

# AD Analytics バージョン 1.2 新機能紹介

株式会社 NTT データ数理システム

異常検知分析ツール AD Analytics (エーディーアナリティクス) をご利用いただき、誠にありがとうございます。本文書では、新機能、追加・修正機能についてご紹介いたします。

## ◆ 新機能

### ■ 「時系列切り出し」アイコンの追加



時系列データを扱いやすい形式に加工するための前処理機能です。長期の時系列データを基に、複数の系列の部分時系列データをずらしながら切り出します。

### ■ 「時系列特徴量作成 (SFA 変換)」アイコンの追加



時系列データの特徴量を作成するための機能です。部分時系列に含まれる波形を抽象化して抽出し、抽出した波形の出現頻度を特徴量として算出します。

### ■ 「時系列特徴量作成」アイコンの強化



単一変数の統計量 (平均/標準偏差/分位点) と、複数変数間の相互の関係性を特徴量として抽出するための相関・相互相関係数の算出機能を追加しました。

### ■ 「モデル構築 (教師あり)」アイコンの強化

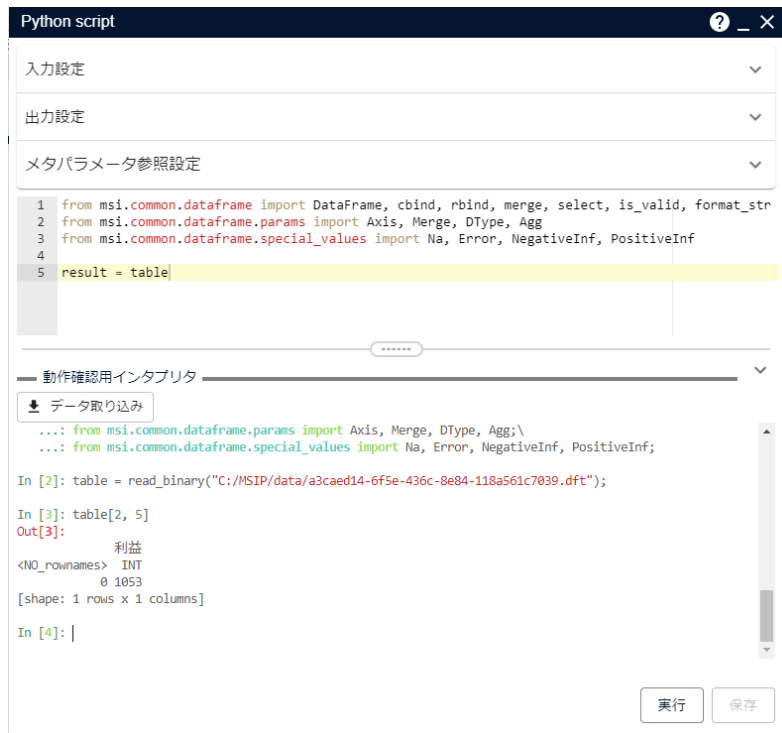


教師あり学習の手法に、勾配ブースティング決定木(Gradient Boosting Decision Tree)のモデリング機能を追加しました。

### ■ Python スクリプトアイコンにインタプリタ機能を追加



Python スクリプトアイコンの下段に表示されているコンソールを使って、IPython が利用できます。結果を確認しながらコーディングを行うことができます。



```
1 from msi.common.dataframe import DataFrame, cbind, rbind, merge, select, is_valid, format_str
2 from msi.common.dataframe.params import Axis, Merge, Dtype, Agg
3 from msi.common.dataframe.special_values import Na, Error, NegativeInf, PositiveInf
4
5 result = table
```

動作確認用インタプリタ

データ取り込み

```
...: from msi.common.dataframe.params import Axis, Merge, Dtype, Agg;\
...: from msi.common.dataframe.special_values import Na, Error, NegativeInf, PositiveInf;

In [2]: table = read_binary("C:/MSIP/data/a3caed14-6f5e-436c-8e84-118a561c7039.dft");

In [3]: table[2, 5]
Out[3]:
      利益
<NO_rownames> INT
              0 1053
[shape: 1 rows x 1 columns]

In [4]: |
```

実行 保存

## ◆ 追加・修正機能

### ■ ワークスペースブラウザ

- ログイン時やブラウザリロード時に作業中のワークスペース・プロジェクトを表示するようにしました。

### ■ データインポート

- Oracle データベースからデータを読みこめるようになりました。
- Excel ファイルから列や型を指定して読み込めるようになりました。
- 拡張子が txt や tsv であるファイルを読みこめるようになりました。

