

# Alkano バージョン1.1.1 のご紹介

**NTT DATA**  
Trusted Global Innovator



 **Alkano**

データ活用の確かなパートナー

株式会社NTTデータ数理システム

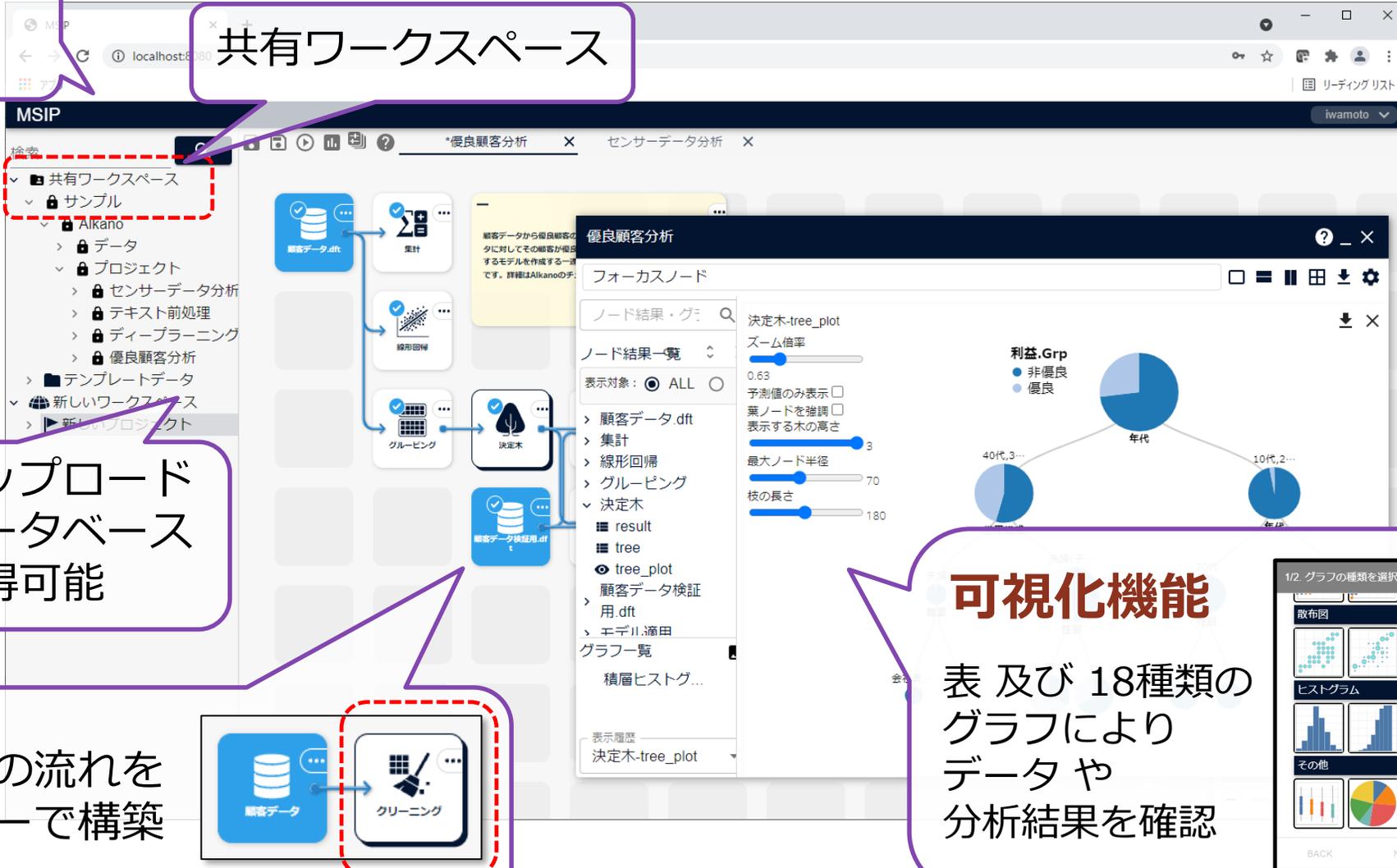
Webブラウザ  
から利用

共有ワークスペース

データをアップロード  
もしくはデータベース  
から取得可能

処理の流れを  
フローで構築

機能の単位：アイコン



可視化機能

表及び18種類の  
グラフにより  
データや  
分析結果を確認

## 多彩な前処理機能・分析機能を備えます

### 前処理アイコン群



### 分析アイコン群



予測分析  
クラスタリング  
関連性分析  
多変量解析  
時系列処理  
ディープラーニング  
テキスト前処理

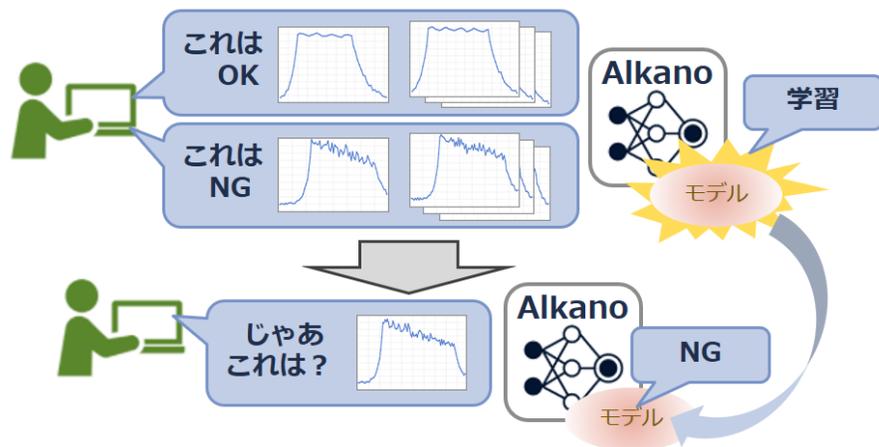
- クリーニング、サンプリング、データの正規化といったデータ分析の場面で必要になるデータ加工 機能を多数搭載しています。
