

NTT DATA

NTT DATA Mathematical Systems Inc.

数理システムユーザーコンファレンス 2018

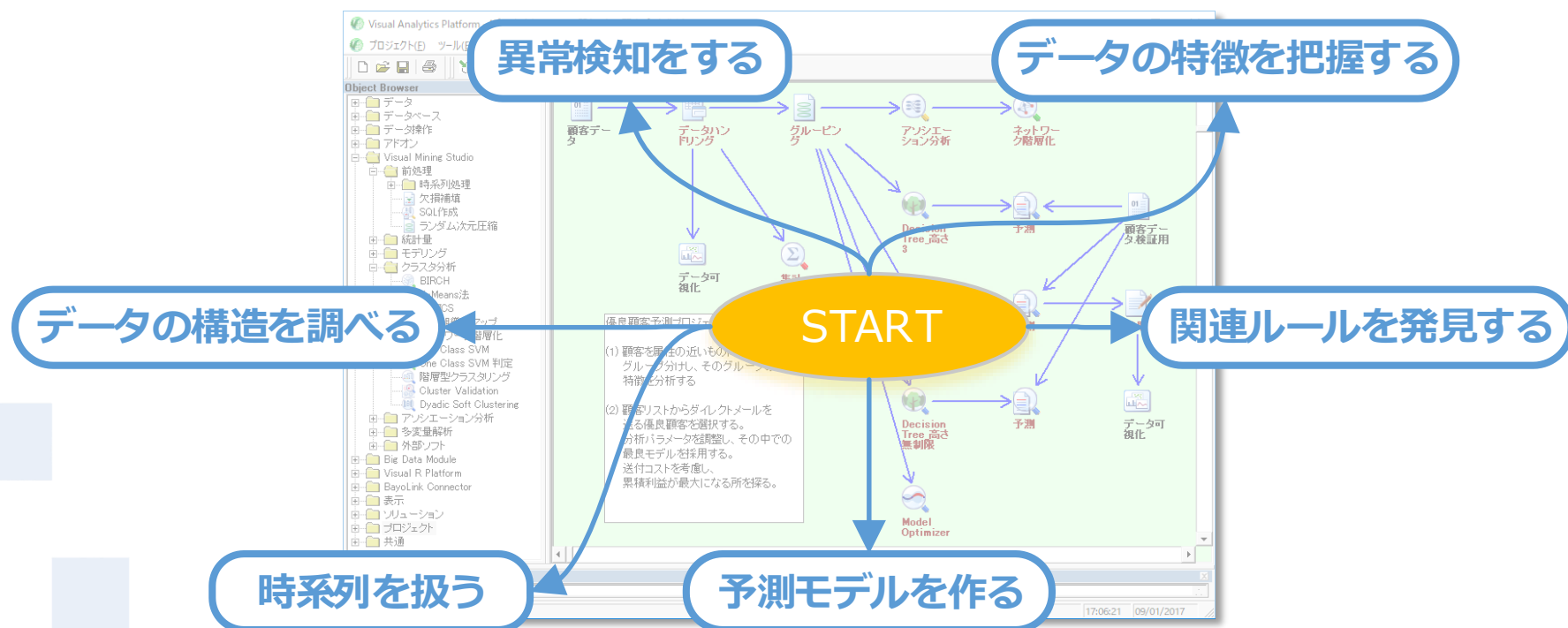
VMS ヒッチハイカーズ・ガイド

あいのり Big Data Module & CRM Insight

2018年11月22日

株式会社 NTTデータ数理システム

データマイニングツール **Visual Mining Studio** は、真にお客様のデータ活用をサポートする環境であり続けるべく、リリース以来進化を重ねてまいりました。

