

Suntory における Citadel AI を用いた 予測モデルの運用事例

SUNTORY SYSTEM TECHNOLOGY

SUNTORY

2022年7月1日

中田恵史・高木基成

本日本話する内容

1 **会社紹介**

2 **機械学習モデルの共通運用基盤**

3 **物流分野におけるモデル運用**

本日本話する内容

1 **会社紹介**

2 **機械学習モデルの共通運用基盤**

3 **物流分野におけるモデル運用**

サントリーグループ紹介

飲料・食品事業



酒類事業



健康食品事業



外食・花・サービス関連事業



サントリーグループ製品紹介



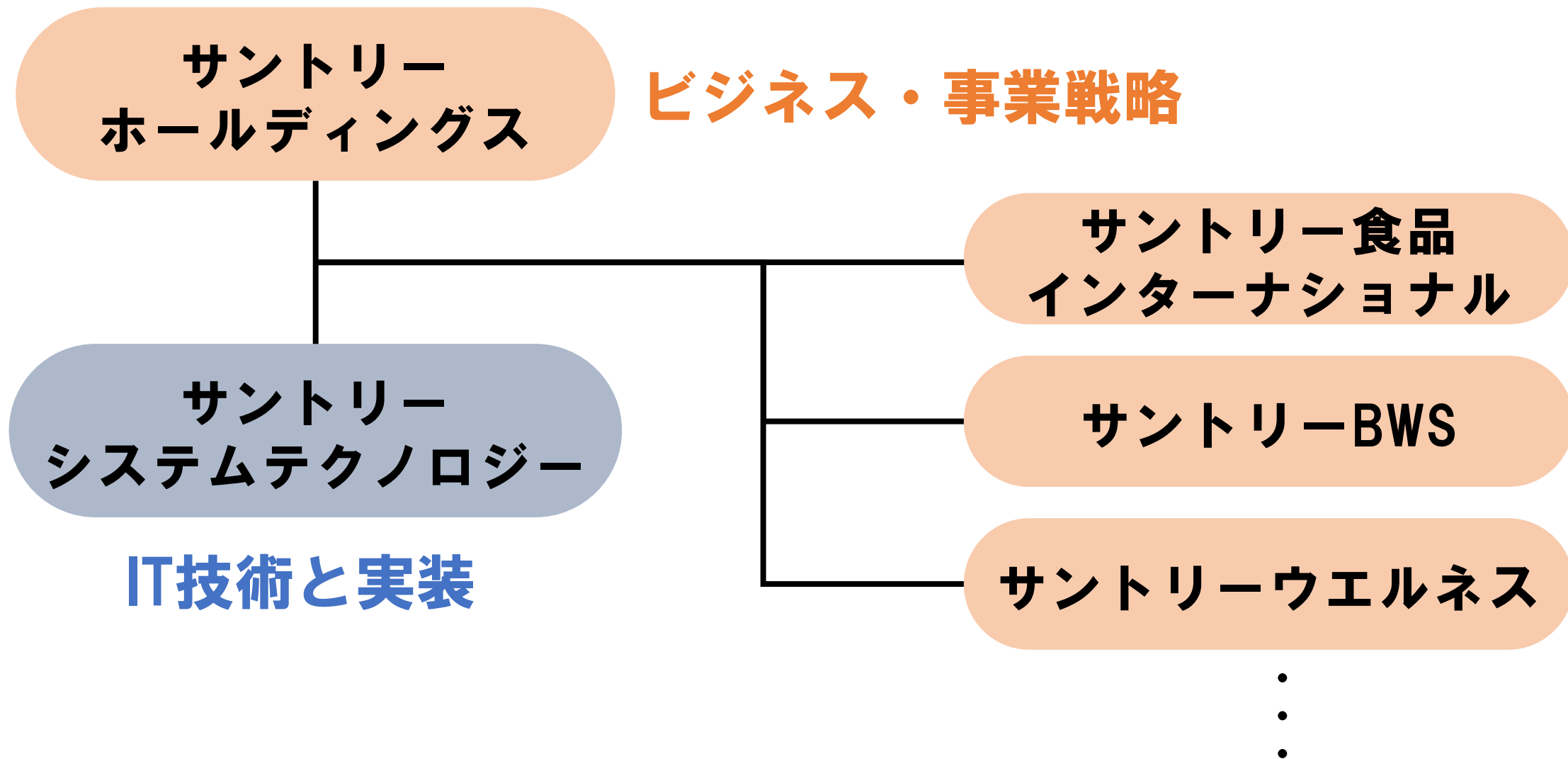
おいしいビールで糖質ゼロ！
ぜひお試しください！



誕生
うまさに驚く
糖質ゼロ*
