

NTT DATA

NTT DATA Mathematical Systems Inc.

新バージョン発表会 2021

異常検知ソリューション AD Analytics 新機能紹介

2021年12月15日

株式会社NTTデータ数理システム

Trusted Global Innovator

NTT DATA Group

NTT DATA

目次

1. 異常検知ソリューション AD Analytics のご紹介
2. 新機能紹介
 - ① 系列データの切り出し機能
 - ② 特徴量作成機能追加
 - ③ 教師ありモデル構築の手法追加

異常検知ソリューション AD Analytics のご紹介

あなたは日常生活や業務で、いかがお過ごしでしょうか。

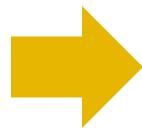
普段の日常生活、または業務においては通常通り遂行できてはいるものの時折、下記のような異常に直面することがあります。

時折経験してしまう「異常」

- ◆ 稼働中のサーバー上でDOS攻撃などの不正アクセスが見つかった。
- ◆ クレジットカードを紛失して不正利用された。
- ◆ 製造した部品を目視確認した結果、ヒビが入るなどの不良品が見つかった。
- ◆ 稼働中のシステムを温度センサーなどで監視した結果、ある一部で異常が見つかった。
- ◆ 病院で精密検査をして心拍数を測定した結果、疾患が見つかった。



このような異常は大抵は滅多に発生しない。
しかし、一度発生すると甚大な被害を被ってしまう。



- 異常が発生したら、早急に特定すべし
- 異常発生前に兆候をつかみ、対策を講じるべし

異常検知とは

異常検知とは、普段とは異なる挙動を見つける技術のことです。
(正常) (異常)

- ◆ 設備保守、製造ライン、社会インフラ、… 様々な分野で活用される
- ◆ データサイエンスや機械学習の技術を活用できる



< 製造業における異常検知の例 >

冷却装置のセンサー情報
と故障履歴



機械学習
モデル



早期の異常状態発見

気圧、湿度、気温などの
環境センサー



機械学習
モデル



配電盤の故障予知により機器の交換
をスムーズに

製造工程内のセンサーデー
タと製品の検査結果



機械学習
モデル



各センサーがどのような挙動を示すと不
良品になるのかを発見

AD Analytics の4つの特徴

時系列データの異常検知分析ツール

AD Analytics エーディーアナリティクス

時系列データの異常検知や要因分析に、機械学習を応用するための分析ツールです

ソフトウェア

- ① 分析シナリオテンプレートですぐにスタート
- ② ビジュアル分析環境と豊富なデータ加工
・可視化機能が、試行錯誤をスムーズにします
- ③ 試す価値のある機械学習や特徴量抽出の手法
を厳選してアイコン化

特別サポート

- +
- ④ 経験豊富な
分析コンサルタントが、導入を
手厚くサポート
(※1)

(※1) 分析コンサルタントによる特別サポートは別途ご契約が必要です。

AD Analytics エーディーアナリティクス

センサーデータなど、時系列データの異常検知や
要因分析に、機械学習を応用するための分析ツールです

1. 「前処理⇒特徴量作成⇒モデル構築⇒予測・評価」の分析シナリオテンプレート

独自の分析シナリオテンプレート用意しています。データを差し替えて、前処理からモデル構築、評価までの処理を実行できます。分析の試行錯誤をここから始められます。

2. ビジュアル分析環境と豊富なデータ加工・可視化機能

データの読み込みや分析に必要なデータ加工、データを理解するための集計や可視化機能が使いやすいアイコンという形で用意されています。典型的な処理ならばプログラミングは不要で、分析の試行錯誤をスムーズに行うことができます。

3. 試す価値のある機械学習手法や特徴量を厳選、アイコン化

自己相関やパワースペクトルによる特徴量作成、異常検知のためにカスタマイズされた Random Forest など、複数の有効な機械学習手法が厳選して実装されています。モデルの定量的評価、定性的評価を行う機能も備えています。

4. 当社の分析コンサルタントが分析ツールの導入を手厚くサポート※1)

AD Analytics の使い方や結果の見方、分析方法、お客様のデータ特有の性質を考慮したデータ加工方法などを、経験豊富な当社の分析コンサルタントが手厚くサポートします。

(※1) ツールの使い方に関するメールによる標準サポートが含まれます。
分析コンサルタントによる特別サポートは別途ご契約が必要です。

AD Analytics 概要紹介

ブラウザからサーバー(※1)にアクセスして利用します。

異常検知専用の分析シナリオを提供しますので、すぐに始められます。

自動でパラメータチューニングします。

試す価値のある機械学習手法を厳選!