- データの傾向把握 から 予測モデル構築 まで、分析に必要な機能を網羅しています。

## 予測分析

予測分析

- 決定木
- ランダムフォレスト
- ニューラルネットワーク
- サポートベクターマシン
- k-NN法
- 単純ベイズ
- 線形回帰
- ロジスティック回帰
- Elastic Net
- 勾配ブースティング決定木
- 説明変数重要度
- 予測精度検証
- 特徴量選択 **NEW**
- 代表点抽出 **NEW**

- データとそれに対する答えの実例を大量に学習させ、その結果に基づいて未知なデータに対する予測を行う **教師あり学習** の機能です。

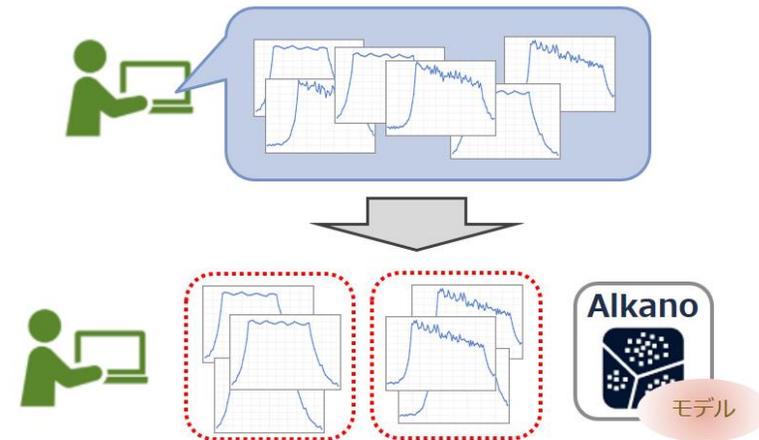


## クラスタリング

クラスタリング

- k-means法
- One Class SVM
- Isolation Forest
- 異常検知フォレスト
- ネットワーククラスタリング
- 階層型クラスタリング
- 二項ソフトクラスタリング

- データのみを与え、自動的にデータのグループ分け（クラスタリング）を行うような **教師なし学習** の機能です。
- **外れ値検出** に有効な手法も収められています。

