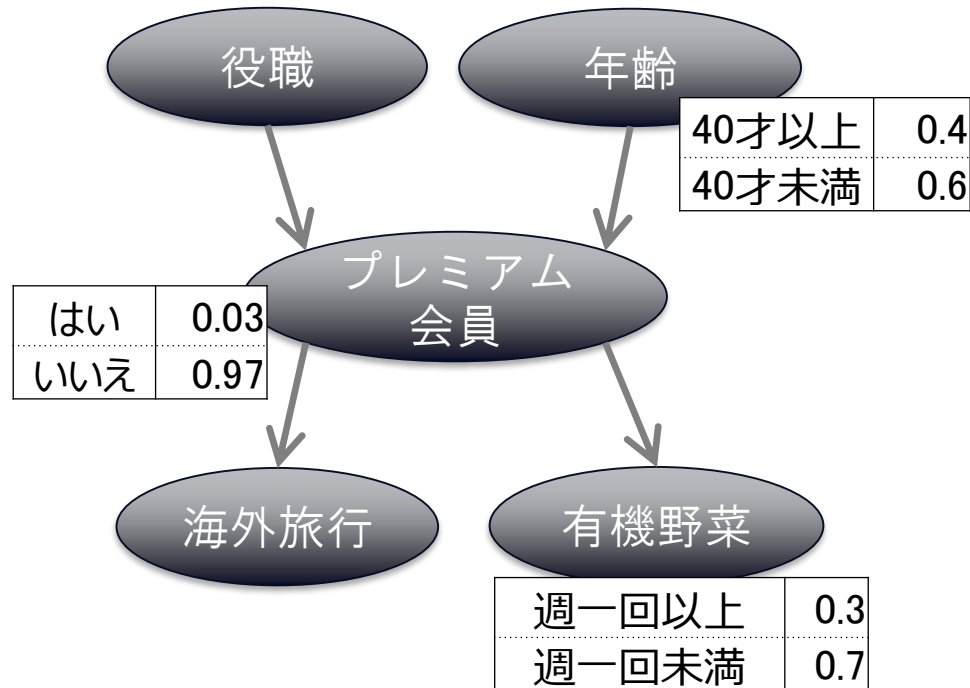


株式会社NTTデータ数理システム

ベイジアンネットワークとは?

- ✓ 確率推論モデルの一つ
- ✓ 事象間の確率的な関係をネットワークで表現する
- ✓ 観測された情報をモデルに入力し、確率を計算する
- ✓ 未知の状況の予測や分析を行う

□ ベイジアンネットの具体例：プレミアムカード会員モデル



『有機野菜』の条件付き確率

有機野菜 (購入頻度)	プレミアム = はい	プレミアム = いいえ
週一回以上	0.65	0.3
週一回未満	0.35	0.7

ベイジアンネットの特徴

- ✓ 自由度が高く柔軟性に優れたモデル
 - モデルに観測値を入力して推論すると観測のない全てのノードで確率値を計算する。このことから**複数の目的変数について同時に推論**を実施することができる。
 - 推論の方向は 矢印の順方向だけではなく逆向きでも可能。原因から結果を予測するだけではなく、原因を推定するような分析にも対応する。
 - 1つのモデルを使って、**様々な方向で分析や予測**ができる
 - **欠損**があっても推論ができる
- ✓ 因果関係の仮説をモデルの構造で表現できる
 - 変数間の因果関係を矢印の向きで表現できる
 - **事前知識**を取り入れることで、より自然なモデルを表現できる
 - 推論の構造を**直観的**に理解できる
- ✓ 変数間の確率的な依存関係のみでモデルを構築できる
 - 論文や経験から得られた変数同士の条件付確率があれば、その変数についてデータがなくてもモデルを構築できる。**スパースな学習データ**にも対応。

BayoLinksS の分析について

- ✓ ベイジアンネットワークの構築・分析支援システム
- ✓ モデルの構築から検証、分析などの各工程を強力にサポート

学習データからモデルを構築・分析するまでの工程

データの前処理

- 数値データの離散化
- カテゴリデータのグループ化

機械学習

- 構造学習
- パラメータ学習

検証

- 予測精度などの指標値の算出

予測 / 分析

- 確率推論
- 感度分析

The screenshot shows the 'Bayesian Learning Wizard' interface. On the left, there is a table with columns for 'Node Name', 'Observed', and four 'View' (観) options. The table lists various nodes like 'アジア経済成長' and '経済危機'. On the right, a complex Bayesian network diagram is displayed with nodes and directed edges. Below the diagram, there are buttons for 'All OK', 'All Required', and 'All Stop'.

