

リテール向けレイクハウス： 変革の時代のためのレジリエンスと俊敏性

データとAIによるコスト削減と
ショッピング体験のパーソナライズ



databricks

目次

小売業界の急速な変化	3
需要と供給のディスラプション	4
最前線でデータと AI を活用	6
小売ブランドが抱える主要な課題	7
リテール向けレイクハウスの必要性	8
ソリューションアクセラレータで、データと AI の成果を加速	12
まとめ	14

小売業界の急速な変化

リアルタイムでパーソナライズされた顧客体験は、過去10年間で「あったほうがよい」から「必ず必要」に進化しています。

パーソナライゼーションは、単にスマートレコメンデーションのためのものから、それぞれの顧客に対してどのような知覚価値を提供するのかを決定するための重要な要素に変化しています。パンデミックの影響によって予測不能となった経済状況下でこの変化は加速し、爆発的な変革を起こしています。多くの小売企業が経済的なリスクを抑えるべく、劇的なまでに、軸足の転換を余儀なくされています。既に強力なeコマースプレゼンスを持っていた大手小売企業は市場シェアを拡大し、差別化に苦心していた小規模ブランドは、多大なコストをかけてオンライン化を強いられるか、リストラの危機にさらされています。デジタル需要が急増し、かつてないほど多くの顧客がeコマースに依存するなか、需要を満たすスケーラビリティを確保することは極めて困難になると同時に、その重要性はますます高まっています。

2020年、eコマースの売上は飛躍的なスピードで加速し、これまで10年かかっていた小売業の成長が10週間単位で行われるようになりました。突然、小売業者は戦略を改革し、データ分析、消費者の取り込み、予測、在庫やサプライチェーンのチェックポイントの追跡など、まったく新しい方法を導入する必要に迫られています。このディスラプションの津波を生き残る企業は、パーソナライズされた魅力的な顧客体験をあらゆる場面で提供する企業でしょう。しかし、そのためには、高度な分析とAIを活用したプラットフォームと手法を採用する必要があり、しかも指数関数的なスケーラビリティと絶えず変化するスコープをサポートしなければなりません。



需要と供給のディスラプション

今日における意思決定の中核には不確実性が存在するため、小売・消費財企業は、データの取り込みと解釈の方法に関して、スピードと俊敏性を重視しなければなりません。顧客が望むものを常に入手できるように、データ収集を高速化する必要があります。サプライチェーン全体のコスト削減、リソースの無駄を省くためのデータ分析の高速化、動的なデータに対応するための柔軟性も必要です。

SKUの合理化

在庫を最適化し、データに基づいて意思決定を行う部門横断的なプロセスです。過去のデータを活用した、生産コストや在庫量と潜在的な販売利益の比較検討は、以前は月に1度程度しか実施されていませんでした。しかし現在では、小売業者は、需要の増大にあわせた戦略的な規模の拡大のため、店舗レベルのSKUを日々最適化する必要があります。小売業者がSKUの合理化を行うために必要なモデルは、1か月あたり数万個から数十万個に増加したと推定されます。

合理化の実行に必要なモデルは、1か月あたり数万個から17億個以上に増加したと推定されます。パンデミックの影響下では、需給パターンが大きく変動しています。消費者の波は、異なる、あるいは不規則な購買決定を行い、需要を不正確に反映させています。また、サプライヤーは、安全面や経済的な理由から人員削減を余儀なくされ、生産が滞る事態を経験しています。

企業が、コストを削減しつつ市場シェアを獲得し、さらに成長するには、消費者行動を理解するための新たな手段が必要です。ビッグデータとAIを活用することで、小売／消費財企業は購買行動をリアルタイムで評価できるようになります。このことは、速やかに価値を創出する、将来的な成長が見込める領域への集中を可能にします。

スターバックス、データを高速かつスマートに活用して俊敏性を維持

コーヒーチェーン世界大手の米スターバックス社は、世界中のコーヒー愛好家に日々サービスを提供しています。スターバックスは、最高水準の顧客サービスを維持するために、顧客との継続的なつながりの構築実稼働環境イノベーションの創出、顧客のデジタル体験の改善に注力しています。データの活用と、従来型のオンプレミスソフトウェアからクラウドベースのプラットフォームへの移行が、スターバックスにとっての成功のカギとなりました。