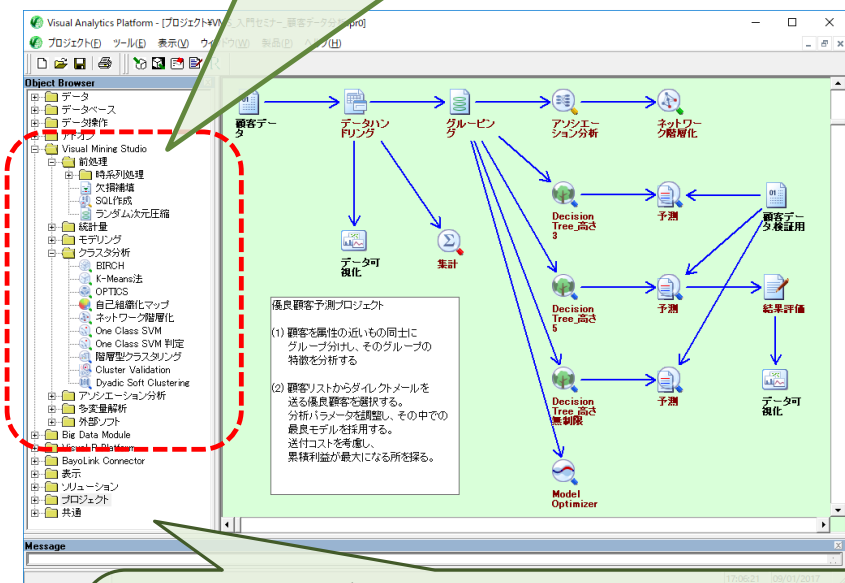


豊富な分析機能をご活用いただくべく、改めて上図のような **6つの目的別** に **VMS** の機能を総覧：**ガイド** し、紹介させていただきます。皆様にとって何か新たな発見がありましたら幸いです。

Visual Mining Studio

■ 特徴

定番から先進までを網羅
マイニング機能群



アイコン アイコン



と を矢印でつないで
処理の流れを定義する

ビジュアルプログラミング環境

業界最多 充実の分析メニュー

- 分類分析・クラスタ分析・アソシエーション分析を中心とした、多彩で実用的な分析機能群

大規模データへの対応

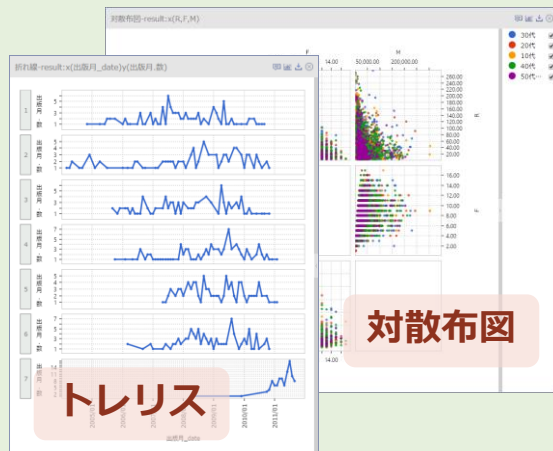
- 並列処理、64Bit対応による大量データの高速処理

マウス操作でデータ分析

- アイコンで処理の流れを定義
- 簡単な操作で試行錯誤が可能

データの特徴を把握する

データ可視化機能が2017年にリニューアルしより強力に、よりモダンに生まれ変わりました。表との連動機能も強化されており、強力なグラフ化機能がデータへの理解を助けます。

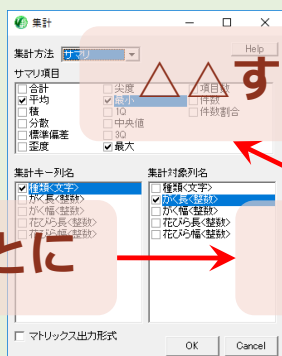


データは
どんな分布をしている？
項目間に関係はあるか？
 などをまず、これらの機能群で確認し、掘り下げていきます。

データ可視化

集計

集計は分析の基本であり最も利用頻度の高い機能です。



項目

する

対象列

××を

キー列

○○ごとに

データの特徴を把握する

データの内容表示・豊富なグラフ化機能

集計・統計量算出

集計&グラフ化

変数間の相関を見る

顧客のグルーピング

群間の差が大きくなる条件を抽出

有効な説明変数の探索

ビジュアル集計

相関

RFM分析

群間比較

Feature Selection

T 目的変数(教師データ)が必要

C カテゴリ値の目的変数に対応

N 数値の目的変数に対応

★は初心者の方にもおすすめ!