*食品表示基準による

パーフェクトサントリー
ビールが飲めるお店

サントリーグループの体制



サントリーシステムテクノロジーのミッション

サントリー システムテクノロジー

ITプロフェッショナル企業として、新技術に挑戦し、
新たな価値を創造することを通じ、サントリーグループの成長、
並びに豊かな社会の発展に貢献する

グローバルインフラ部

システム品質部

ビジネスソリューション部

先端技術部

・
・
・

✓ 業務アプリの運用開発

✓ AI、IoT、xRの業務適用

✓ 新規事業開発

✓ デジタル人材育成

などを担当

本日本話する内容

1 会社紹介

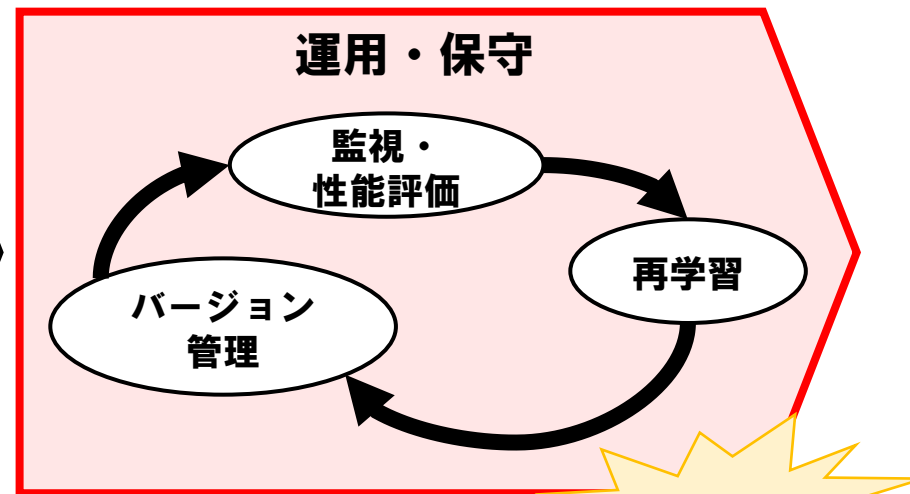
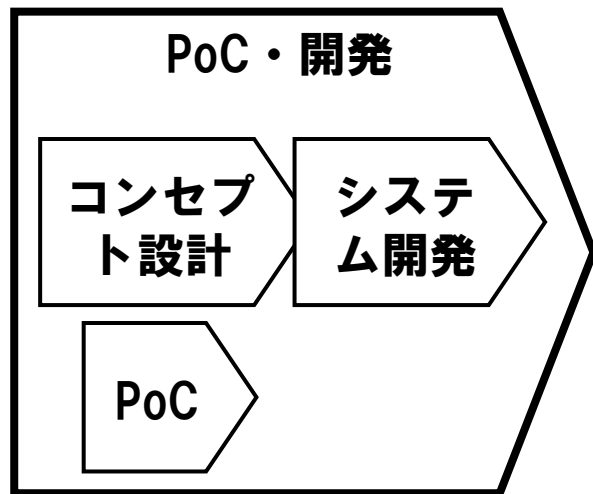
2 **機械学習モデルの共通運用基盤**

