

NTT DATA

NTT DATA Mathematical Systems Inc.

BayoLinkS 新バージョンのご紹介

数理システム 新バージョン発表会 2021

2021年12月15日
(株)NTTデータ数理システム

Trusted Global Innovator

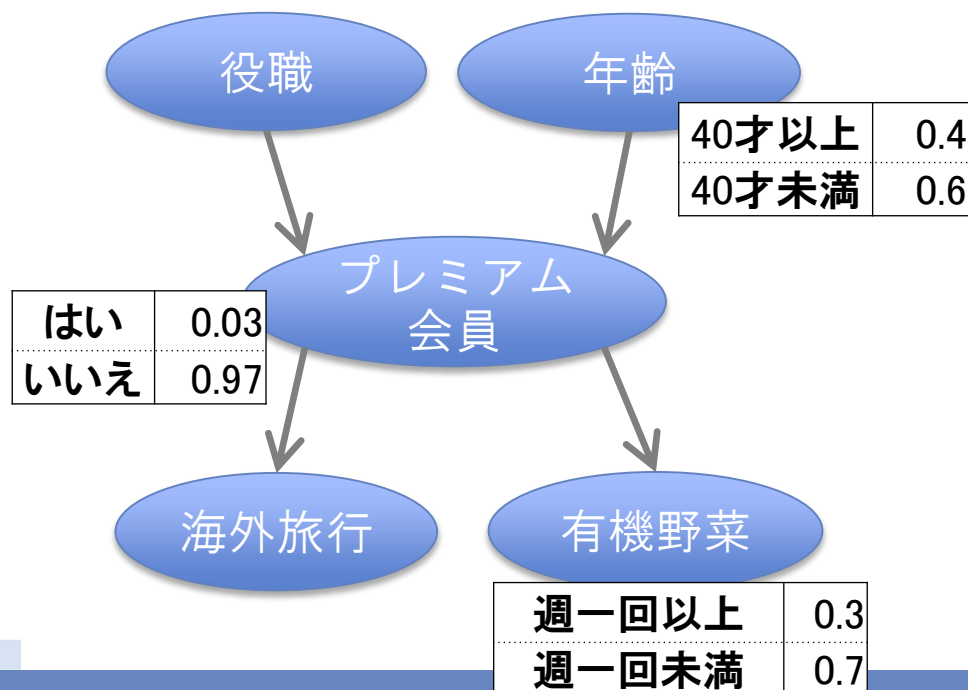
NTT DATA Group

NTT DATA

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。

BayoLinkS 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

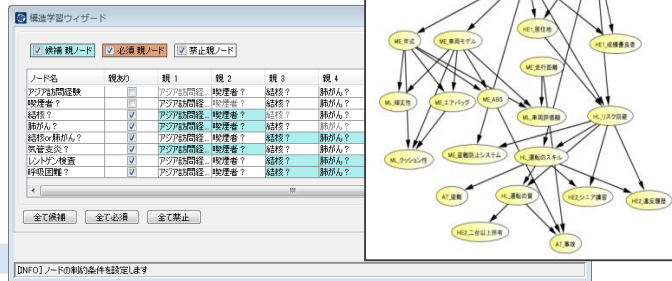
- 構造学習
- パラメータ学習

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



BayoLinkS 製品概要 (既存リリース版)

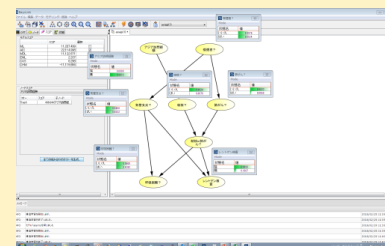
BayoLinkS

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

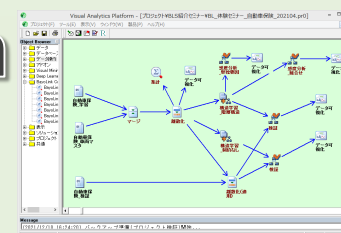
(開発元：(独) 産業技術総合研究所)



Visual Analytics Platform

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
- テキストマイニング・データ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