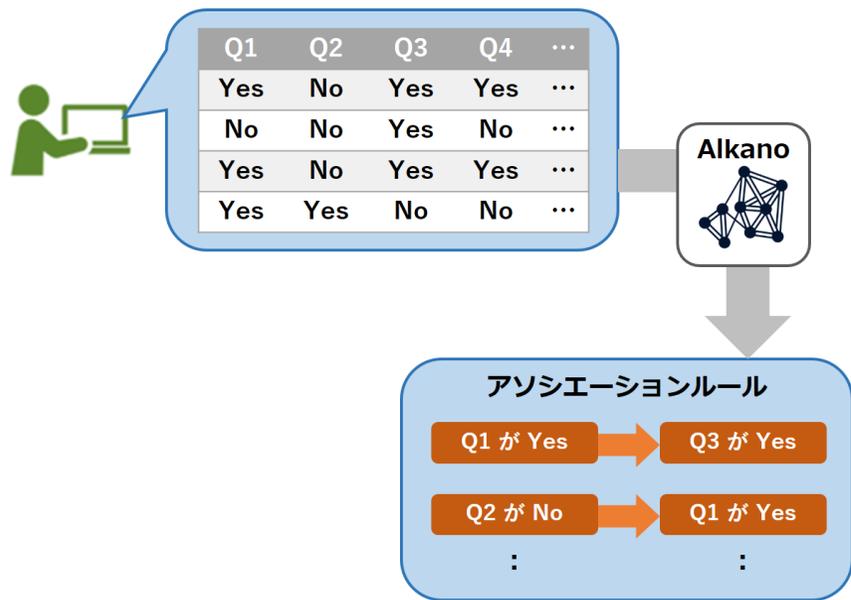


## 関連性分析

関連性分析

相関分析    アソシエーション分析

- **変数間の関連性**を分析するための機能です。データの内容把握のために大変有用です。



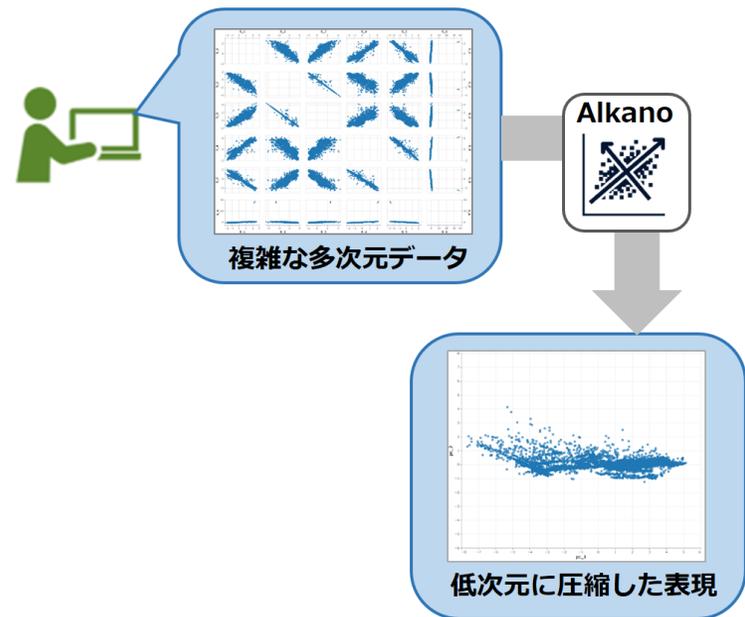
アソシエーション分析 利用イメージ

## 多変量解析

多変量解析

主成分分析    対応分析

- 主成分分析、対応分析 といった **多変量解析**の手法を利用するための機能です。



主成分分析 利用イメージ

## 時系列処理 NEW

時系列処理

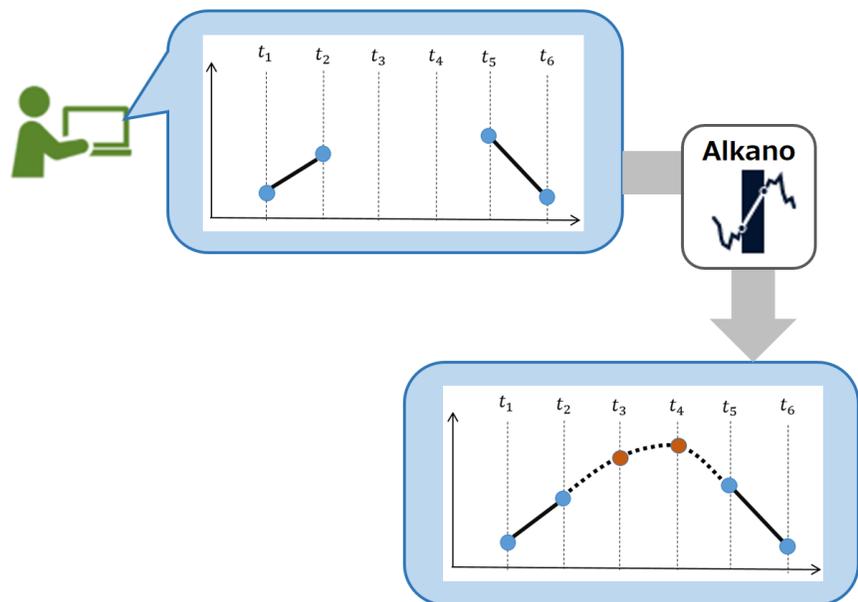
 時系列切り出し

 時系列欠損補間

 時系列特徴量作成

 時系列特徴量作成(SFA変換)

- **時系列データ** を分析に適した形に加工するための **前処理機能** と、時系列データを分類や予測、クラスタリングで利用するための **特徴量作成機能** を搭載しています。



時系列欠損補間 利用イメージ