Databricksの導入により、企業全体で活用できるデータレイクハウスを整備し、ペタバイト規模の高速データパイプラインを構築して、在庫管理を改善するMLモデルを迅速に構築し、新しい製品やサービスのイノベーションを実現することが可能になりました。

1000 以上

データ
パイプラインの数

50~100 倍

データ処理を
高速化

15 分

MLモデルの
展開に要する時間

需要と供給のディスラプション

小売業界では既に、ソーシャルディスタンスのガイドラインを守りつつ、どこでどのように顧客にサービスを提供すべきかを再考しており、その結果、ほとんどの小売業でBOPIS/BORIS取引（オンラインで購入し、店舗で受け取り/返却する）が急増しています。パンデミックの影響下で、カーブサイドオーダーは208%増加し、自宅待機中の消費者が買い物のためにデジタルストアフロントを利用したため、オンラインセールスが約50%増加しました。

“

Databricksを導入したことで、データ分析に戦略的な視点を取り入れられるようになりました。データを単にA地点からB地点に移動させるだけでなく、バリューチェーンの上位にあるビジネス上の課題解決に注力できるようになりました。”

Vishwanath Subramanian 氏
スターバックス データエンジニアリング・アナリティクス部門ディレクター



パンデミックが小売業の長期的成長におよぼした影響

パンデミック以降、これまで活況を呈していた多くの分野が、販売量の大幅かつ急激な減少や取引先の消失により、苦境に立たされました。このような収益と利益の損失は、サプライチェーン全体の労働力の減少につながり、景気後退に伴って消費者の財布の紐が固くなるにつれ、不安定さを助長しています。しかし、一部の業種では、ニューノーマルの直接的な影響として成長が見られ、今後12~24か月はその成長が継続すると予測されています。

- 食料品店
- ドラッグストア
- ホームセンター
- eコマース
- 消費財（例：食品、日用品など）
- 飲食店 eコマース

最前線でデータとAIを活用

小売・消費財市場は流動的であり、正確な予測を行うには、日々の商品需要や流通の変動を考慮したソリューションが必要です。しかし、従来のデータウェアハウスのツールのケイパビリティではこのニーズを満たすことができません。日々増え続けるデータを扱う組織には、サプライチェーンの全段階でのデータの取り込みに必要なテクノロジー、ツール、プロセスのオーケストレーションを可能にする、一元化されたハブ、すなわち、サプライチェーンコントロールタワーが必要です。時系列予測の変革は、より信頼性の高い需要予測を得るのに役立ちます。タイムリーで詳細な予測により、高精度なインベントリ管理が可能になります。

Databricks は、小売業や消費者ブランドのデータチームが必要とする全ての情報とツールを統合し、データとAIを簡素化し、オムニチャネルイノベーションを加速させます。

Databricks で精度を磨き、予測精度を向上させる

消費者は、モバイルや店舗でのショッピングを問わず、パーソナライズされたオムニチャネル体験を期待しています。Databricks は、顧客が何を求めているのかについての優れた知見と、サプライチェーンのリアルタイムビューを提供します。

ユースケース

消費者は、店舗、モバイルを問わず、パーソナライズされたオムニチャネル体験を求めています。

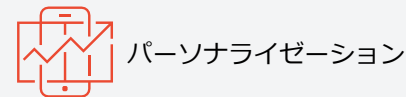
Databricks によって、小売業者や消費者ブランドは、顧客に対する 360 度の理解を真にサポートし、サプライチェーンをリアルタイムに把握できます。



サプライチェーン
インベントリ

顧客の需要を把握することで在庫予測の精度を高め、余剰在庫の削減と売上低下を回避

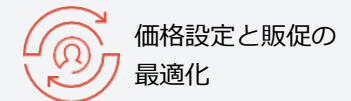
- サプライチェーンコントロールタワー
- 時系列予測
- 因果的予測
- 安全在庫
- 店頭在庫管理 (OSA)
- SKU の合理化