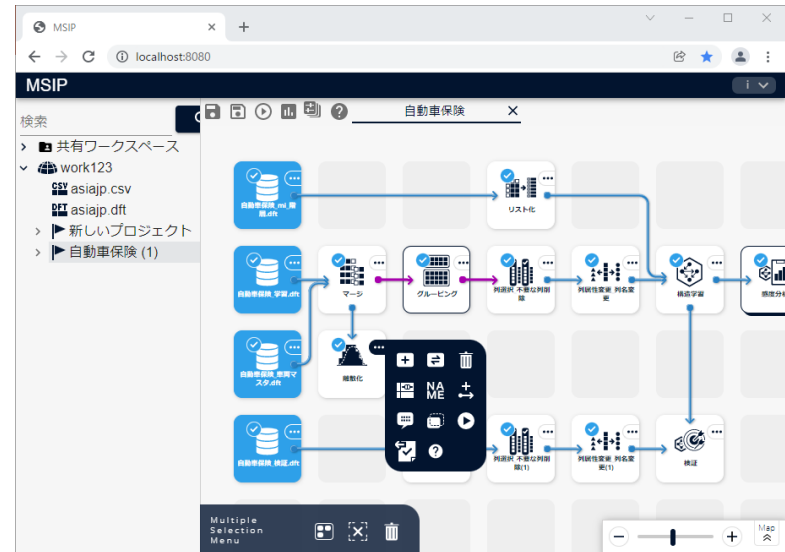
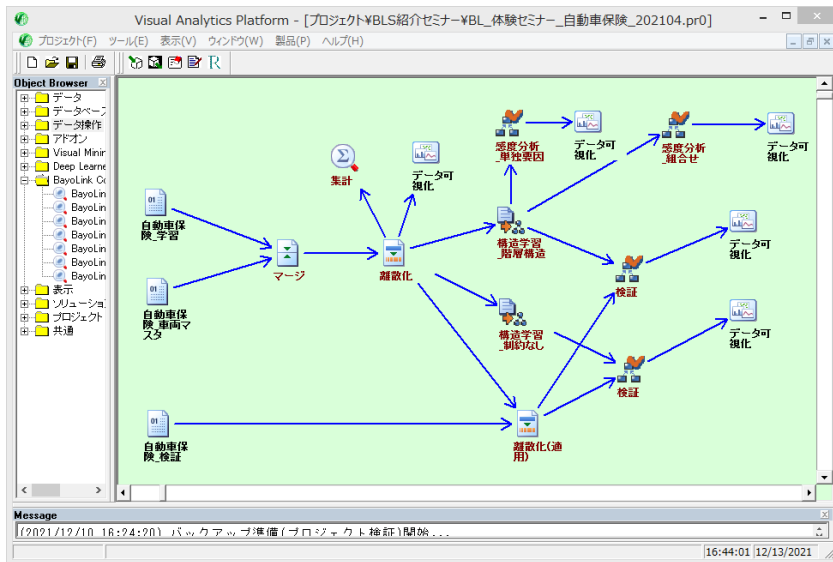
2002年
BAYONET
販売を開始

2017年
BayoLink に
名称を変更

2018年
VAPとセット
BayoLinkS
販売開始

2021年
ver.8.0
リリース

分析基盤を Visual Analytics Platform から MSIP に移行します！



- ・デスクトップアプリケーション
- ・VMS と共通基盤



- ・ブラウザベースアプリケーション
- ・Alkano と共通基盤



データ活用の確かなパートナー

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。

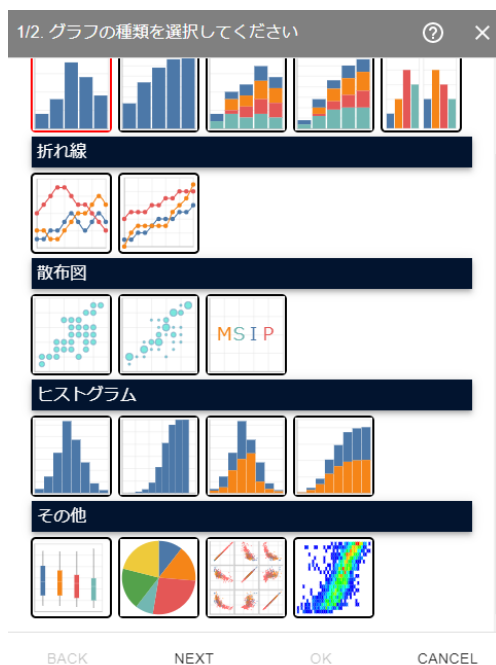
豊富なデータの前処理メニュー

- ・ 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



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	いいえ	いいえ
8	無	はい	いいえ	いいえ
9	無	いいえ	いいえ	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 ERROR : 0 +Infinity : 0 -Infinity : 0	カテゴリ総数 : 2 無 : 4957 有 : 43	
	NULL : 0 NA : 0 ERROR : 0 +Infinity : 0 -Infinity : 0	カテゴリ総数 : 2 はい : 2512 いいえ : 2488	