データの読み込みから、前処理、異常検知モデリング、予測、結果出力までを兼ね揃えた、データ分析環境です。

ビジュアル分析環境と豊富なデータ加工機能・前処理機能を搭載。アイコンとして自由に配置できます。

結果をすぐに確認するための可視化機能を搭載

No.	ID INT	label CATEGOR	v1_acf01 FLOAT	v1_acf02 FLOAT	v1_acf03 FLOAT	v1_acf04 FLOAT
1	1	abnormal	0.896762	0.790838	0.684543	
2	2	abnormal	0.910672	0.810306	0.708679	
3	3	abnormal	0.885772	0.767499	0.649026	
4	4	abnormal	0.914827	0.795962	0.676881	
5	5	abnormal	0.906581	0.805671	0.703714	
6	6	normal	0.885772	0.779673	0.667925	
7	7	normal	0.885772	0.779673	0.667925	
8	8	normal	0.885772	0.779673	0.667925	
9	9	normal	0.885772	0.779673	0.667925	
10	10	normal	0.885772	0.779673	0.667925	
11	11	abnormal	0.895616	0.784631	0.672392	
12	12	abnormal	0.915207	0.823488	0.729647	
13	13	abnormal	0.908789	0.800136	0.691264	
14	14	abnormal	0.913189	0.812128	0.709560	
15	15	abnormal	0.900811	0.797514	0.694095	

※1 サーバーはオンプレミスです。サーバーにインストールしてお使いいただけます。

新機能紹介

- ① 系列データの切り出し機能
- ② 特徴量作成機能追加
- ③ 教師ありモデル構築の手法追加

異常検知分析で重要な「前処理」「時系列データの特徴抽出」「機械学習手法」のそれぞれで機能強化を行います。

- 新機能①「前処理」 系列データの切り出し機能
 - 分析の前処理にかかる時間を最小化するため、典型的なデータ加工をアイコン化
- 新機能②「時系列データの特徴抽出」 特徴量作成機能追加
 - 機械学習手法に入力する学習データの質を向上させるため、時系列データの特徴を抽出する機能を強化
- 新機能③「機械学習手法」 教師ありモデル構築の手法追加
 - 予測精度が高いことで知られる機械学習手法を簡単に利用したいというご要望にお応えするため、GBDTをアイコン化