パーソナライゼーション

消費者の行動に関するより深い知見に基づくセグメンテーションで、収益の増大を促進

- 消費者セグメンテーション
- 顧客生涯価値
- 生存期間分析と離脱予測
- 購入傾向
- A/B テスト
- 推薦のパーソナライズ



価格設定と販促の
最適化

顧客ライフサイクルの重要な時期やシーズン価格の設定を最適化

- 動的な価格設定
- 価格の最適化
- 販促の最適化
- 販促の効果

小売ブランドが抱える主要な課題

大規模なデータ管理は容易ではありません。AIの管理は、それ以上に複雑です。顧客体験を取り巻く期待は確かに新たな高みに達していますが、こうした新しい基準に対応するための適切なテクノロジーに投資している小売業者はほとんどありません。まず、プライバシーとセキュリティは、現代のAIイニシアチブの基礎となるものです。ブランドロイヤルティはブランドに対する信頼に依存しており、顧客に対して価値を提供する限り、顧客は企業にデータを託します。自分のデータが安全であると感じられなければ、顧客は企業を信頼しなくなる（そして競合他社に移る）でしょう。

また、多くの企業が抱える問題として、スキルギャップがあります。優秀なデータサイエンティストの採用は、現在非常に競争が激しく、大企業でさえ、ビジネス全体のニーズに見合うだけの人材を確保するのに苦労することが多いのが実情です。同様に、企業には優秀なデータサイエンティストだけでなく、業界を理解する人材が必要です。規制の力学、競争の力学、消費者の行動などを考慮すると、データサイエンスは単に大量のデータを取得してアルゴリズムを書き出すだけのものではありません。データチームが相応の影響力を得るには、コストの削減や収益の増加など、ビジネス上の問題を解決する必要があります。

最後に、ビジネスのスピードが変化したことが挙げられます。組織は、PoCのテストに数か月や数年、そしてスケールアウトにさらに時間をかけている場合ではありません。数日から数週間で価値を提供する必要があります。価値創出までの時間を短縮する主な障害の1つは、過度に複雑で、性能、俊敏性、およびスケールのために構築されていない IT アーキテクチャです。



他の多くの業種と同様に、小売ブランドはこれまで、オンプレミスのITインフラ環境を個別のストレージと個別のサーバーを持つ別々のソリューションから構築し、その結果、アクセスや管理が複雑すぎる迷路のような状態を作り上げてきました。しかし、これらの環境のほとんどは、混沌としたテクノロジーの集まりと化して制御不能に陥り、維持するために膨大なリソースを必要とするようになりました。データエンジニアリングチームは、分析のためのデータを準備するよりも、レガシーインフラの管理に多くの時間を費やすことを余儀なくされることが多いのです。

基本的にレガシーIT環境は基礎的なレベルにおいて非効率的です。データを収集しても、現代の顧客のリアルタイムのニーズにはほとんど利用できないのが実情です。

リテール向けレイクハウスの必要性

顧客データは今日、急速に流入し、常に進化しています。毎日新しいチャネルが登場し、そこから得られる情報の賞味期限はどんどん短くなっているように思われます。貴重な知見を得るためのウィンドウは非常に小さいため、ブランドはその可能性を最大限に活用するために、データをジャーニーの最初の段階で統一する必要があります。

小売・CPG向け Databricks が実証される理由



厳しい SLA 内で高インパクトな分析を実現

Databricks は、お客様が最も要求の厳しい分析を、サービスウィンドウの最前線に提供することを可能にします。すぐに全ての店舗と SKU の在庫の割り当てや棚上の在庫の予測を開始できます。



リアルタイムなデータで体験をパーソナライズ

リアルタイムの認識は、レコメンデーションにおけるより高い関連性をもたらし、より高い増分と顧客満足につながります。Databricks は、あらゆる種類のバッチデータおよびリアルタイムデータを取り込み、eコマースやモバイル体験を強化することを可能にします。



俊敏性を高めてレジリエンスを強化

このような変動の激しい時代において、小売業や消費財メーカーには、変化にリアルタイムで対応できるプラットフォームが必要です。新しい知見と分析結果を、従来の倉庫のような数か月ではなく、数日から数週間で提供します。

Databricks は、小売・CPG 企業に対し、データ、分析、AI の全てのワークロードを統合し、イノベーションを加速させるフルマネージドのクラウドプラットフォームを提供します。Databricks は、全てのデータを実用化し、データ探索・分析の新しい方法を開放することで、新たなユースケースを推進し、信頼を築き長期的価値を創造する超パーソナライズされた体験を可能にし、イノベーションを加速させます。

