

DataRobot が時系列データの予測分析を完全自動化 ビジネスパーソンが自ら未来予測を行える新製品を発表

DataRobotの機械学習自動化技術と昨年買収したNutonian社の未来予測技術が生み出した、時系列ビジネス問題解決のための新しいソリューション

米国DataRobot, Inc. (本社: マサチューセッツ州ボストン) は、本日、新製品「DataRobot 時系列アドオン」を発表しました。世界各国の70以上のユーザーと、世界最高クラスのデータサイエンティストとのコラボレーションによって生まれた本製品は、あらゆる種類の時系列データのモデリングを自動化します。これにより、現場の第一線で活躍するビジネスパーソンが自ら未来予測を行うことが可能になります。

時系列予測モデル構築の自動化の開始画面

これまで、時系列データの分析には非常に高度な技術が必要とされ、問題の定式化にも多くの困難がありました。典型的な機械学習の手法は、データセットから無作為に抽出した記録を用いて、予測モデルの構築と評価を行います。その時、学習用データは比較的近似度の高いレコードの集積であることを前提としていました。

時系列分析はそうではありません。時間的關係が重要となります。また考慮すべき多くの要素が存在します。例えば今から3日間の製品需要を予測したい場合、それは何曜日なのか？その日の需要はその週の他の日の需要とどれほど違いがあるのか？それ以前の週の同じ曜日はどうだったのか？その週はその月のいつに当たるのか？それは重要なのか？四半期内ではどうか？先月および前四半期と比較して、今月と今四半期は現在のところどうか？1年前の同じ月と四半期ではどうか？2年前から10年前まではどうか？その日は休日なのか、休日と平日はどう違うのか？どのような販促活動を行い、それをいつ実施したのか？過去の販促活動は、同様の時間的状況のもとでどのような成果をもたらしたのか？などの要素を考慮する必要があります。



特徴量名	インデックス	特徴量の型	ユニーク数	欠損	平均	標準偏差	中央値	最小	最大
マーケティング (actual)	0	テキスト	677	0					
売上 (28 day min)	0	数値	63	0	11,270	3,835	11,876	0	22,788
売上 (7 day diff) (28 day min)	0	数値	51	0	-1.27e+4	8,883	-1.31e+4	-3.31e+4	-1,383.70
売上 (diff) (28 day min)	0	数値	43	0	-1.07e+4	7,502	-8,916.79	-3.18e+4	-2,061.08
クリスマス期間 (28 day mean)	0	数値	45	0	2.85	6.48	0	0	24
従業員数 (14 day mean)	0	数値	97	0	7.32	1.66	6.79	5.07	15.29