関連ルールを発見する

データの傾向把握を推し進め
項目間にどんな関連があるか？
そこから何らかのルールが抽出できないか？
といった観点で、
有用な知見の発掘を行います。

関連ルールを発見する

対話的に前提を探索

インタラクティブルール分析

時系列を考慮

時系列アソシエーション分析

指定アイテム間の
関連性を表示

関連性ダイアグラム分析

基本のアソシエーション分析

アソシエーション分析

結論の列を指定して
予測モデル作成

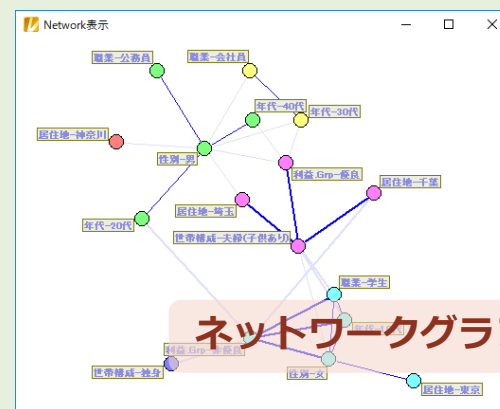
クラスアソシエーション分析

アソシエーション分析は、

事象A が起きていた時に
事象B も起きている、
その確率や頻度が高い
事象A, B のペア情報

を抽出します。

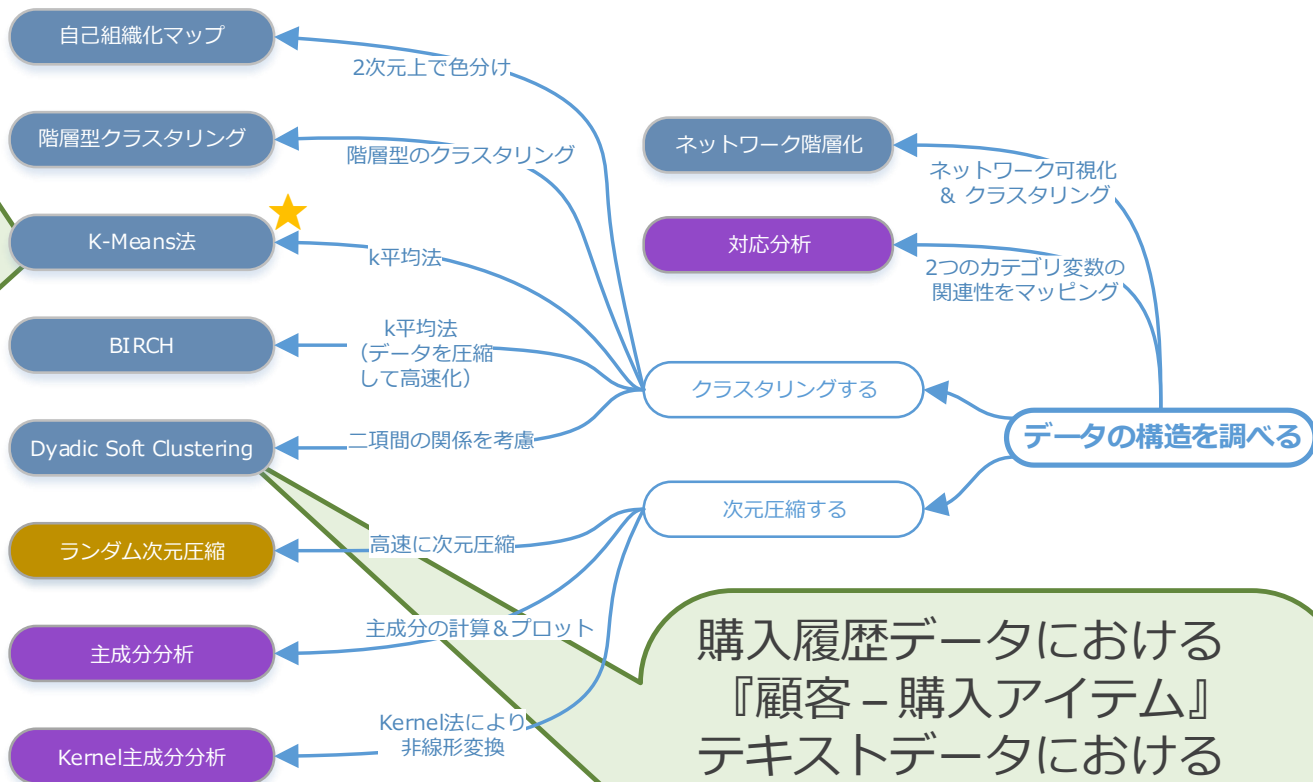
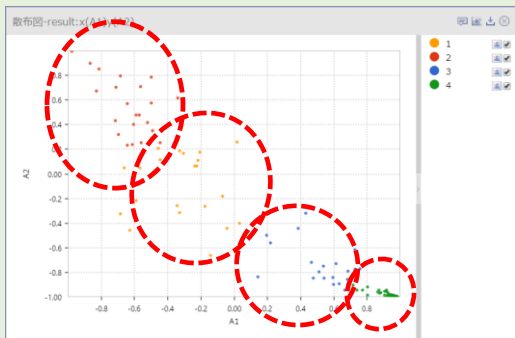
結果を **ネットワーク階層化**
を用いて可視化することも
理解の上で大変有効です。