3 物流分野におけるモデル運用

機械学習モデルの運用・保守における課題

機械学習モデルは、リリース直後から性能監視業務がはじまり、「再学習による精度向上」が必要となる。

機械学習モデルを持つシステムの開発開発～運用までのプロセス



AIシステム運用・保守の課題

- AI精度の向上
- データ傾向変化の検知
- AIモデルの更新と本番化
- ビジネス変化によるAIモデルの変更

課題

Citadel AIで実現できること

機械学習モデルを組み込んだシステムの運用・保守に必要な機能が網羅されており、既存システムに組み込むことができる。

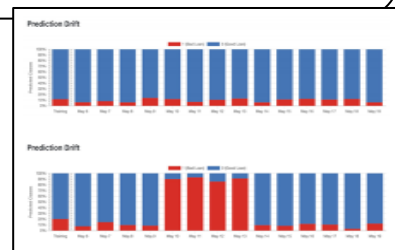
運用・保守

監視・
性能評価

バージョン
管理

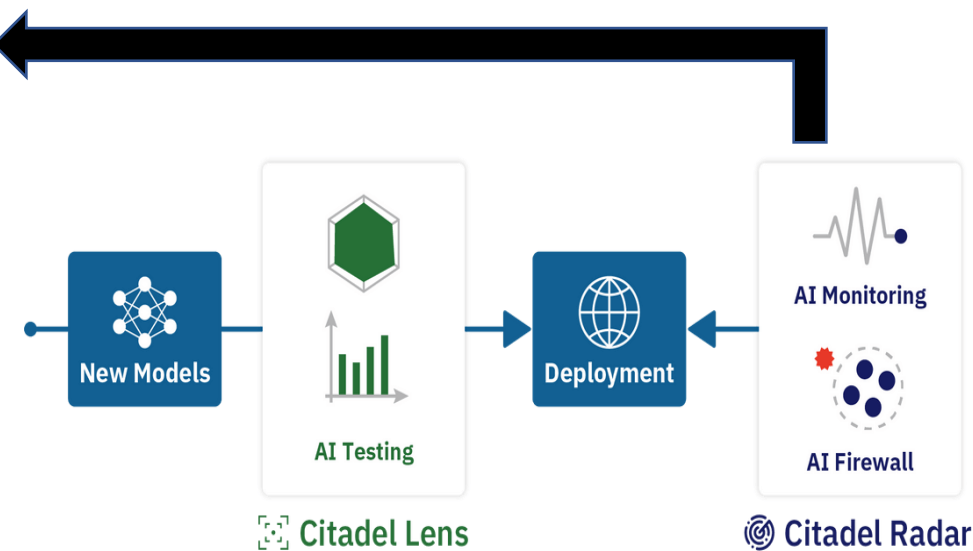
再学習

[予測結果のモニタリング]
予測結果の変化を監視し、
再学習のタイミングを把握



[モデルのバージョン管理]
過去すべてのモデルと比較して、
精度が改善⇒AI賢くなっている
ことを確認

[学習データの分布可視化]
学習データを分析し、性能劣化に