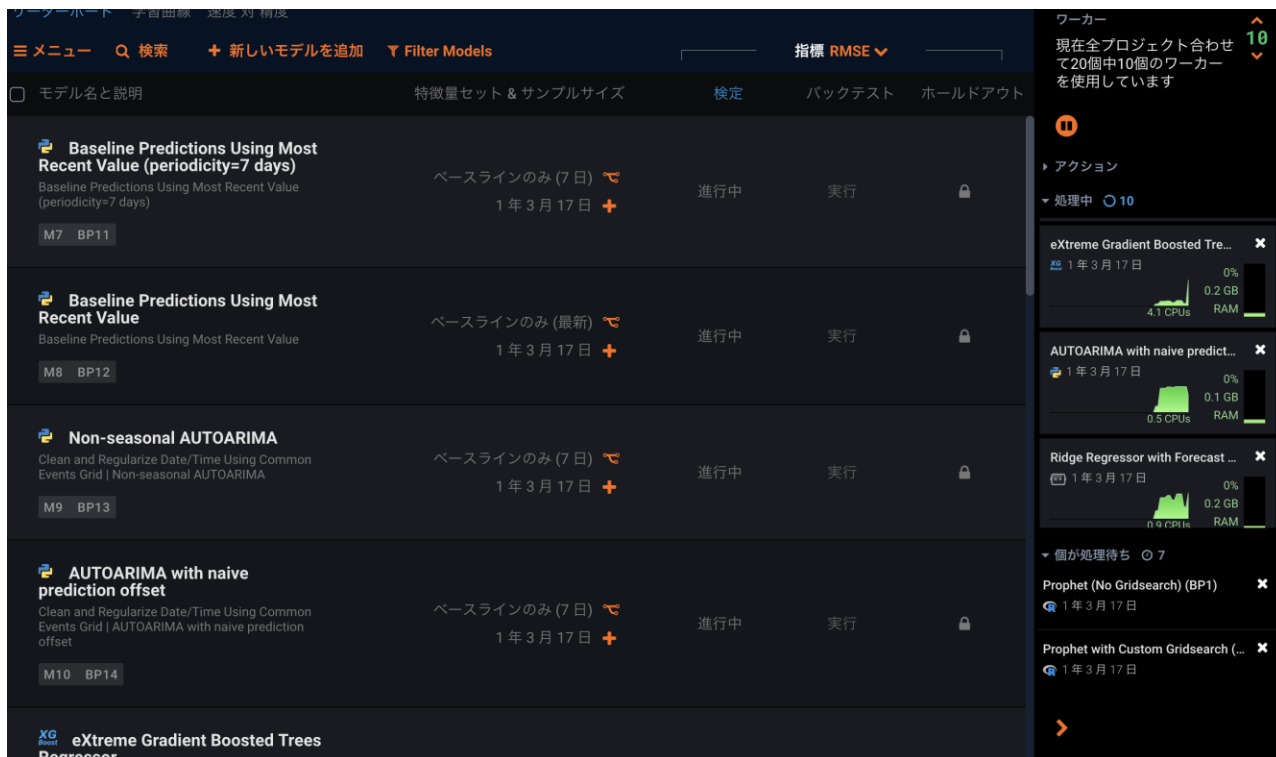
入力データの前処理の自動化により、様々な仮説に基づく特徴量を自動生成する

今回発表した「DataRobot 時系列アドオン」は、上記のような質問と、その回答に基づいていかに問題を整理すればよいのかを自動的に理解します。その後、DataRobot の自動化プラットフォームが多様な時系列モデルを構築し、学習データの各期間において安定した精度の予測をできるかを検証します。

「DataRobot 時系列アドオン」のベータ版ユーザー（フォーチュン2000に名を連ねる小売業者、銀行、病院グループを含む）は、既に人材配置、在庫管理、需要予測などにおける高精度な予測モデルの構築に成功しています。そのため手作業で行われるルール決めや、データサイエンスに関する特殊な専門知識、さらにプログラミングは必要ありません。

DataRobot のチーフデータサイエンティストであるMichael Schmidtは、次のように述べています。「予測はビジネスにおいて最も重要なアクションの1つです。未来を予測できれば、多くのケースで勝利を収めることができます。しかし、予測はデータサイエンスにおける最も難しい課題の1つでもあります。DataRobotは、昨年5月に Nutonian社を買収して以降、NutonianとDataRobotのイノベーションを統合して世界最高の時系列製品を作るという大規模プロジェクトに取り組んできました。開発過程においてユーザーとの大規模なテストが行われた本製品は、特徴量エンジニアリングなどの分野において最先端の自動化技術を駆使することで、超高精度の予測モデル構築を実現しています」

現在発売中の新アドオン製品には、ARIMAやFacebook Prophetといった時系列分析手法に加え、最新の機械学習モデルが含まれています。モデル生成と予測を実際の業務プロセスに活用するなど、既存システムへのインテグレーションを担うAIエンジニアは、APIから本製品の各機能を呼び出すことができます。