ネットワークグラフ

データの構造を調べる

K-Means法は類似したデータのまとめ上げを自動的に作成する**クラスタリング**手法の代表的なものです。



さらに踏み込んで、次のような考え方でデータ内に潜む構造の抽出を試みます。

類似するデータのグループを自動的に作れないか？
複雑なデータをもっと単純に表せないか？

購入履歴データにおける『顧客 - 購入アイテム』
テキストデータにおける『発言者 - キーワード』
など、**2種類の項目の組**で特徴づけられるデータの把握には**二項ソフトクラスタリング**が大変有効です。

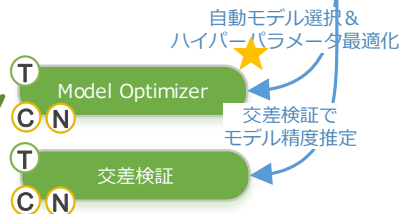
予測モデルを作る

機械学習の手法群を用いて
予測モデルを作成します

例：正常/異常を判別（分類）

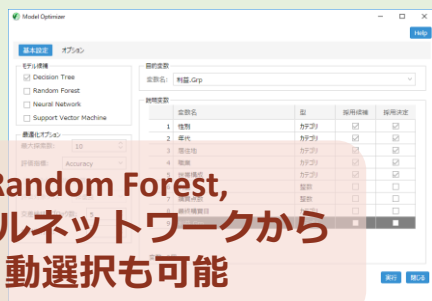
売上高を予測（回帰）

構築したモデルは
未知な状況の予測に
利用できます。

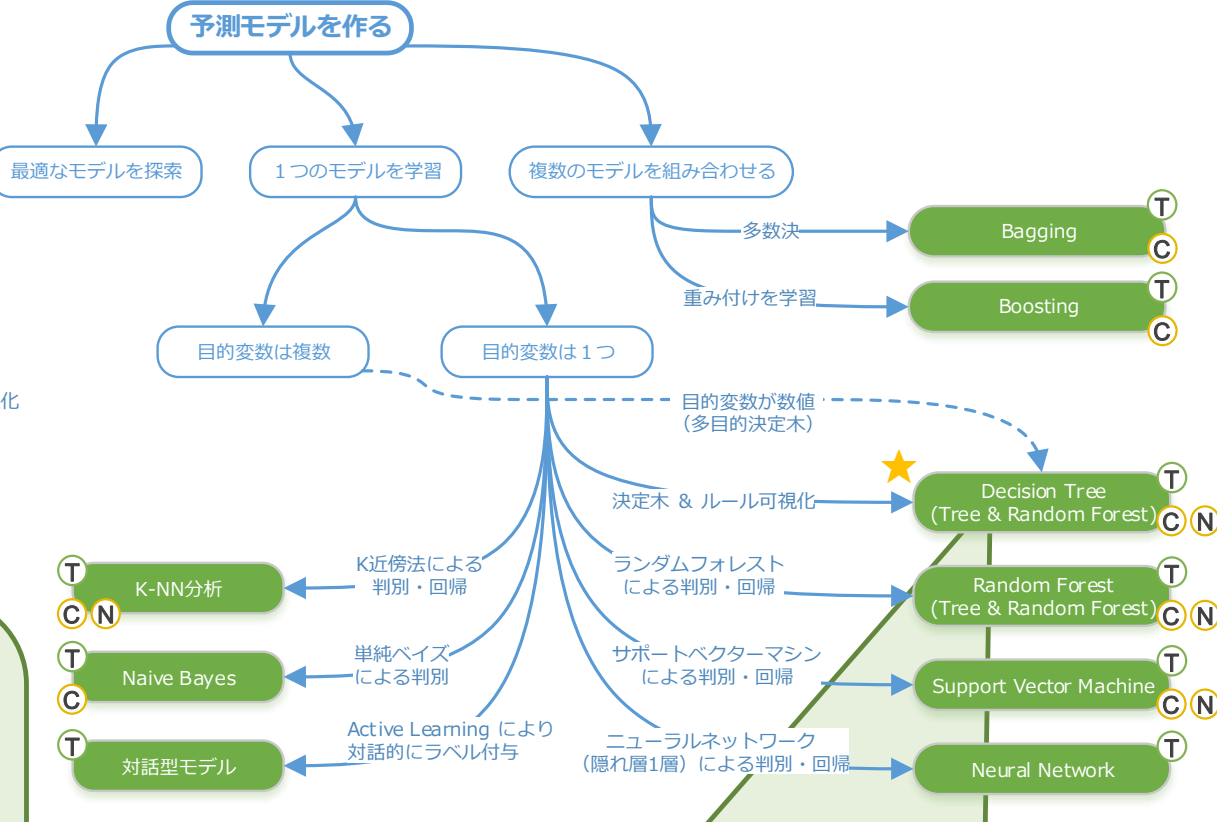


Model Optimizer では、
自動で効率的なパラメータ
探索を
実現します。

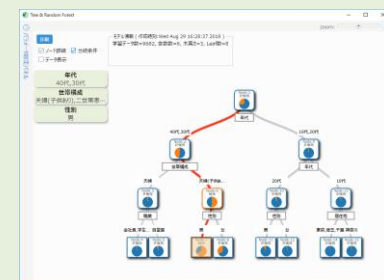
決定木, Random Forest,
SVM, ニューラルネットワークから
手法の自動選択も可能



予測モデルを作る



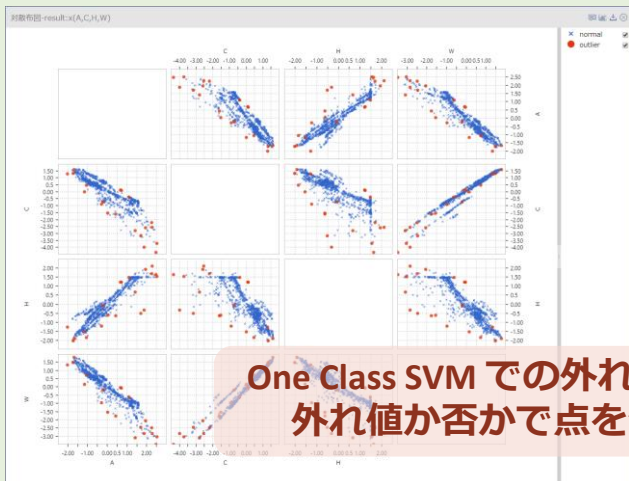
Tree & Random Forest では、
決定木により分類ルールの
抽出が可能、データの
傾向把握にも役立ちます。