The screenshot shows the 'Model Verification' window. It displays a table of performance metrics for different models. Below the table, there is a section for 'Probability of Increase in Error Rate' with a dropdown menu and a numerical value. At the bottom, there is a table with columns for 'Rank', 'HE1_成績優良者', 'HE1_年齢', 'ME_進行距離', 'HE2_シニア講習', and 'HE1_居住'.

目的変数	データ数	正解データ数	正解率	平均対数
経済危機?	5000	4049	0.8098	0.6895
経済?	5000	4194	0.8388	0.6265
災害発生?	5000	4210	0.842	0.6028

状態	適合率	再現率	F-Measure
しない	0.9855	0.9845	0.984
はい	0.7865	0.7294	0.7299

順位	HE1_成績優良者	HE1_年齢	ME_進行距離	HE2_シニア講習	HE1_居住	値
1		Adolescent	Domino			5814
2			Domino		Secure	2.43421
3	TRUE		Domino			2.43292
4			Domino		Suburb	2.41374
5			Domino		Rural	2.38365
6			Domino		City	2.34473

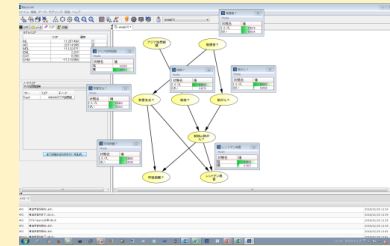
BayoLinkS

BayoLink

ベイジアンネットワークのモデル構築から確率推論まで

- ・ 自動構築・確率推論・モデル検証・モデル編集

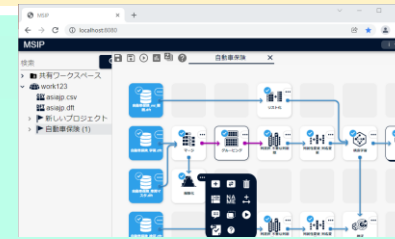
(開発元：(独)産業技術総合研究所)



MSIP

データの前処理から、分析の実行まで

- 充実のデータ前処理機能・可視化機能
- 分析工程を俯瞰できるプロジェクトボード上での分析フロー作成
- 統計・機械学習・テキストマイニングをはじめとするNTTデータ数理システム他製品とのシームレス関係