## ディープラーニング

ディープラーニング

 ディープラーニング  
テーブル

 ディープラーニング  
時系列

 ディープラーニング  
テキスト

 ディープラーニング  
テキスト属性

- 深層学習（**ディープラーニング**）によるモデル構築を行うための機能です。**時系列データ**、**テキストデータ** といった **可変長の系列データを自然に扱える** ようなアルゴリズム を搭載しています。
- **GPU計算** に対応しています。

## テキスト前処理

テキスト前処理

 文章ベクトル化

 Word Embedding

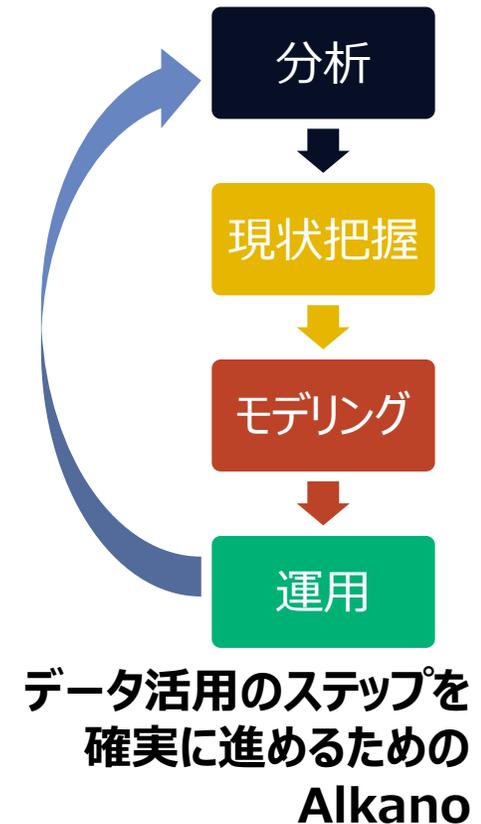
 形態素解析

 区切り文字を指定して  
分割

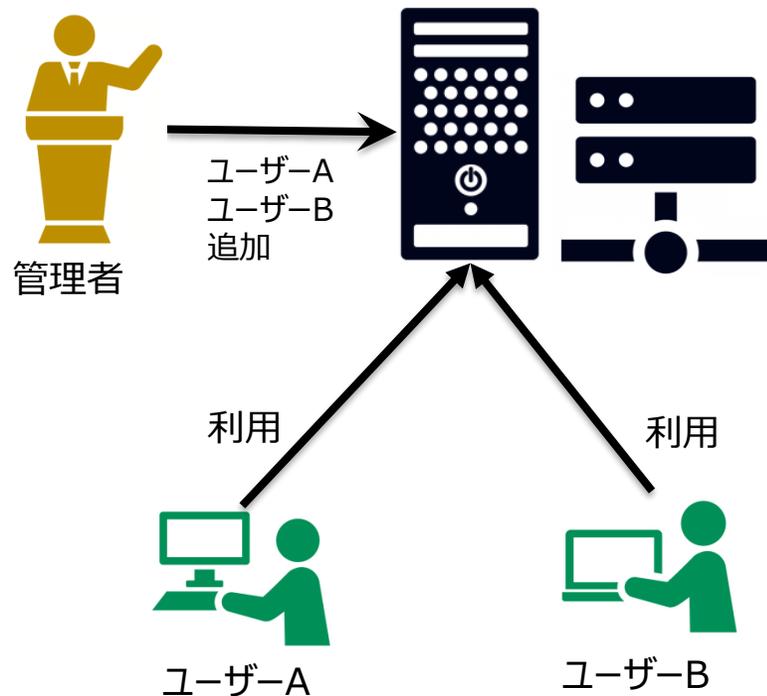
- **テキストデータ** に対する前処理を行う機能です。
- この機能で **「文章」のデータを「単語」の情報に変換する** ことで、Alkano の各種手法がテキストデータに対しても適用できるようにす。