異常検知をする

特にデータに正常・異常のラベルが与えられていない
教師なし異常検知の場合は、

データ全体の分布と照らし合わせて、外れているデータをいかに抽出するかという課題になります。



One Class SVMでの外れ値抽出例
外れ値か否かで点を色付け

主要データと外れ値の2群に分ける
ことに特化したクラスタリング手法
であると考えることもできます。

Isolation Forest

木構造ベースで高速処理

One Class SVM

教師なしSVM

OPTICS

密度ベース

異常検知をする

時系列を扱う

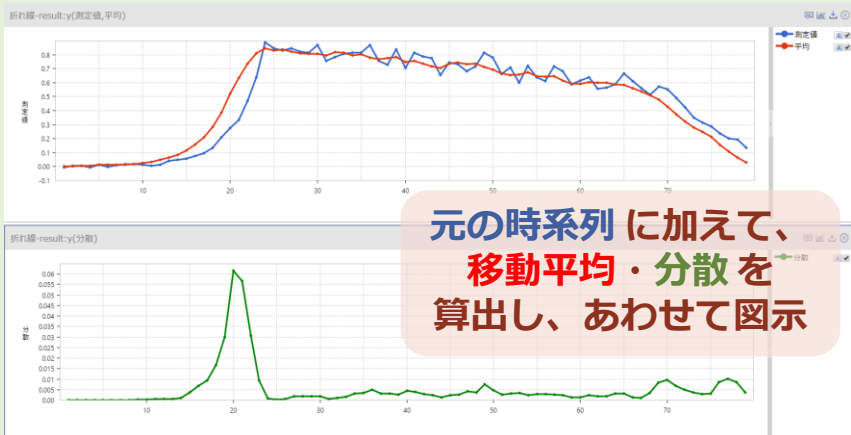
時系列データは、

各データ点同士が独立ではなく、その推移にも意味がある

という点に特徴があります。

時系列データの前処理・扱いに特化した機能を備えています。

時系列基本操作 機能で、移動平均や変動率などの情報を算出することができます。



時系列処理 / 基本操作

時系列処理 / 線分表現

時系列処理 / 距離計算

HSMM

推移列の展開・
移動平均等の計算

区分線形化

線分パターン
との距離

隠れセミマルコフモデル

時系列を扱う

- ご利用者の方に、**データに対する洞察**をより深めていただくための新機能を計画しております。
 - ▶ ある分類モデルができたとき、**そのモデルにおける説明変数の重要度合い**を評価する
 - ▶ データのみから、ある変数の挙動に対して特に大きく**影響を及ぼしている変数**はどれか把握する
- 本日紹介した **VMS**機能ガイド は、**VMS**ウェブサイト → サポート → Tips集 から **Visual Mining Studio チートシート** として公開しています。
- **VMS8.4** に同梱していた Deep Learning 機能は、新製品 **Deep Learner** として販売を開始いたしました。