2002年
BAYONET
販売を開始

2017年
BayoLink に
名称を変更

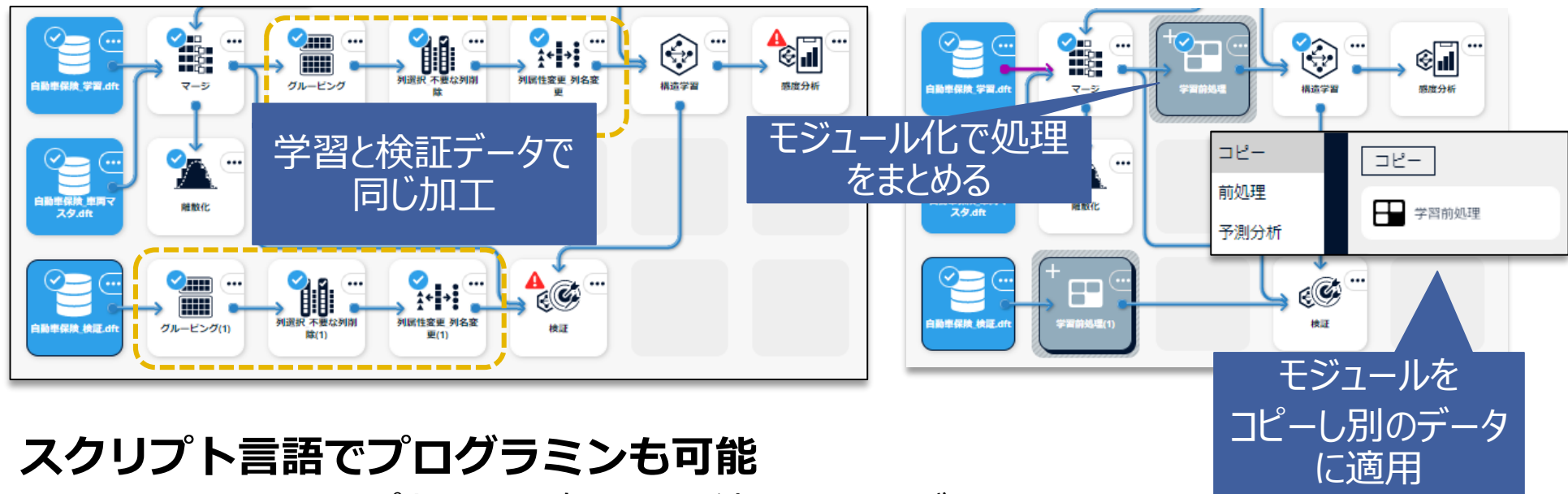
2018年
VAPとセット
BayoLinkS
販売開始

2022年
ver.9.0
リリース

MSIP の強味 ①

ノーコードの分析プラットフォーム

- ・データの取り込みから前処理、モデル構築、検証、結果の可視化までを簡単なGUI操作で実現します。
- ・設定済みのアイコンをコピーして、別の分析で再利用できます。
- ・さらにフローのモジュール化機能で 利用頻度の高い処理フローをまとめることができます。



スクリプト言語でプログラミングも可能

- ・ Pythonスクリプト言語で処理を記述することができます。

MSIP の強味 ②

豊富なデータの前処理メニュー

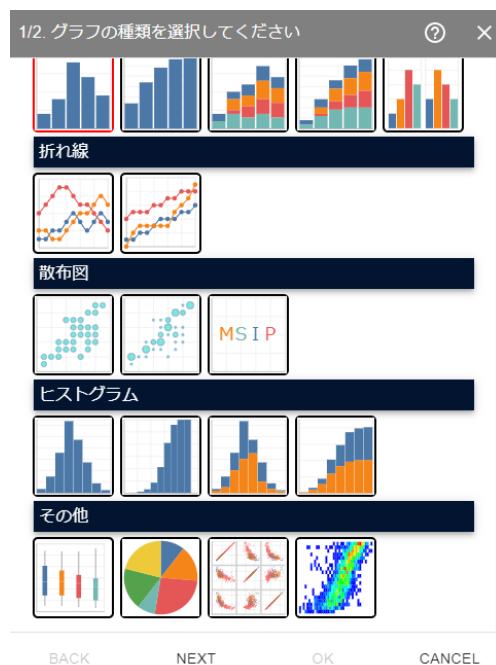
- ・ マージ、クリーニング・サンプリングなどの基本的な加工機能の他、**集計、正規化、クロス集計**などの統計処理にも対応
- ・ ベイジアンネット分析には欠かせない、**数値データの離散化**を搭載



MSIP の強味 ③

データ可視化

- 基本的なヒストグラムや散布図の他、箱ひげ図など多種多彩なグラフを描画できます
- 可視化画面でデータのサマ리를ワンクリックで表示します



asiajp.dft-data 列数: 8 行数: 5,000

No.	アジア訪問経験 CATEGORY	喫煙者? CATEGORY	結核? CATEGORY	肺がん? CATEGORY
1	無	はい	いいえ	いいえ
2	無	はい	いいえ	いいえ
3	無	はい	いいえ	いいえ
4	無	はい	いいえ	いいえ
5	無	いいえ	いいえ	いいえ
6	無	いいえ	いいえ	いいえ
7	無	はい		
8	無	はい		
9	無	いいえ		

データサマリの切り替えボタン

[サマリ表示] asiajp.dft-data 列数: 8 行数: 5,000

列名	特殊値	統計量	グラフ
アジア訪問経験 <CATEGORY>	NULL	: 0	
	NA	: 0	カテゴリ総数 : 2
	ERROR	: 0	無 : 4957
	+Infinity	: 0	有 : 43
	-Infinity	: 0	
喫煙者? <CATEGORY>	NULL	: 0	
	NA	: 0	カテゴリ総数 : 2
	はい	: 2512	
	いいえ	: 2488	
	+Infinity	: 0	