繋がる学習を防止



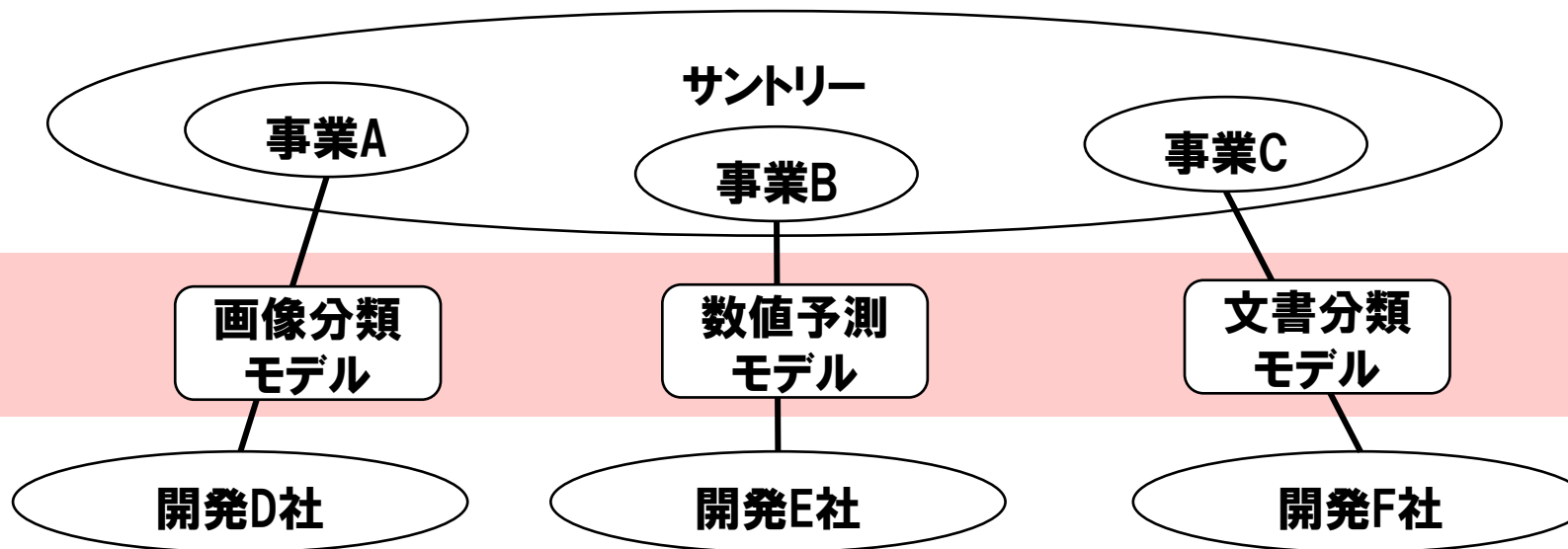
Citadel
サービス構成

サントリーにおける活用方針

AIベンダー各社が作ったモデルを横串で管理して
モデル性能を引き出していく

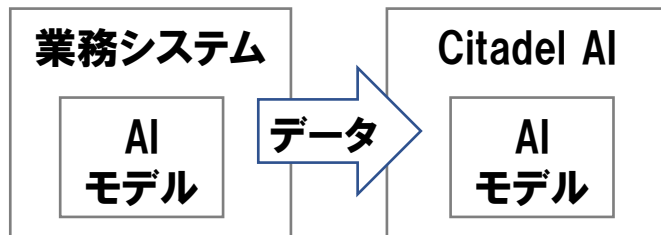


横串で監視

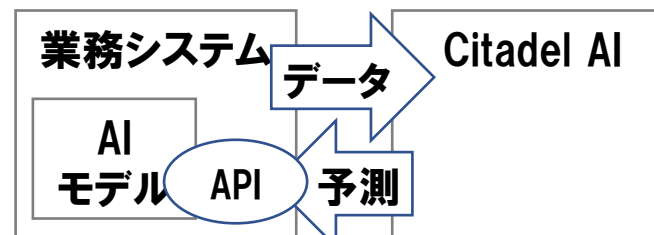


Citadel AIの
組み込み方式

Citadel Lens

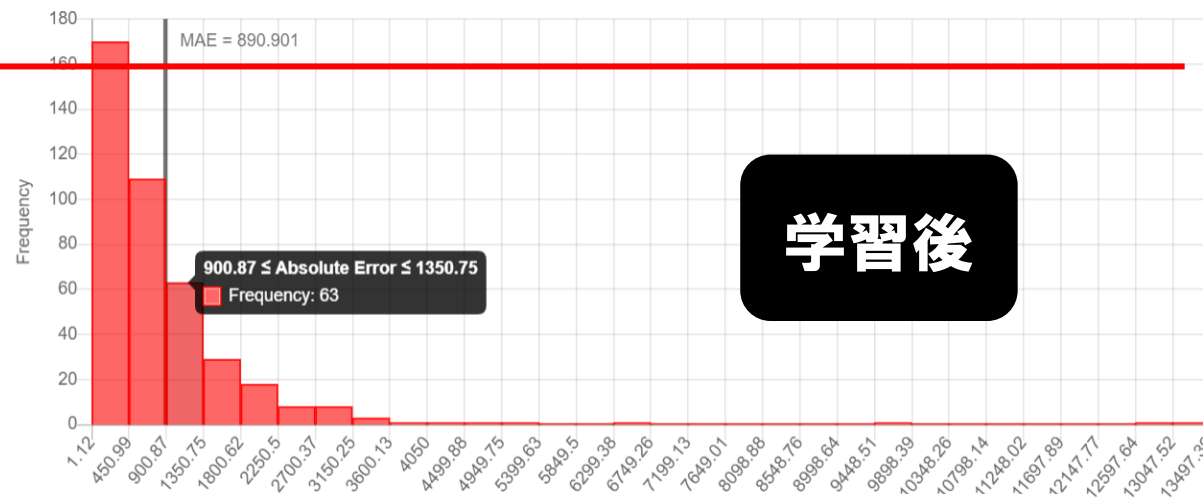


Citadel Radar



Citadel Lens

学習前後のモデル性能を詳細に比較することができる



小 予測値と実測値の誤差 大

改善

悪化

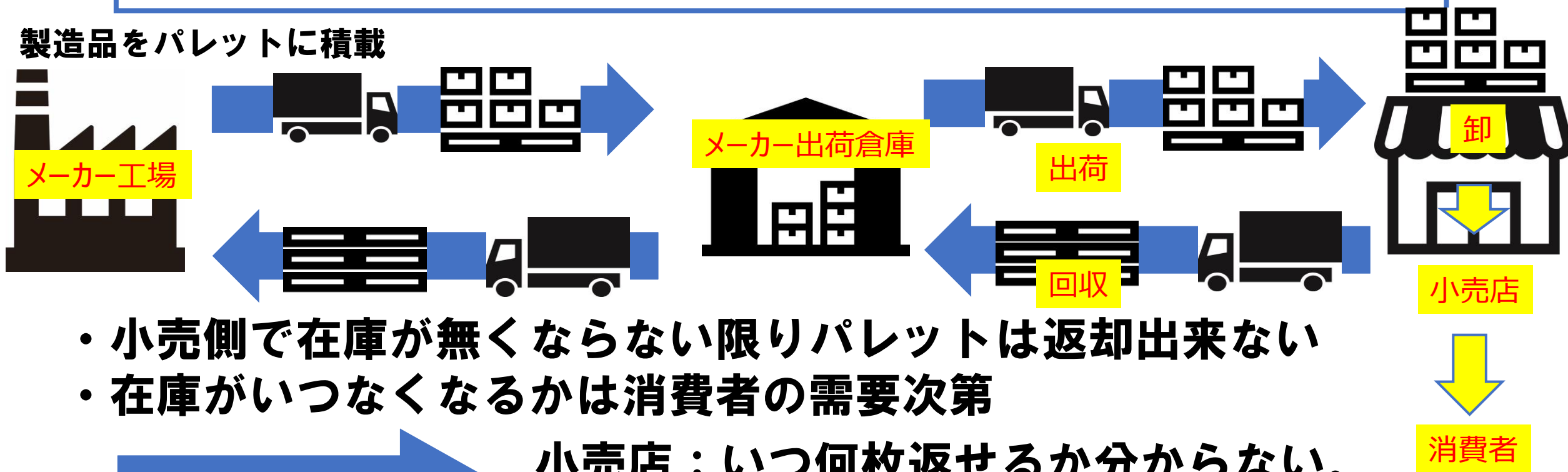
Less Error?	Prediction Change	Delta	Truth	Row Index	
✓	4520.811 ⇒ 8553.571	+4032.760	11596.000	Row #237	[2022050]
✓	11681.881 ⇒ 14214.124	+2532.243	15733.000	Row #187	[2022042]
✗	15661.964 ⇒ 17492.302	+1830.339	16020.000	Row #243	[2022050]

本日本話する内容

- 1 会社紹介
- 2 機械学習モデルの共通運用基盤
- 3 物流分野におけるモデル運用

物流機材需給業務の背景

パレットとは製品の輸送に用いる機材。出荷後に回収を行い、再利用している。



つまり

小売店：いつ何枚返せるか分からない。
メーカー：いつ何枚返ってくるかわからない。

回収予測を見誤るとパレット不足で工場が稼働できず甚大な被害に。

回収予測について

- **従来：物流実務チームでの予測**
 - ・現場の長年の感覚による予測
 - ⇒そのため、属人化・精度に課題があり、対面のアプリチームに相談
- **昨年：アプリ運用チームで予測モデルを開発**
 - ・アプリ運用チームで開発、業務適用可能な精度のモデルが完成
 - ⇒しかし、**運用に工数がかかっている状況**