## チームで分析→運用をサイクルで実現

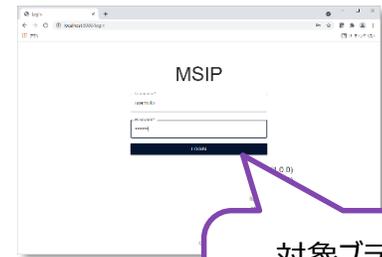
- 探索的な分析試行環境として
  - チームでの利用により適した形に進化しました
    - クライアント/サーバ形式 による 機動性
    - データ・プロジェクト の 共有 と 権限管理
- 組み込み～運用環境として
  - 試行・検証の結果をそのまま同一環境でデプロイできます
    - 作成したフローを RESTful API から呼出し可能
- 変化を見据えた**拡張性**も備えます
  - Pythonアイコン, Rアイコン による 処理の拡張
  - プラグイン機構 による 独自アイコン拡張機能
  - ※今後、別途 開発者用ライセンス をご用意させていただく予定です



## 利用環境の自由度が高くなります



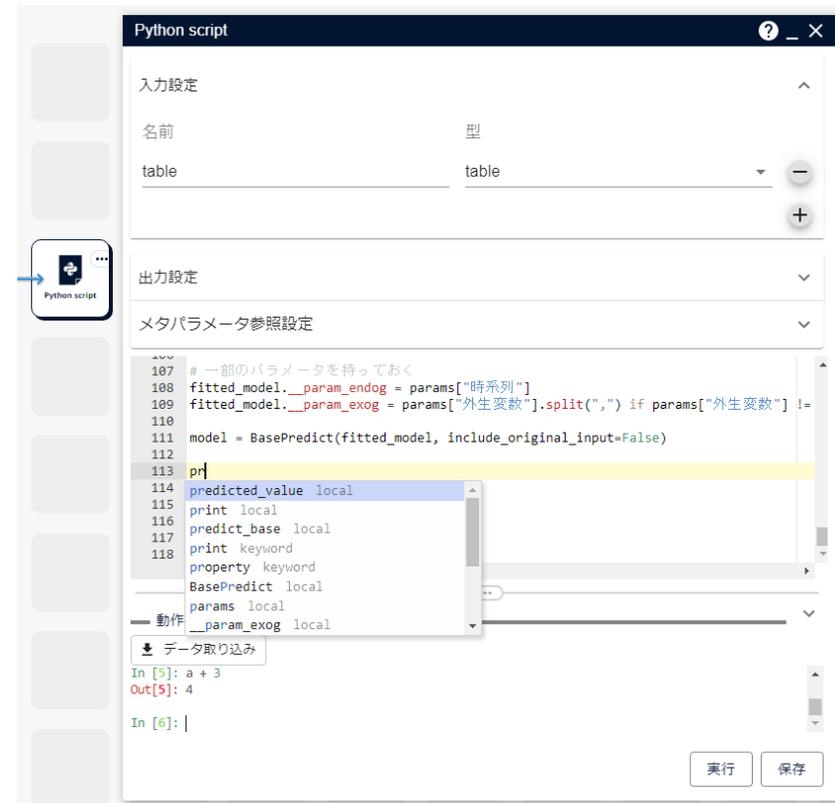
- ライセンス形態 により、MSIP へ追加できる ログインユーザー数 が制御されます。
  - named での管理を行います。
- お客様環境に サーバー (Windows) を設置のうえご利用いただけます。
  - 1台のPC で、自分自身をサーバーとして閉じた形での利用も可能です。
  - クラウド上での利用をお考えの方は、当社よりプライベートクラウドのご紹介も可能です。お問合せ下さい。



