

※本ニュースリリースは DataRobot, Inc.、株式会社 N T T データ、Cloudera 株式会社 が共同で配信しています。重複して配信されることがありますが、ご了承願います。

2018 年 4 月 19 日

報道関係各位

DataRobot Japan
株式会社 N T T データ
Cloudera 株式会社

Hadoop 対応版の「DataRobot」を日本国内で提供開始
～増加する企業のビッグデータを迅速に分析可能に～

DataRobot Inc.（以下、DataRobot）、株式会社 N T T データ（以下、N T T データ）、Cloudera 株式会社（以下、Cloudera）は、4 月 19 日より、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である「DataRobot」の Hadoop（注 1）対応版の提供を開始します。企業の業務改革や価値創造に対する AI 活用の期待によりビッグデータ活用が求められる中、データを蓄積するプラットフォームとしてビジネス規模やデータ量の拡大に合わせて柔軟に拡張できる Hadoop 基盤を利用する企業が増えています。一方、データ分析を行う際には、データ抽出や集計する個別の DataMart（注 2）を構築したうえでデータを取り込む必要があり、DataMart に取り込まれない Hadoop 基盤上のデータを活用しきれないという課題がありました。Hadoop 対応版の「DataRobot」は、Hadoop 基盤から直接データを取り込み、Hadoop 基盤の処理能力を使って分析（各種の予測モデル構築、予測処理など）を実施できます。これにより、これまで企業の中で眠っていた膨大なデータに対しても、機械学習の自動化によるデータ活用が飛躍的に進歩することが見込まれます。今回、Hadoop 対応版の「DataRobot」は、Cloudera が提供するデータプラットフォーム上での動作検証を完了しています。

【背景】

昨今、企業の業務改革や価値創造に対する AI 活用の期待が高まりつつあり、膨大に蓄積されていくビッグデータを活用していくことが求められています。これまでは自社で得たデータのみを対象に分析を行っている企業が大半でしたが、これからはネット上の顧客行動履歴や SNS などの社外データ、IoT 技術で新しく獲得できるようになったデータなどを分析で活用し、高度な分析を行うことが必要になってきています。

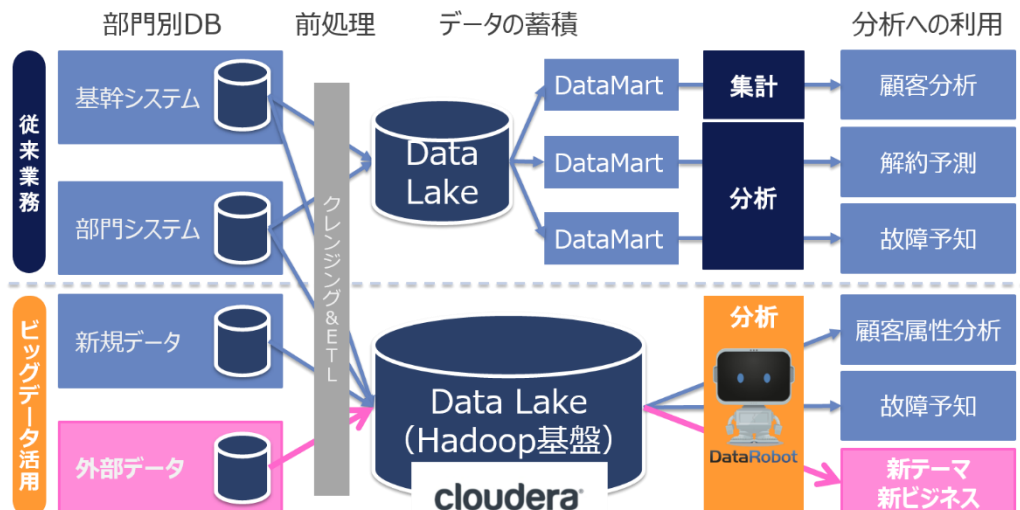


【概要】

Hadoop 対応版の「DataRobot」は、従来の「DataRobot」の特長をそのまま兼ね備えながら、データを蓄積している Hadoop 基盤から直接データを取り込んで分析できます。利用企業は分析用の DataMart を個別に構築する必要なく、Hadoop 基盤上で必要なデータの準備をして分析を行います。また、Hadoop 基盤上の処理能力を活用し、高速な予測モデル構築や予測処理が可能になります。