### ■ 前処理アイコン

- 集計アイコン
  - ◇ 不偏分散・不偏標準偏差を集計できるようになりました。
  - ◇ 集計対象列および集計キー列のそれぞれに対して、特殊値ごとに取り扱いを設定できるようになりました。
- クロス集計アイコン
  - ◇ 行ラベル・列ラベル列および内容列のそれぞれに対して、特殊値ごとに取り扱いを設定できるようになりました。
- 表の列名からソート（昇順、降順、クリア）や全選択、全解除（列がチェックボックス形式の場合のみ）、フィルタを行えるようになりました。
- 列名を選択する表で列型を確認しやすくなりました。

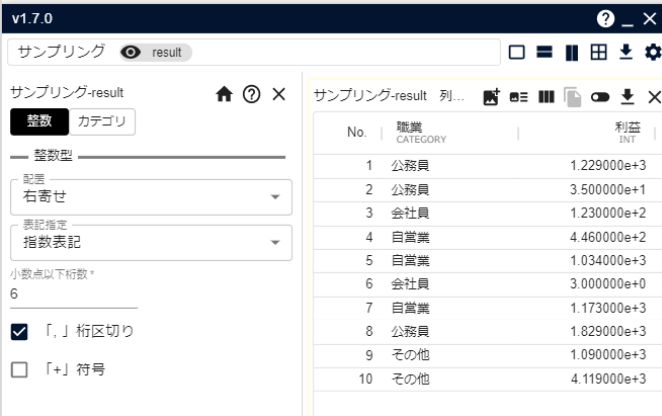
## ■ ワークフロー

- フロー実行時に各ノードのステータス（実行中や実行待機中）をワークフロー画面上で確認できるようになりました。
- ミニマップの拡大・縮小を保存できるようになりました。

## ■ 可視化機能

- テーブル

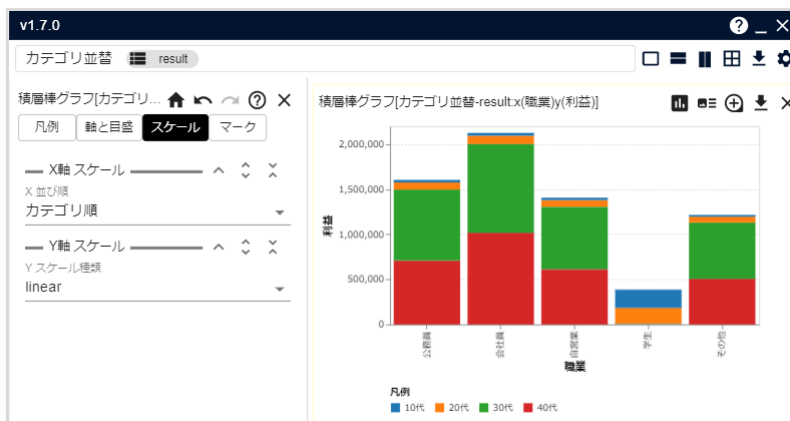
書式を設定できるようになりました。またカテゴリ列ではカテゴリ順に対して昇順・降順でソートできるようになりました。



No.	職業 CATEGORY	利益 INT
1	公務員	1.229000e+3
2	公務員	3.500000e+1
3	会社員	1.230000e+2
4	自営業	4.460000e+2
5	自営業	1.034000e+3
6	会社員	3.000000e+0
7	自営業	1.173000e+3
8	公務員	1.829000e+3
9	その他	1.090000e+3
10	その他	4.119000e+3

- グラフ

スケール変換、描画の点の大きさや線の太さ、補完、フォント種類などを設定できるようになりました。



- 可視化画面の結果表示高速化

結果を再表示する際、素早く表示されるようになりました。

## ■ MSI DataFrame API

### ➤ 新規 API

- ◇ pandas データフレームとの変換関数 `pandas_to_dataframe` と `to_pandas` を追加しました。
- ◇ 指定された 2 つの DataFrame の内容が異なるか判断する関数 `content_diff` を追加しました。
- ◇ 各行について複数の列の値の文字列表現を区切り文字で連結する関数 `strcat` を追加しました。
- ◇ 対象列の各行の値と相対行の値の差分や比率を計算する関数 `diff` を追加しました。
- ◇ VAP スクリプトの `chain_table` に相当する関数 `chain_table` を追加しました。
- ◇ 各列を指定した型に変換する関数 `as_type`、`as_int`、`as_float`、`as_string` を追加しました。

## ■ その他

- MSIP インストール環境の全データの保存 (dump) / 復元 (restore) が行えるようになりました。
- オフラインでライブラリをインストールすることで、インストールの安定性を向上しました。

以上