関連ソリューション
Big Data Module
CRM Insight



Big Data Module

大規模な専用
コンピュータ
がいるんで
しょう？

ビッグデータ分析に
こんなイメージを
持っていませんか？

プログラムが
書けないと
分析できない
んでしょう？

操作が
難しいん
でしょう？

いいえ。BigDataModuleなら、プログラミング不要。
マウスだけの簡単操作でビッグデータ分析ができます！



簡単な使い方でビッグデータ分析

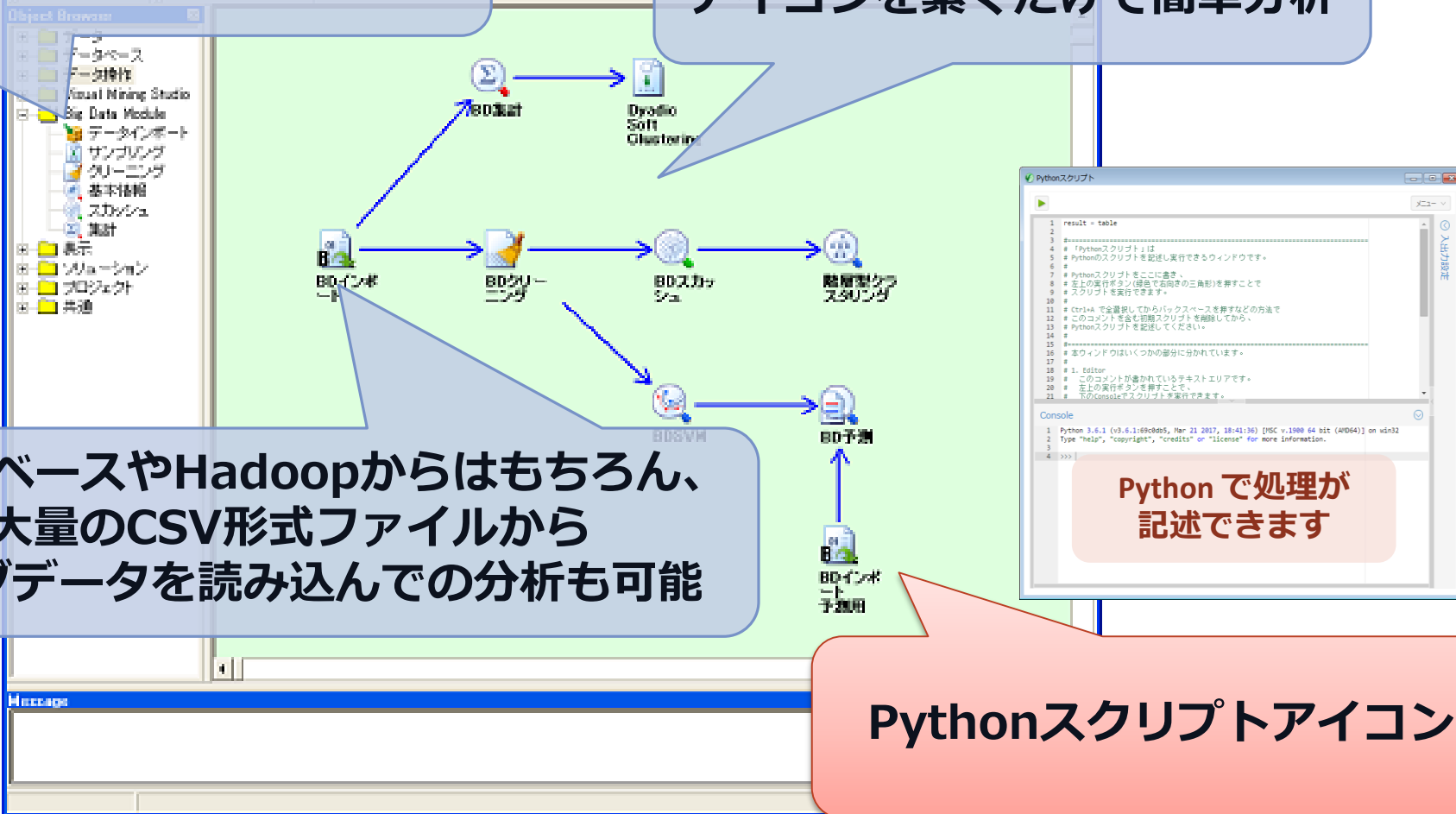
オンラインアルゴリズムによる
ビッグデータ分析アイコン群

Visual Mining Studio 同様
アイコンを繋ぐだけで簡単分析

データベースやHadoopからはもちろん、
大量のCSV形式ファイルから
ビッグデータを読み込んでの分析も可能

Pythonで処理が
記述できます

Pythonスクリプトアイコン





【使用例】大規模データの数値予測



売上予測



株価予測



電力需要予測

予測精度を高めるには...

- データ数を増やす
- 説明変数を増やす



計算時間の
爆発的な増加で
計算不可能

Big Data Moduleなら...

オンライン線形回帰でビッグデータでも予測可能

- データ数の線形オーダーの計算時間 → **超高速**
- データ数に依存しないメモリ使用量 → **超省メモリ**

CRM Insight とは？

ID-POS・顧客データに特化した分析システムです。
分析メニューをクリック！いつでもどこでも使えます。

ブラウザ上で利用できる
手軽なシステム

メニューを選ぶだけで
本格的な分析を実現

分析条件は
マウスで選択