Hadoop 対応版の「DataRobot」は、日本国外では提供されていますが、2018 年 3 月末時点で

日本国内では提供されていません。今回、DataRobot の販売代理店パートナーであるNTTデータにより、Cloudera プラットフォーム上での Hadoop 対応版の「DataRobot」の動作検証が完了し、提供できる体制が整備されたため、4月19日より正式にサービス提供を開始することになりました。



【導入メリット】

デジタルマーケティングの領域においては、従来の大まかな市場予測を行うマスマーケティングではなく、多様なチャネルから集まる膨大な情報を基に、個人ごとの特性や好みに合わせて個別最適な提案を行う One to One マーケティングでの活用が期待されています。

また製造業の領域において、何千種類もある部品やプロセスごとに需要予測や不良品検知を行うことができるようになり、納品までのリードタイム短縮やこれまで以上の業務効率化が可能となります。今回拡張される機能と利用上のメリットは下記の通りです。

1. 外部データ/新規データを活用した新テーマ検討

Hadoop 基盤上に新しく蓄積した外部データや新規データを活用するためには、データから個別の DataMart を構築し分析を行う必要がありました。Hadoop 対応版の「DataRobot」は直接 Hadoop 基盤から分析データを取り込んで利用できるため、従来の抽出や集計、ダウンサイジングなどの DataMart 構築にかかる処理を簡素化し、これまで使えていなかったデータを使って、「新規ビジネスの創出」や「既存業務の改善」に向けた新テーマ検討を加速させます。

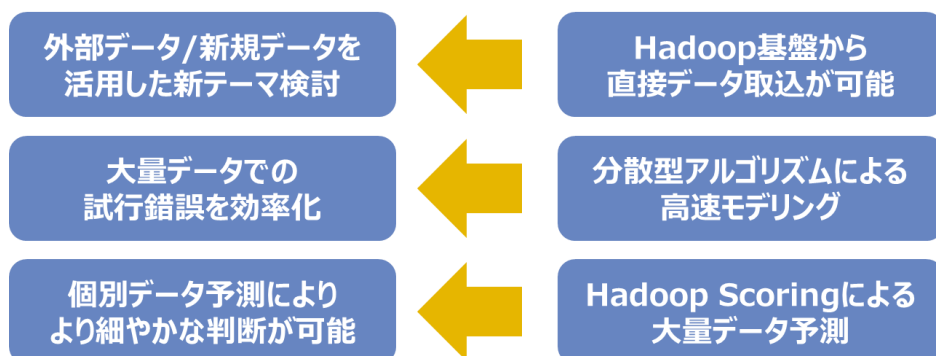
2. 大量データでの試行錯誤の効率化

予測モデルの精度を上げるためには新しいデータを入れて試行錯誤したり、学習データ量を増やしたりといった対応をとる必要があります。従来の機械学習アルゴリズムは単一マシンでの動作を前提としていたものが多く、単一マシンの処理能力が制約となっていたため、現実的なデータ量に落として予測モデル作成を行っていました。

分散型アルゴリズム(Spark MLlib, H2O 等)は Hadoop の規模が大きくなればなるほど分散処理によって短い時間で予測モデル作成を行うため、大量データに対しても現実的な時間で予測モデル作成が可能になります。これにより、精度の良い予測モデルを作るための試行錯誤が効率的にできるようになるというメリットがあります。