各カテゴリの件数を表示

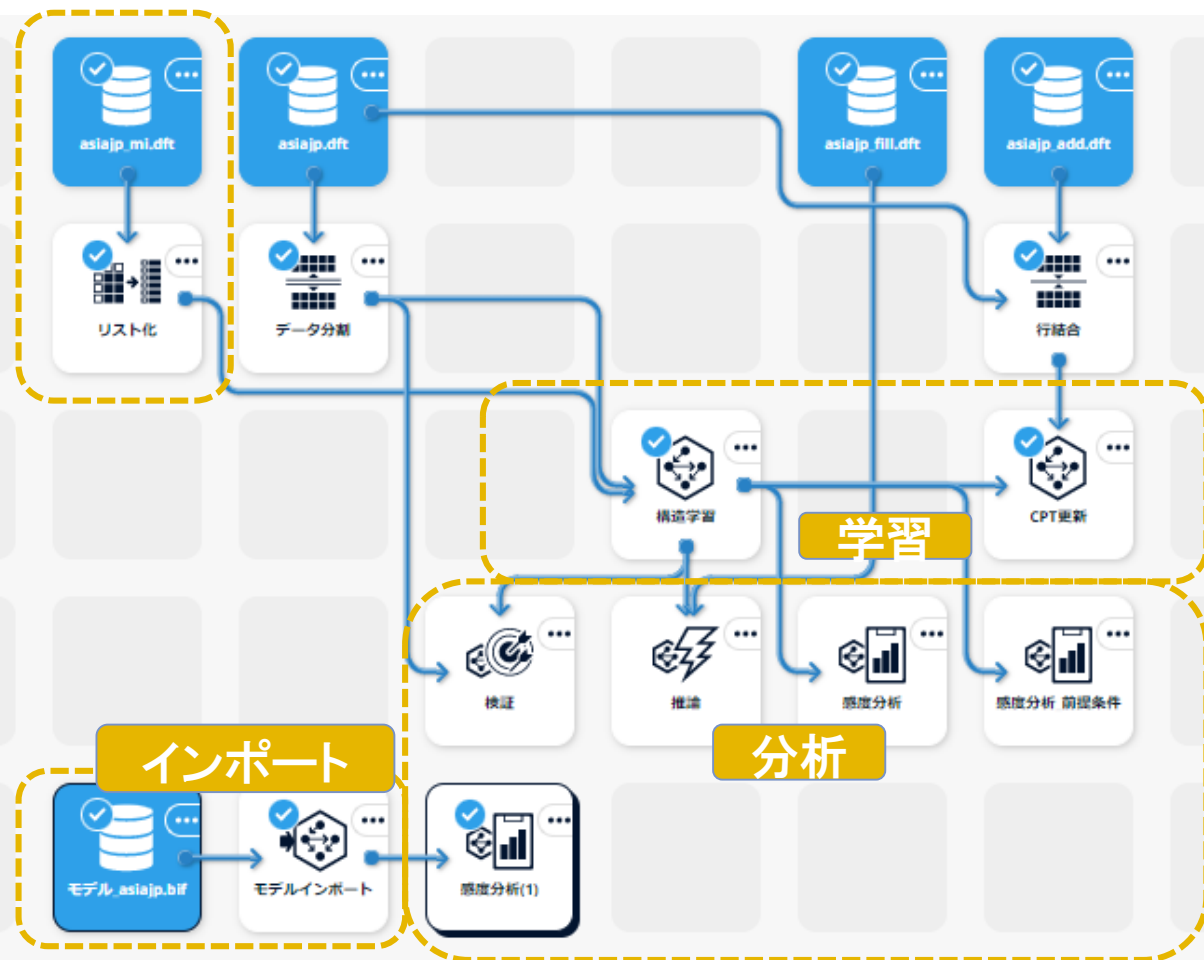
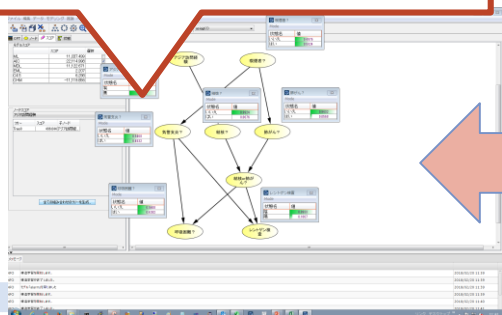
MSIP の BayoLinkS プラグイン

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



□ 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

- 2022 年 春
分析基盤を MSIP に移行
- 2023 年 春 (予定)
ネットワーク編集画面に対応



NTT DATA

NTT DATA Mathematical Systems Inc.

Trusted Global Innovator

NTT DATA Group

NTT DATA