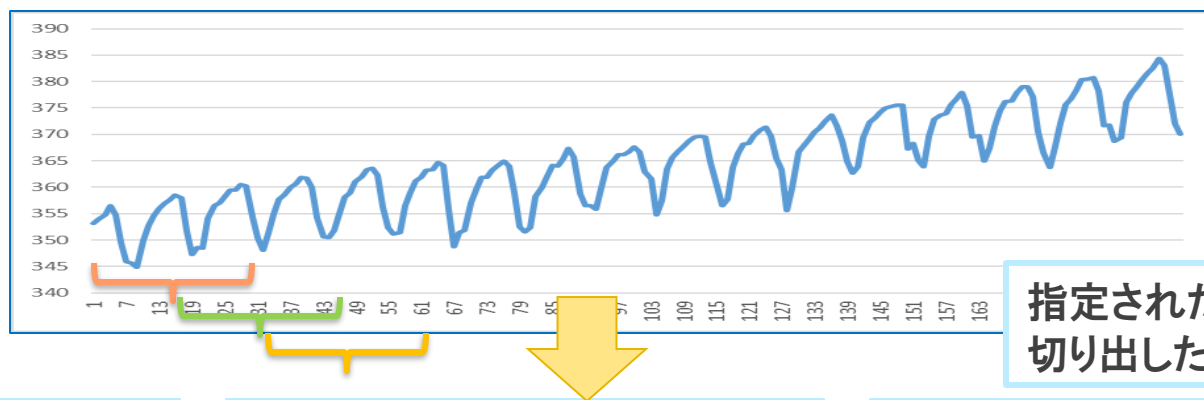
AD Analytics 新機能① 系列データの切り出し機能

AD Analytics 1.1.0(2021.12 時点最新バージョン) では、テンプレートの想定するデータは系列ごとに並んだ時系列データです。

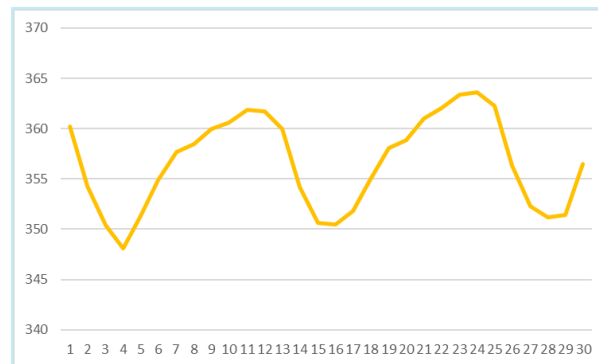
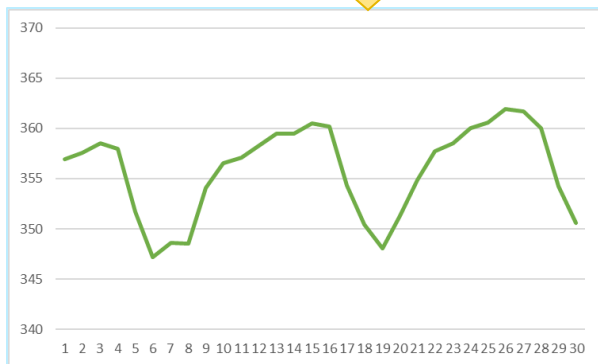
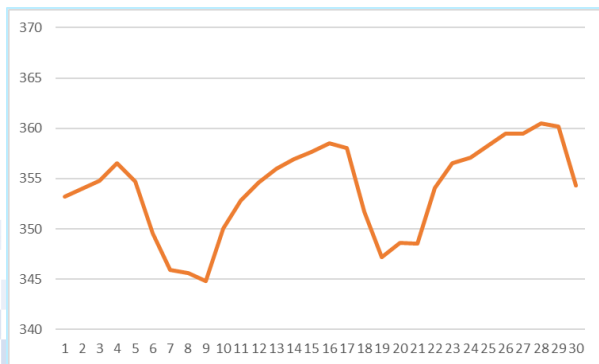
しかし、実際の計測データは、系列を切り出す前の計測値が行方向に計測された順に並んでいるデータが典型的であると考えられるため、系列の幅やずらし幅などを指定して**系列を切り出す、時系列データに対する前処理機能を追加**します。

この新機能により、テンプレートに入力するデータの加工の負担が軽減します。

計測された順に
並んだ計測データ



指定された幅でずらしながら
切り出した時系列データ



AD Analytics 新機能② 時系列特徴量作成機能追加

AD Analytics 1.1.0(2021.12 時点最新バージョン)の①「時系列特徴量作成」アイコンへの機能追加、さらに②局所的な周期の特徴を抽出する「時系列特徴量作成_BOSS」(仮称)アイコンを新しく追加します。

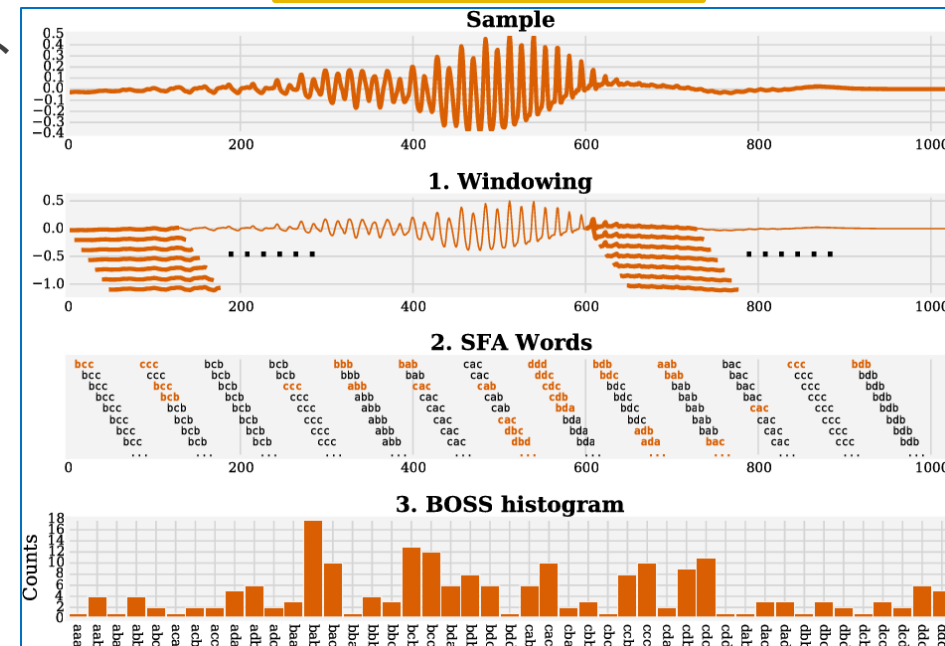
①では基本統計量(標準偏差、分位点など)、変数間の関連性を表す特徴量(相関、相互相関)を抽出する機能を追加します。

この新機能追加により、これまでは別アイコンなどを利用しないと作成できなかった特徴量が「時系列特徴量作成」アイコンで手間なく追加できます。

②では BOSS(Bag-of-SFA-Symbols) という手法を用い、時系列の局所的な特徴を抽出する機能を追加します。BOSS では時系列データの一部を順に切り出しながら文字列に置き換え、その出現頻度を特徴量として扱います。