対象ブラウザ：  
Google Chrome  
Microsoft Edge

## Python 資産を活用できます Python 既学習者にも利用の幅が広がります

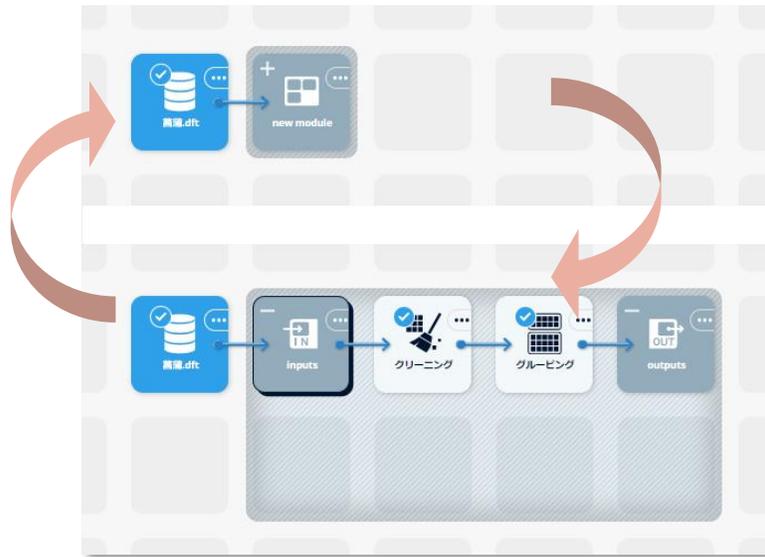
- データ分析の分野で急速に利用が広まってきたプログラミング言語 **Python** をアイコン内でそのまま記述し、実行できます。
- 次のような場面で威力を発揮します。
  - お手持ちのデータに強く依存した、Alkano 標準機能 を超える範囲の **カスタム な 前処理 の実現**
  - 巷の **Python資産 を活用しての、さらなる 機能拡張**



Pythonスクリプトアイコン と コードエディタ

## フローやモデルをシンプルに扱えます

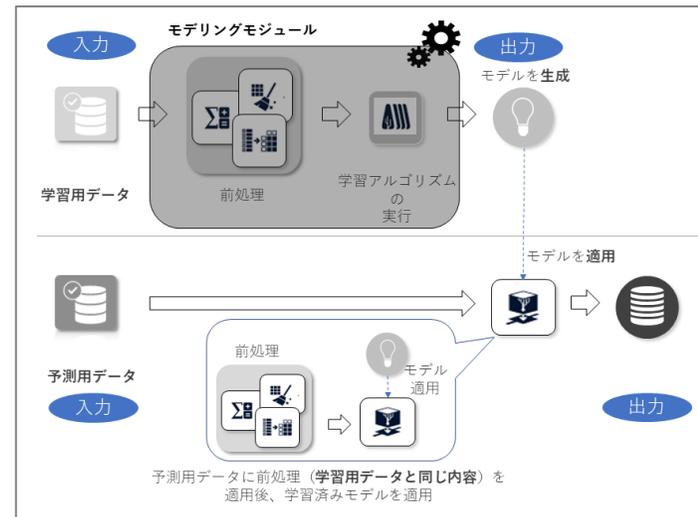
- 複数の処理アイコンをまとめて **モジュール** という 1つのアイコンとして扱うことができます。