Databricks プラットフォームは、レイクハウスのオープン性とシンプルさをもたらします。レイクハウスは、データレイクとデータウェアハウスの優れた要素を取り入れた新しいオープンアーキテクチャです。新しいオープンで標準化されたシステム設計によって実現されます。データウェアハウスと同様のデータ構造とデータ管理機能を、データレイクに使用されるような低コストのストレージに直接実装します。

高性能・信頼性データパイプラインを大規模に構築

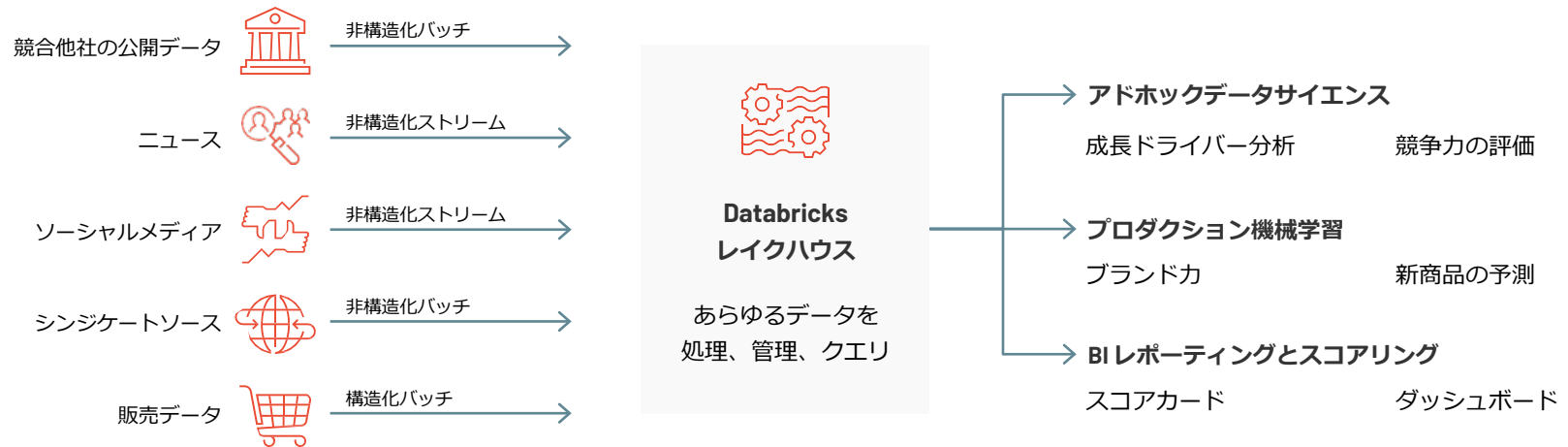
それぞれ異なる複数のソースからの膨大な実データを扱う組織の多くが、データの信頼性と品質の低下によって管理が複雑になるという共通課題を抱えています。多くの組織が、ビッグデータをコスト効率よく集約するためにデータレイクを利用していますが、これにも課題があります。

Delta Lake は、データレイクに信頼性と性能をもたらすオープンソーステクノロジーであり、Databricks にネイティブに統合されています。The Linux Foundation の中でも特に急成長しているビッグデータのオープンソースプロジェクトの 1 つです。Delta Lake は、データレイク上に位置するソフトウェアレイヤーで、データレイクに入るデータに対してデータ品質を強制的に向上させます。小売・CPG 企業は、構造化、半構造化、非構造化データを、バッチでもリアルタイムでも、単一の Delta Lake に取り込み、データの供給をクリーンで使用可能な状態にできます。

リテール向けレイクハウスの必要性

BI、データサイエンス実験、機械学習のためのデータレイクの価値を解き放つ

Databricks のレイクハウスプラットフォームは、小売および消費者ブランドにおけるデータと AI の取り組みを簡素化し、オムニチャネルイノベーションを加速させます。