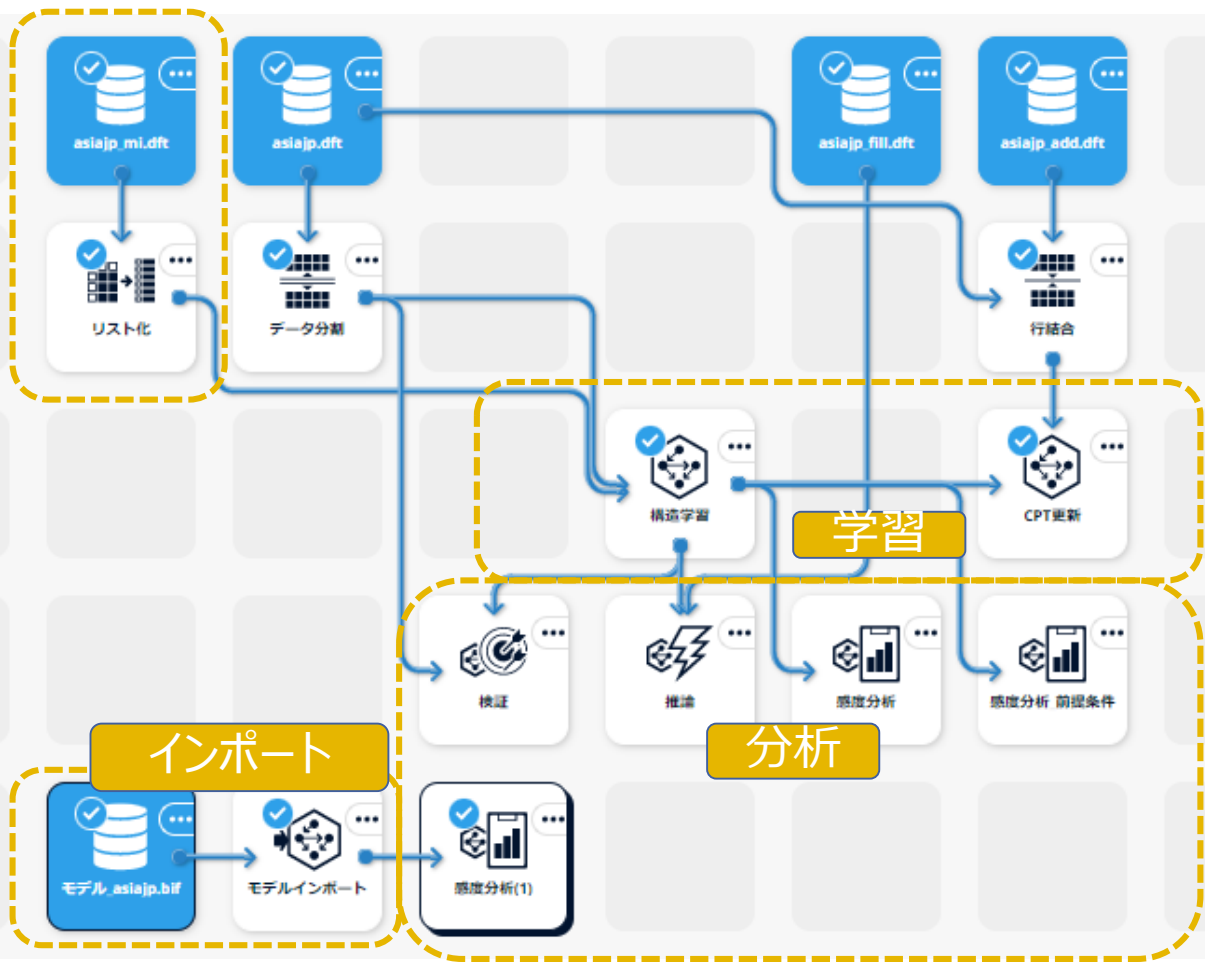
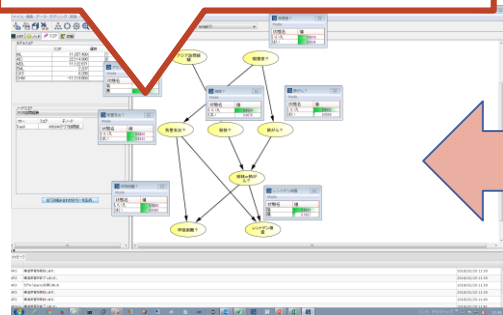
各カテゴリの件数を表示