モジュールの開閉操作

- さらに **モデリングモジュール** という機構を用いると、前処理・予測モデル構築 といった 一連の処理 をまとめてモデルとして扱うことができます。

- この枠組みを用いて パラメータチューニングの自動化 (最適化実行) が可能です。



モデリングモジュール概略

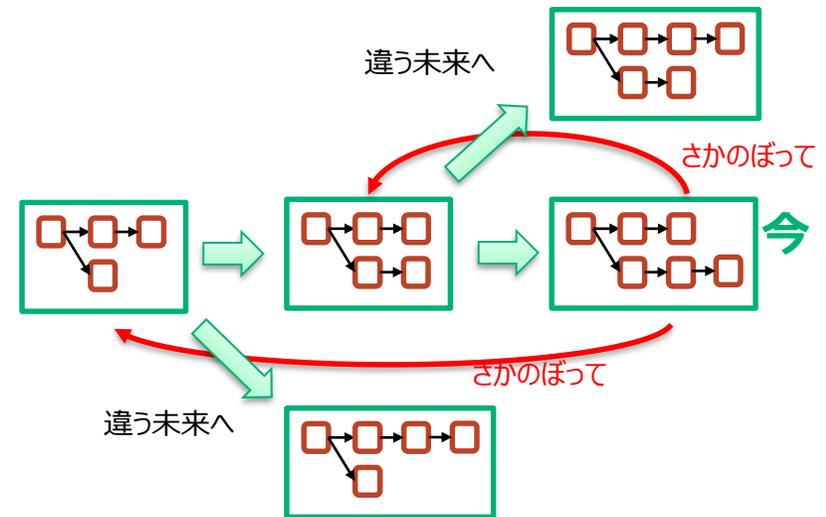
## マルチな実行を実現します

- 処理の実行は完全にバックグラウンドで行われます。
  - ユーザーの手を止めません
- 並列実行で処理を効率化します。
  - 同一フロー内の分岐も
  - 同一フローを分析パラメーターを変えて実行した場合も



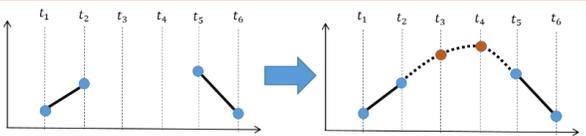
ジョブ一覧					
シナリオ名	リクエスト時刻	開始時刻	終了時刻	ステータス	進捗
<input type="checkbox"/> sw1	2021/6/13 19:10:16	2021/6/13 19:10:16		実行中	<div style="width: 50%;"></div>
<input type="checkbox"/> sw1	2021/6/13 19:10:09	2021/6/13 19:10:09		実行中	<div style="width: 50%;"></div>

- 結果は全て後から振り返ることが可能です。(タイムシフト機能)
  - さらに、振り返った過去から異なる分析を実行可能です。



## 時系列処理

- 時系列データをより容易に取り扱えるようになるための機能群を追加いたしました。

 <p>時系列切り出し</p>	<p>ウィンドウを設けて時系列データを切り出します。 長大な時系列データを、Alkano で扱いやすい形にします。</p>
 <p>時系列欠損補間</p>	<p>欠損値や異常値を検出して、前後の値を考慮し補間を行います。</p> 
 <p>時系列特徴量作成</p>  <p>時系列特徴量作成(SFA 変換)</p>	<p>時系列データから様々な方法で特徴量を抽出し、元々 1件1件 の長さが異なるような 時系列データを統一した形に整形します。</p>

## 予測分析

- 特徴量選択 と 代表点抽出の 2機能を新規に搭載いたしました。

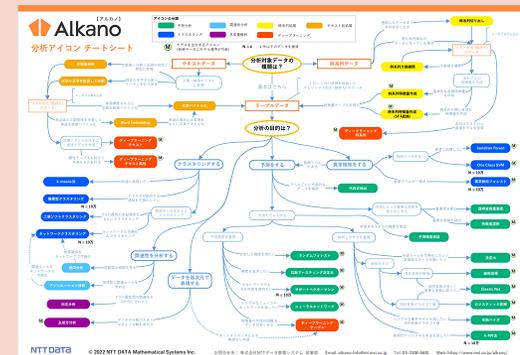
 <p>特徴量選択</p>	<p>相関係数の高い変数同士があったときに一方を除外する、また目的変数を指定して説明変数重要度の高い変数を採用する、といった手法で重要な特徴量を選択します。</p>
 <p>代表点抽出</p>	<p>データに付随するラベル毎に、分布を代表する少数個の代表点を抽出することができます。</p>

## ■ テクニカルサンプルプロジェクト の公開を開始いたしました。

- Alkanoの基本的な使い方から典型的な分析の利用例、さらに一歩進んだ分析方法の手順を表現したテクニカルサンプルプロジェクトの公開を開始しました。
- **新バージョン1.1.1 でご利用いただけます。**  
**ログイン画面 や メニュー から公開サイトにアクセスできます。**