分析結果がその場で
ぱっと表示



分析ラインナップ 全27種類 + α

分類	メニュー	わかること・できること
サマリ	全店概況集計	全店の重要指標が見える化
	月次概況集計	重要指標の月次推移
	基本集計	各種属性別集計
	売上クロス	売上との関係
	来客クロス	来客とその他の関係
顧客分析	デシル分析	お客様を売上で10分割して評価
	RFM分析	最近・沢山・高頻度購入のお客様
	優良顧客分析	ロイヤリティの高いお客様
	購買傾向 セグメント分析	購入商品を基準にした セグメンテーション
	会員ステージ分析	お客様をステージ分け
	顧客抽出	購買の様子をもとにお客様を抽出
商品分析	指定商品分析	この商品の顧客像を把握
	商品分類分析	分類ごとの購買傾向をチェック!
	売れ筋分析	売れ筋? 死に筋? 持続中?
	ABC分析	売れてる商品を見つけ出す
	併売分析	一緒に買われる商品を見つけ出す
	トライアル・ リピート分析	リピーターが多い? トライアル で終わってしまう商品?
	商品抽出	対策を練る商品を抽出

分類	メニュー	わかること・できること
トレンド分析	売上推移	これまでの売上変化と これからの変化
	来客推移	これまでの来客変化と これからの変化
	時間帯別 曜日別傾向	曜日と時間帯の深い関係
ポイント分析	ポイント集計	ポイントの利用集計
	ポイント推移	ポイントの変化
	ポイント 効果分析	ポイントの利用は次の 購買に結び付いたか?
キャンペーン 分析	キャンペーン集計	キャンペーンの利用者集計
	キャンペーン推移	キャンペーンの利用者推移
	キャンペーン 効果検証	キャンペーンの前後比較と 効果の検証
オリジナル 分析	独自分析・ 定常業務の分析	定常実行処理や、ならでは の分析を追加して運用