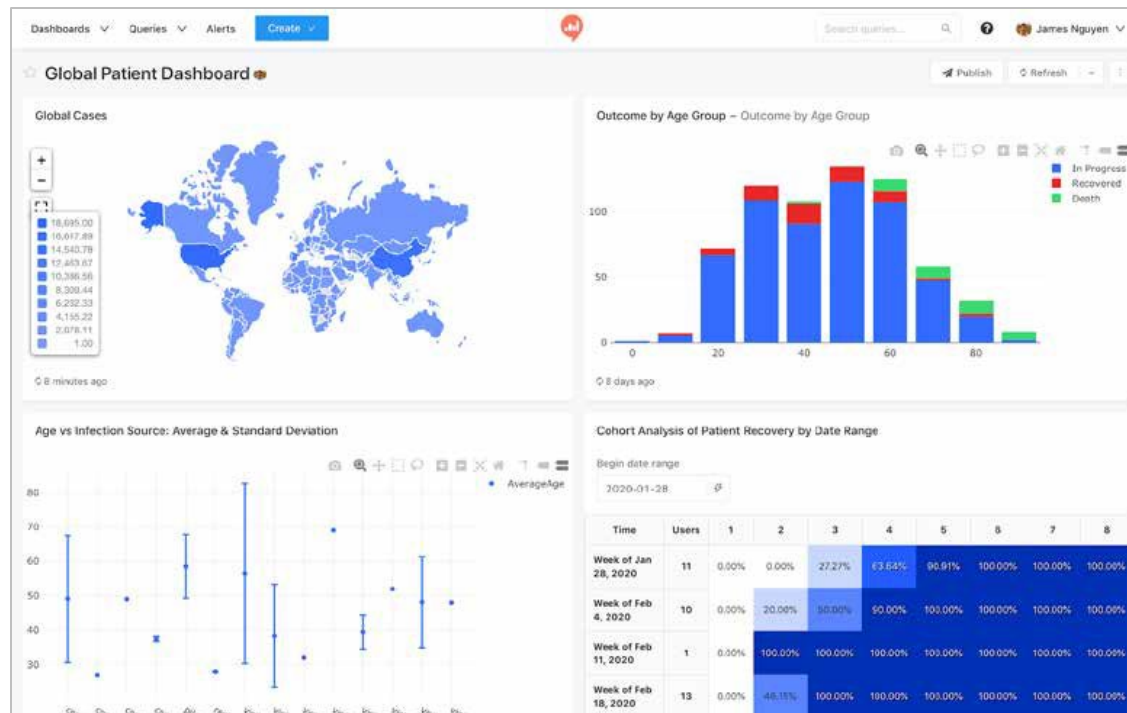
データの課題	Databricks のソリューション
データの取り込み ：バッチデータ/ストリーミングデータの処理が低速でエラーが発生しやすく、下流の分析に影響を及ぼすおそれがあります。	リアルタイムのインベントリデータとリアルタイムの顧客体験データを連携させます。
データレイクの管理 ：データのサイロ化が、顧客に関する完全なビューの取得能力を制限するおそれがあります。	複数のソース（クリックストリーム、POS、ソーシャルなど）からの膨大なデータを、堅牢なプライバシー基盤上で容易に扱うことができます。
データクエリ ：BIやデータサイエンスに必要なデータソースが分散、サイロ化し、一貫性を持たなくなります。	単一のプラットフォームから、迅速かつ安価な実験、管理、実運用規模のプッシュアウトが可能です。

リテール向けレイクハウスの必要性

分析のためのビジネスインサイトを抽出

Databricks ではデータの一元化と容易なアクセスが可能です。データアナリストは、Delta Lake と Spark SQL を使用してデータレイク内の完全で最新のデータに直接アクセスしてクエリを実行できます。また、コネクタが Tableau や Power BI などの主要な BI ツールとの連携を可能にします。アナリストは、任意の BI 可視化・レポートツールを介して顧客に関する知見をリアルタイムに取得できます。

シームレスなエクスペリエンスを実現するために、ネイティブに統合された可視化ツールである Redash を活用し、直感的なダッシュボードやクエリによってデータを容易に可視化し、共有できます。