用途から分析機能を  
逆引きできる  
Alkanoチートシートも  
公開しています



## ■ ユーザビリティに関する修正を多数実施いたしました。

# 2023年3月 新規リリースのご紹介

1. 予測モデルの精度検証のための可視化機能を追加

2. モデル解釈のための新分析アイコンを追加

3. 対応データベースの拡充

**2023年3月リリース予定です**

**1.** 分析機能「予測精度検証」の出力に、精度検証のための各種のグラフを追加します。

➤ 回帰分析・分類分析の結果評価が、より直感的に可能になります。

**2.** モデルを解釈するうえで有用な「XAI手法」を備えた新アイコンを追加します。

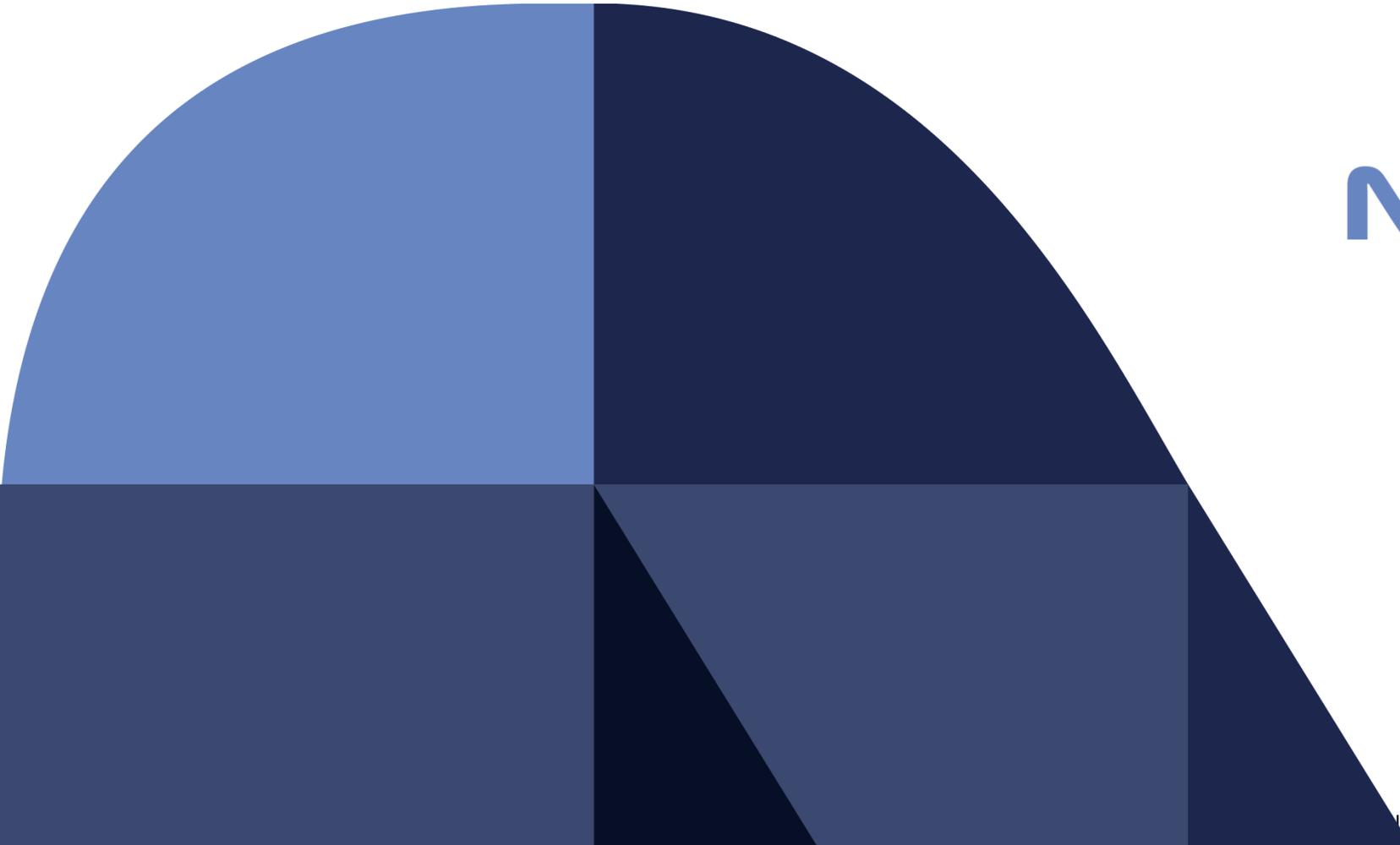
➤ 「Permutation Importance」「ALE Plot」といったアウトプットから、モデルの解釈が可能になります。

## 3. 対応データベースを拡充します。

- ODBC接続 に対応します。これにより、ODBC で接続可能な DB は、統一的に利用できるようになります。

## その他の取り組み

- 2022年9月の公開開始時から、ご好評をいただいている「テクニカルサンプルプロジェクト」をさらに拡充してまいります。



**NTT DATA**  
Trusted Global Innovator