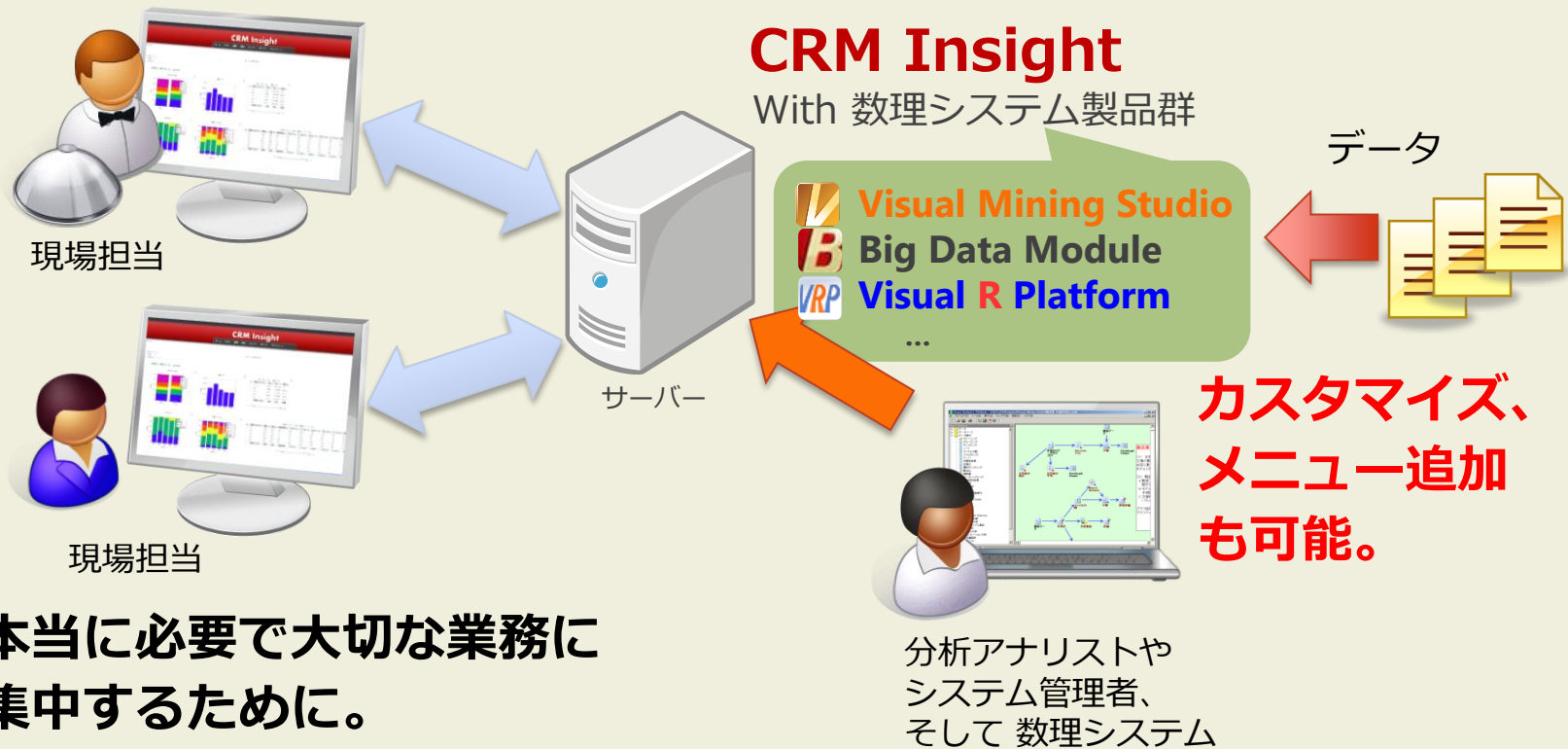
Excelなどで行っていた月次レポート
なども組み込んでしまえます

CRM Insight 2.4
皆様のご要望を反映しました

システム構成 ～ 場所を選ばず しかもパワフル ～

Webブラウザだけで
いつでもどこでも。
操作も簡単。

多くのメニューと
分析ノウハウを。





NTT DATA

NTT DATA Mathematical Systems Inc.

**Visual Mining Studio /
Big Data Module / CRM Insight 担当**

TEL : 03 - 3358 - 6681

FAX : 03 - 3358 - 1727

【URL】 <http://www.msi.co.jp/vmstudio/>

【E-mail】 vmstudio-info@msi.co.jp

Trusted Global Innovator

NTT DATA Group

NTT DATA