Redash は、データの探索、クエリ、可視化、共有を容易にします。

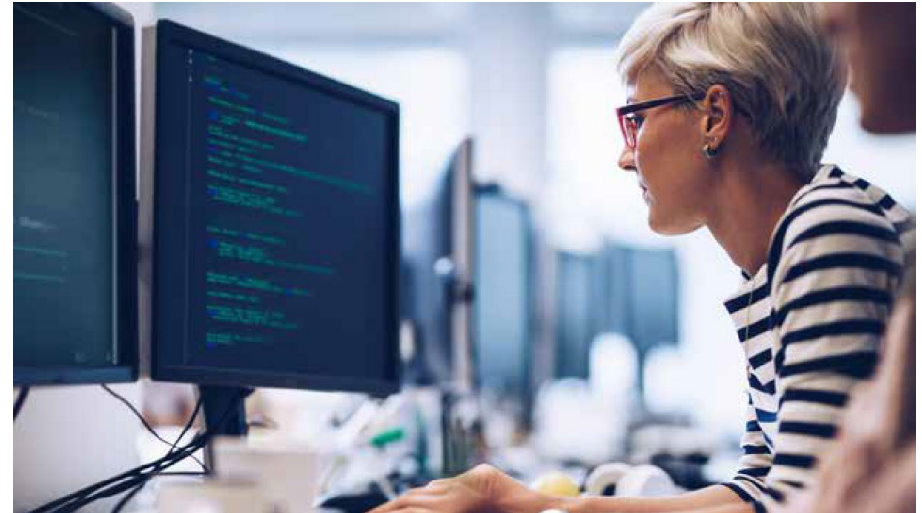
リテール向けレイクハウスの必要性

データサイエンスと分析チームをまとめ、イノベーションを加速

データドリブなイノベーションを迅速に実現するためには、データチームが企業全体でよりよく協力し合えるようなコラボレーション環境を醸成することが重要です。データチームが効果的に連携することで、アイデア、スキル、エネルギーの集約が容易になり、画期的な成果を達成できます。インタラクティブなワークスペースを通じて、データを介したコラボレーション、モデルやコードの共有、機械学習のライフサイクル全体の管理の一元化が可能になります。Databricks Notebook は、Python、R、SQL、Scala をネイティブにサポートしているため、ユーザーは任意の言語やライブラリを使用して共同作業ができます。また、組み込みのダッシュボードや視覚化機能を利用して、結果をビジネス関係者に示すことができます。

機械学習のライフサイクルを効率化して顧客価値を創造

機械学習モデルの構築とデプロイを成功させることは困難ですが、小売業者にとっては不可欠です。MLflow は、機械学習のライフサイクルを効率化するオープンソースのフレームワークで、データサイエンティストがパイプラインを再現し、異なるバージョンの結果を比較し、何がどこで実行されているかを追跡し、更新したモデルを再展開およびロールバックできるようにするものです。MLflow は Databricks レイクハウスプラットフォームにネイティブに統合されており、チームはデータパイプラインを開発中のモデルにシームレスに接続し、ML ライフサイクル全体でモデルを追跡することができます。



これらは全て、データサイエンティスト、データエンジニア、アナリストがビッグデータを探索し、パーソナライズされた魅力的な顧客体験を提供できる実用的な知見を引き出すことができる統合データ分析プラットフォームによって実現されます。

ソリューションアクセラレータで、データと AI の成果を加速

大手企業との取り組みから得たベストプラクティスに基づき、一般的なデータ分析や機械学習のユースケースに対応したソリューションアクセラレータを開発しました。データエンジニアやデータサイエンティストは、開発時間を数週間から数か月間短縮することができます。

リテール向けソリューションの最新リリースについては、[Web ページ](#)をご覧ください。

在庫の最適化



因果関係に基づく需要予測

eコマースの拡大、サプライヤーとの取引の変動、世界的なパンデミックのリスクは、サプライチェーンに衝撃を与え、需要を加速させています。企業は、需要予測やインベントリ管理に関する既存のモデルや手法が、小売業のニューノーマルに対して不十分であることに気づいています。しかし、消費者が簡単に購入先を変えることができる eコマースの時代で競争するには、企業は日単位および店舗レベルで SKU の需要を予測する能力を持つ必要があります。

ブログ、Notebook : [サプライチェーンの需要予測を改善する新手法](#)
オンデマンド Web セミナー : [実稼働規模での細粒度の需要予測](#)



時系列予測

時系列分析の速度と確度を向上させ、商品やサービスの需要予測をよりの確に行うことは、小売業者の成功にとって極めて重要です。この Notebook では、時系列予測の重要性について説明し、サンプル時系列データを視覚化した後に、簡単なモデルの構築を通じて Facebook Prophet の使用方法を解説します。1つのモデルの構築に慣れたら、次は、Prophet を Apache Spark™ と組み合わせて数百のモデルを一度にトレーニングします。これにより、今までほとんど実現されていない細かい粒度で、個々の商品と店舗の組み合わせについての正確な予測を取得できます。