利用上のメリット

拡張機能



3. 個別データ予測により、より細やかな判断が可能

大量のデータに対して予測を実施したい場合にも、Hadoop Scoring という分散型バッチ予測の機能により、Hadoop 基盤上の分散処理の仕組みを有効活用して従来の「DataRobot」の予測エンジンよりも高い処理能力で大量の予測を行います。これにより、個々のデータ特性に合わせて予測を行い、より細やかな判断ができます。例えば、マーケティングにおいて何万件もいる顧客ひとりひとりに対して購買予測を実施することで、個別の傾向に合わせたアクションが可能になります。

【各社の役割】

<DataRobot>

- ・機械学習自動化プラットフォーム「DataRobot」の Hadoop 対応版提供

<NTTデータ>

- ・Hadoop 対応版の「DataRobot」に関するインテグレーションサービス提供

<Cloudera>

- ・機械学習と分析のためのデータプラットフォーム「Cloudera Enterprise」、ならびにプロフェッショナルサービスの提供

【参考】

<DataRobot について>

「DataRobot」は、世界トップレベルのデータサイエンティストの知識、経験、ベストプラクティスを組み込んだ機械学習自動化プラットフォームです。ユーザーは非常に正確な機械学習モデルを短期間で構築し、即座にシステムへ展開できます。Hadoop 対応版の「DataRobot」に対応する Hadoop ディストリビューションとしては、Cloudera と Hortonworks が選択可能です。

<Cloudera Enterprise について>

Cloudera が提供するエンタープライズ向けの機械学習と分析のためのデータプラットフォームです。今回提供を開始する Hadoop 対応版の「DataRobot」は、「Cloudera Enterprise」と組み合わせることにより、追加のモジュールを必要とせず Cloudera の管理ツールから一元的にインストール、ランタイムライブラリの配布、セキュリティ設定、データリネージ（データ追跡）などが容易に行え、安全かつ安定したエンタープライズ向けの運用管理環境が提供できます。

<Cloudera 株式会社について>

社名：Cloudera 株式会社

所在地：東京都中央区

Cloudera は、データの力によって、今日不可能なことを明日可能にできると考え、明確でかつ行動につながるインサイトを複雑なデータの中から得ようとする企業を支援し、クラウドに最適化された機械学習および高度な分析のための最先端のプラットフォームを提供します。Cloudera は最も困難なビジネス上の課題を解決する際のパートナーとして世界中の大企業から信頼を得ています。

<NTTデータの関連サイト>

NTT DATA Enterprise AI ～NTTデータが考える企業における AI 活用とは～

<http://www.nttdata.com/jp/ja/services/sp/ai/>

NTTデータ DataRobot 特設サイト

<http://nttdata-aistudio.com/datarobot/>

(注 1) Hadoop とは、大規模データの蓄積・分析を分散処理技術によって実現するオープンソースのミドルウェアです。

参照 URL：<https://oss.nttdata.com/hadoop/hadoop.html>

(注 2) DataMart とは、利用目的に合わせて個別に構築するデータベースを指します。目的に合わせているためデータ量も必要最小限で扱いやすい一方、運用にコストがかかる傾向があります。一般的にデータ量を圧縮するために、利用頻度が高い直近のデータのみを持ったのり、明細データを集計して持つケースが多いです。

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

* Cloudera および Cloudera のロゴは Cloudera Inc.の商標として米国その他の国に適用されています。

* その他の商品名、会社名、団体名は、各社の商標または登録商標です。

<DataRobot Japan について>

DataRobot Japan は、組織におけるデータ利活用を飛躍的に向上させる先進的なエンタープライズ機械学習プラットフォーム「DataRobot」の開発・提供を行う DataRobot Inc.の日本法人です。

2012年に米ボストンで設立された DataRobot inc.は、設立以来、一貫して機械学習の自動化による AI の民主化を提唱しており、現在では、アメリカ、イギリス、ウクライナ、シンガポール、オーストリア、ベラルーシ、インド、そして日本でビジネスを展開しています。

わかりやすい UI による優れた操作性と約 2000 を超える高度な分析アルゴリズムを備えた DataRobot 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幅広い企業で導入され、成果をあげています。

さらに詳しい情報はウェブサイトをご覧ください。 <https://www.datarobot.com/jp/>

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