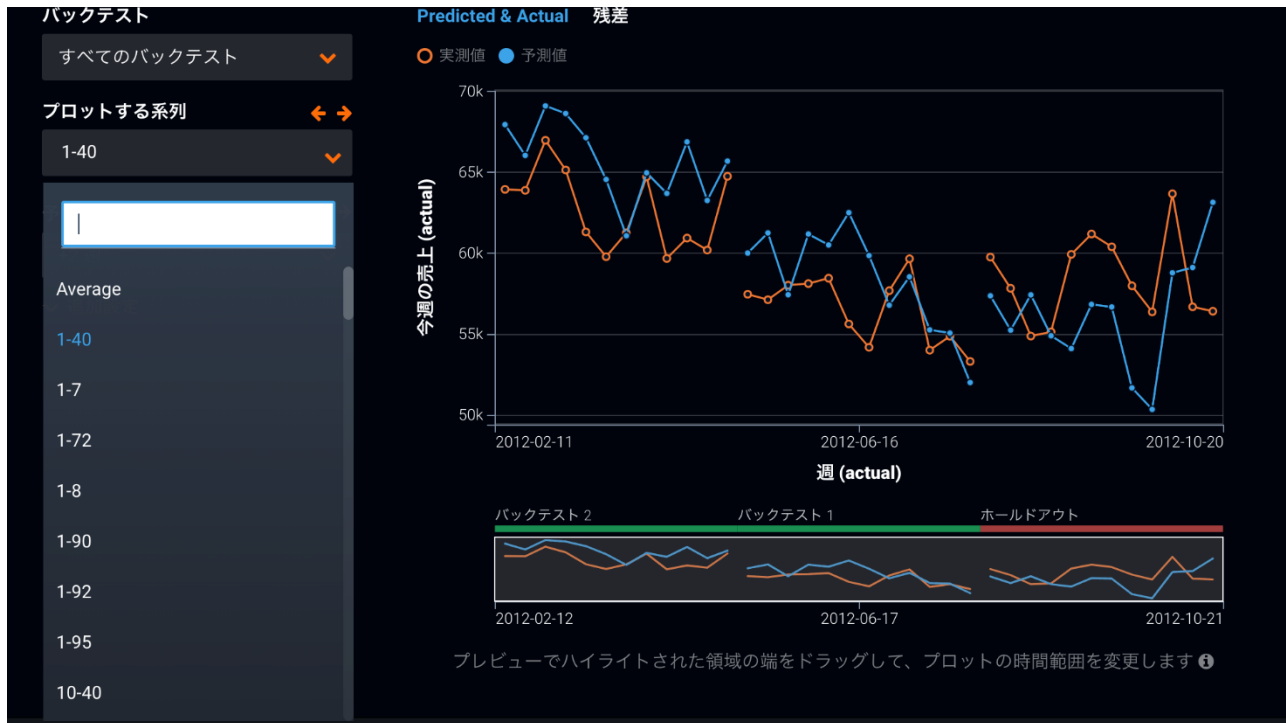


多種多様なアルゴリズムが同時並列で学習を進行

経営学の教授で、ハーバード・ビジネス・スクールのハーバード・イノベーションラボ所長Srikant Datar氏は次のように述べています。「時系列機械学習の自動化の試みはこれまで困難の連続でした。しかし、ここ数か月間、DataRobotのTime Series製品を利用し、金融アプリケーションの開発を行う中で、この製品の革新性とモデル構築の容易さに繰り返し感動しました」

米国で最大の営利私立病院を運営するSteward Health Careは、DataRobotを活用することで、国内38か所の病院ネットワークで大幅に業務効率を改善し、コスト削減を実現しています。同院の病院経営で発生する費用は、人件費のみで60%が占められています。DataRobotで患者数の予測精度を上げることにより、Stewardは38か所ある病院の8か所で1%の過剰職員配置を見直し、計200万ドルの人件費が削減される見込みです。こうしたコスト削減により、人命救助や医療の質の改善により多くの資源を投入することが可能となります。

Steward Health Careの情報システム兼ソフトウェア開発部門の総責任者、Erin Sullivan氏は次のように述べています。「我々は膨大なデータを持っており、そのデータの活用を目指してきました。DataRobotを利用することで、過去のデータから多くを学ぶことができるようになりました。DataRobotのおかげで、既にコストや時間の大幅な節約に成功しています。今回のリリースにも含まれる新しい機能により、これから看護職員の配置や患者の滞在時間の予測がさらに向上すると思われます。これらの課題の解決は、今後我々の病院に大きな利益をもたらします」



時系列予測の検証を複数系列データにおいても容易に行える

「DataRobot 時系列アドオン」や同製品のモデル生成および予測の機能についての詳細は、近日中にウェブサイトで公開予定です。英語の詳細は、<https://www.datarobot.com/add-on-products/time-series/>をご覧ください。

<DataRobot Inc.について>

DataRobotはあらゆる技術レベルのユーザーがAI・機械学習を迅速に活用しビジネス展開するためのエンタープライズ向け機械学習自動化プラットフォームを提供しています。数千種類におよぶパワフルなオープンソースの機械学習アルゴリズムから成るライブラリを組み込んだDataRobotのプラットフォームは、予測モデルのトレーニングと検証のプロセスを自動化し、またそれらを同時並列に行うことで、大規模なAI構築を可能にします。DataRobotはあらゆる規模の企業が、最速でAIの導入と活用を成功することを可能にします。

日本では、すでに大阪ガス株式会社、トランスコスモス株式会社、パナソニック株式会社、三井住友カード株式会社、株式会社リクルートホールディングスを含む国内100社を超える企業で導入され、成果をあげています。

詳細は<https://www.datarobot.com/jp/> ご覧ください。

※DataRobotは、日本国内におけるDataRobot, Inc.の登録商標です。