ブログ、Notebook : [Facebook Prophet と Apache Spark による時系列予測・分析](#)
導入事例 (オンデマンド Web セミナー) : [スターバックス](#)



安全在庫

自然災害、パンデミック、社会不安など、最近、私たちのグローバルなサプライチェーンに混乱が生じています。需要に応えるために十分な製品を確保しつつ、過剰なインベントリを回避することは、あらゆるビジネスにとって重要な課題です。このソリューションは、小売業者やメーカーが、運転資金を解放すると同時にビジネスの停滞を回避するための安全在庫を特定する最新の方法を提供します。

ブログ、Notebook : [安全在庫分析の新しいアプローチでインベントリを最適化](#)

ソリューションアクセラレータで、データと AI の成果を加速

パーソナライゼーション



顧客生涯価値

マーケティング部門は、エンゲージ度が高く、より大きな価値をもたらす顧客にリソースを投入したいと考えています。このような顧客に投資することで、より大きな成長と ROI が期待できます。このソリューションアクセラレータは、マーケティング部門が消費者を生涯価値でセグメント化し、製品開発や販促のパーソナライズに関する意思決定を改善する方法に焦点を当てています。

ブログ、Notebook : [顧客の生涯価値 第1部：生涯価値の予測](#)

[顧客の生涯価値 第2部：消費の予測](#)

オンデマンド Web セミナー : [顧客価値の最大化](#)



顧客の残存率と顧客リテンションの予測

サブスクリプションサービスを提供する小売業者や消費者直販ブランドが増えていきます。サブスクリプションサービスは、消費者に利便性を提供すると同時に、安定した年間収益源となります。サブスクリプションモデルの会員数が増加するにつれ、顧客維持による収益性の確保が重要になります。このソリューションは、顧客維持につながる要因、顧客が離脱する時期や理由を特定するための、顧客分析の新たな方法を提供します。

ブログ、Notebook : [顧客離脱の分析方法](#)

オンデマンド Web セミナー : [顧客維持率向上のための離脱予測](#)



利益主導型のリテンション管理

効率的なリテンション管理を行って離脱率を低下させるには、顧客のライフタイムがどのように進行するのかを理解し、ライフタイムのどの時点で顧客が離れていきやすいかを検討することが必要です。より信頼性の高い離脱リスクの予測ができれば、個々の顧客に関連する残存 CLV をより慎重に検討し、いつ、どのように介入するかについて、的確な意思決定ができるようになります。

ブログ、Notebook : [機械学習を活用した利益主導型の顧客維持](#)



推薦エンジン

レコメンダーは、適切なコンテンツや SKU の発見、カスタマイズ可能な製品のパーソナライズなど、顧客の選択のナビゲートに重要な役割を担っています。Databricks を使用してさまざまなレコメンダーを開発する方法を理解していただくために、一連の詳細な Notebook をご用意しました。Notebook では、実世界のデータセットを利用して、未加工データを1つまたは複数のレコメンダーソリューション（協調フィルタ vs コンテンツベース）に変換します。

ブログ、Notebook : [レコメンデーションを活用した顧客体験のパーソナライズ化](#)

オンデマンド Web セミナー : [データドリブンのパーソナライゼーション](#)

まとめ

小売ブランドは、顧客を正確かつリアルタイムに理解するための多様なデータセットをリアルタイムに把握することで、直近の行動に基づいて適切な体験を提供し、カスタマージャーニーをパーソナライズできます。データエンジニアリングとデータサイエンスの統合が、変動の激しい今日の状況下における成功のカギとなります。

Databricks のリテール向けレイクハウスソリューションの詳細、導入事例、小売・消費財に関する最新のイベント情報などについては、[Web ページ](#)をご覧ください。

データブリックス・ジャパン株式会社
お問い合わせ：03-6821-1670（代表）
<https://databricks.com/jp/company/contact>

©Databricks 2022. All rights reserved. Apache、Apache Spark、Spark および Spark のロゴは、Apache Software Foundation の商標です。[プライバシーポリシー](#) | [利用規約](#)