この機能追加により、これまでは系列全体の特徴しか抽出できなかったものが、局所的な特徴を抽出できるようになります。

BOSSの処理イメージ

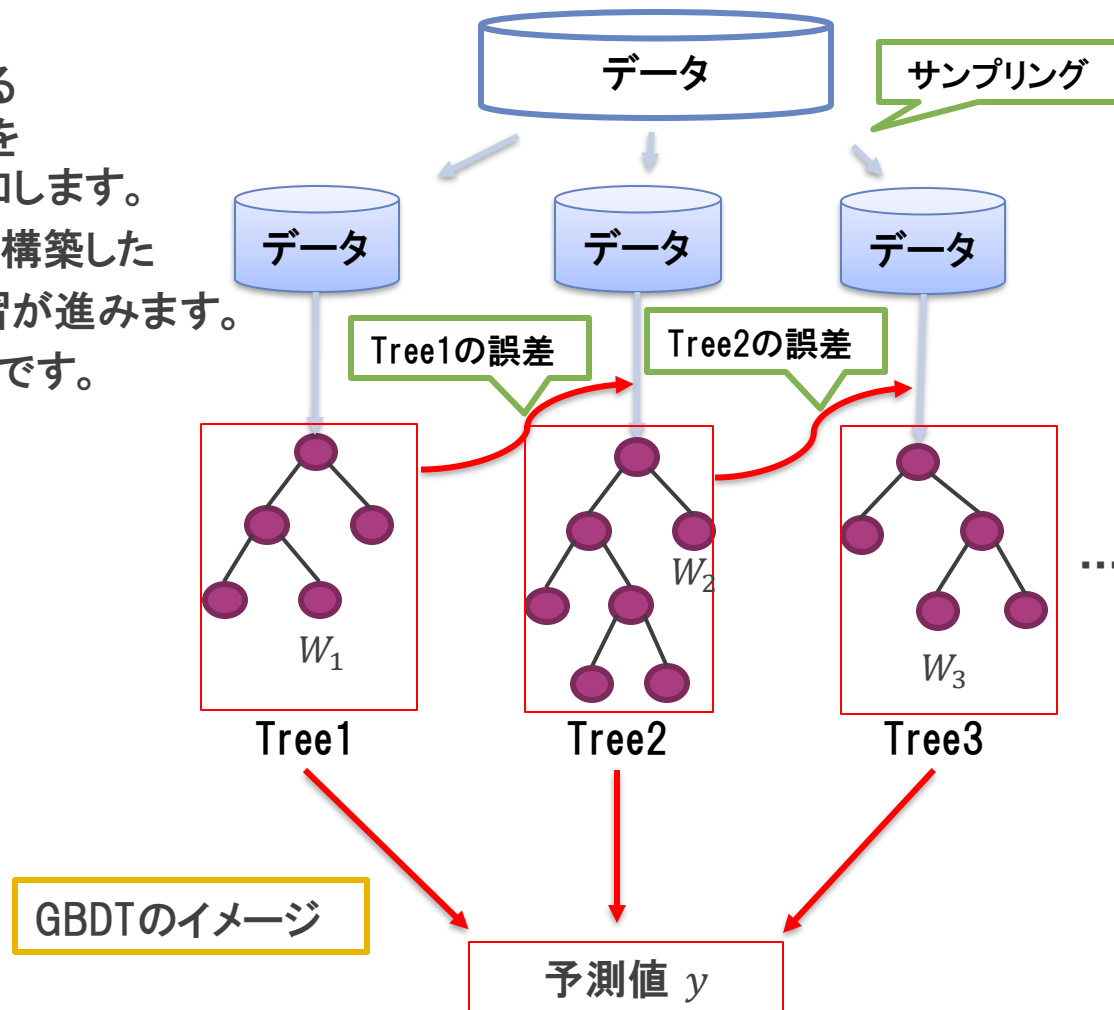


Schäfer, Patrick. "Bag-Of-SFA-Symbols in Vector Space (BOSS VS)." (2015).

AD Analytics 新機能③ 教師ありモデル構築の手法追加

AD Analytics 1.1.0(2021.12 時点最新バージョン)の「モデル構築(教師あり)」アイコンの機能追加を行います。

決定木にブースティングを適用した手法であるGBDT(Gradient Boosting Decision Tree)を教師ありモデル構築アイコンの機能として追加します。GBDTは決定木を構築する際に、それまでに構築した決定木による予測誤差を小さくするように学習が進みます。そのため、高精度のモデルが期待できる手法です。



AD Analytics 紹介セミナーを実施しております。(ご参加無料)
<https://www.msi.co.jp/adanalytics/seminar.html>
にて内容、お申し込み方法をご案内しております。

ご参加をお待ちしております。

NTT DATA

NTT DATA Mathematical Systems Inc.

NTTデータ数理システム AD Analytics 担当
adanalytics-info@ml.msi.co.jp

Trusted Global Innovator
NTT DATA Group

NTT DATA