ベイジアンネットワークの学習から分析までをMSIPで実現します

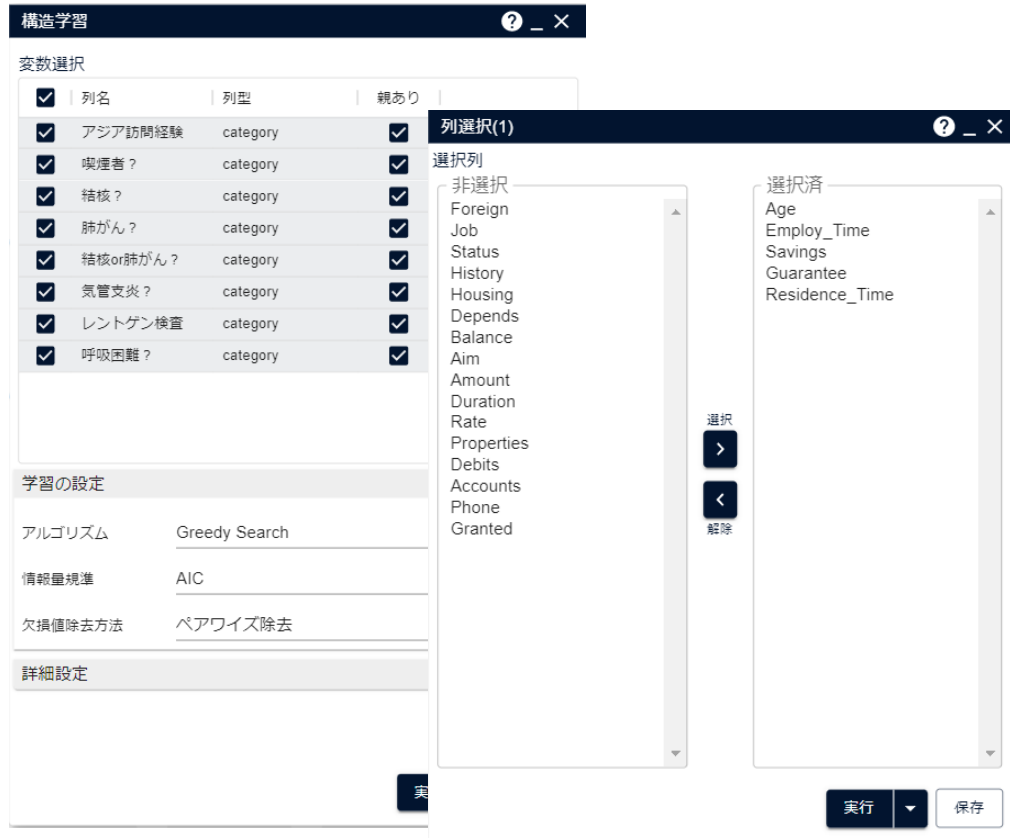
BayoLinks
分析メニュー



従来の単体機能にも
関係可能



(1) 各種アイコンの設定画面のリニューアル



前処理アイコンやBayoLinkS分析アイコンの設定画面がリニューアルをリニューアル。

一部のアイコンにおいて操作性の改善も行いました。

(2) データベース連携機能の強化

ODBC 対応データベース製品からのデータインポートに対応しました。

・対応データベース

Snowflake/MariaDB/Oracle/PosgreSQL/Access など

(3) 一部の前処理アイコンについて、大規模データへの対応、及び高速化を実現しました。

・大規模データ対応

ソート/クリーニング/グルーピング/離散化/クロス集計 など

・高速化対応

グルーピング、クロス集計、マージ、列選択

(4) 操作マニュアルを完全リニューアルしました。

BayoLinkS 適用事例

□ マーケティング分野

- アンケート分析を活用し ヒトの評価と行動をモデル化する
- 商品の関心が高いヒトの属性やライフスタイル、ターゲット層の絞り込みに活用
- 従来のエクセルによるクロス集計分析より短時間での分析が可能
また膨大な質問項目の組合せの中から知見を得る

□ 製造分野

- 製造機器のセンサデータや業務日報を紐付け、さらに現場のノウハウなどを取り込みモデル化する
- 異常の発生から、どの部分に原因があるかを推論によって探る
- 製造ラインで障害が発生する根本原因を究明

□ 医療/ヘルスケア分野

- 症例データや生活習慣のデータを分析することで、疾病のリスクを推論
- 診察結果や周辺データなどを階層構造でモデル化し、疾病の因果関係をわかりやすく表現する
- 確率推論により、疾病のリスクを高める原因を探ることができ、リスクを回避するための具体的な行動を示すことが可能
- 医療現場における治療効果を確率推論



NTT DATA

Trusted Global Innovator

お問い合わせ：

NTTデータ数理システム 営業部

Mail : bayonet-info@ml.msi.co.jp

Tel: 03-3358-6681