弊社の基本方針

予測モデルの運用には貴重なデータサイエンティストの工数を要するため、運用業務をできるだけ自動化・効率化したい。

運用課題

- 予測結果を実績と照らし合わせるまで精度劣化に気づけない
- 精度劣化の原因が分からない
- 原因が分かってもモデルを再構築が困難・工数がかかる



Citadel AIを活用し、
精度劣化を検知 & 原因を可視化

Data Drift

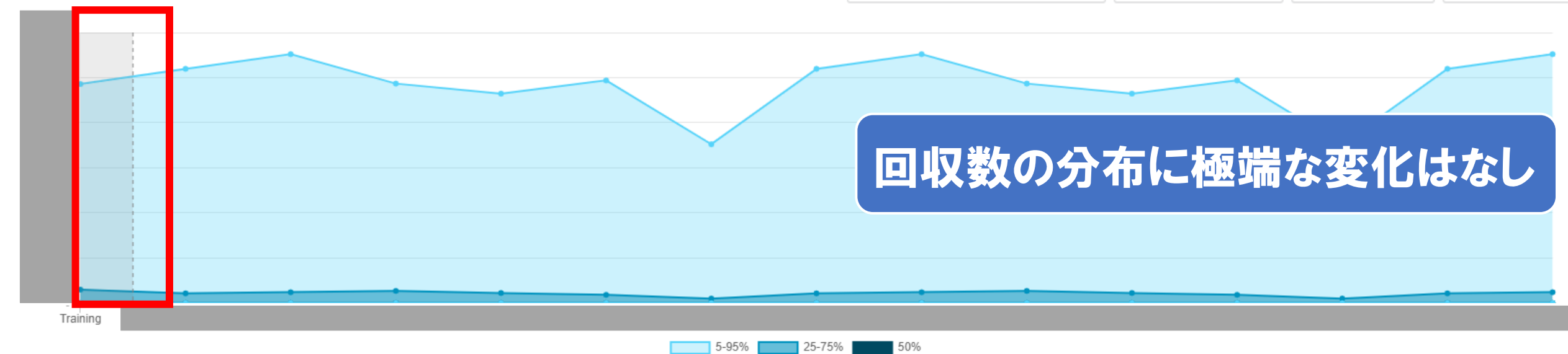
学習データと運用データで各変数の分布を比較、
分布の傾向変化を監視する

Status	Feature	Type	Drift ⓘ	Threshold
Skew	変数名	Float	0.84	0.7
Skew		Float	0.75	0.7
Healthy		Float	0.58	0.7
Healthy		Int (Numeric)	0.58	0.7

➡ ビジネス変化により、
学習時点とデータ傾向が変化

➡ 通年通りの傾向

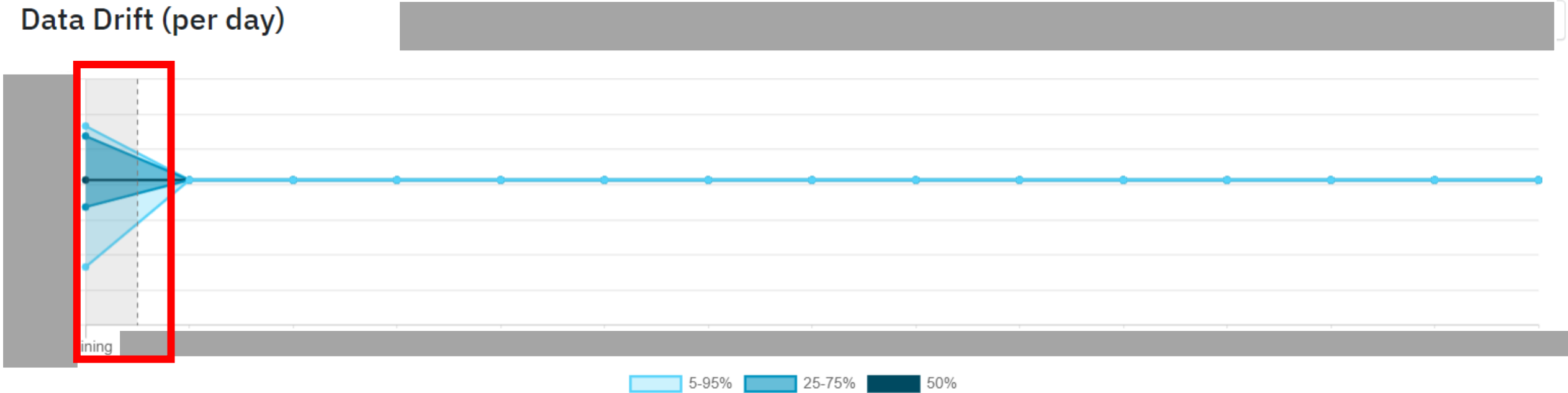
回収数の曜日平均



エリア（エリアごとのデータ数（=倉庫数）の分布）



とあるジャンルの月間販売計画



学習データは2年分、運用データは1か月分
変数：月次で変化するデータ（同月は一定の値）

時系列データで、一定期間同じ値を取る変数や季節による変動が大きい変数では、Data Driftでうまく精度劣化を見ることが出来ない。

Data Gaps

学習データと運用データの出現割合のGapを検知

Data Segment	Gap ②	Serving Freq.	Training Freq.
年月 ∈ []	99.96%	99.96%	0%
変数名 ∈ [1.2, 1.5]	99.96%	99.96%	0%
日付 ∈ []	99.96%	99.96%	0%

③運用データの99.9%は
学習データに存在していない値

②運用データの
1.2~1.5の値の頻度は99.96%

①学習データの
1.2~1.5の値の頻度は0%

そのため、学習データに含まれていないデータが大量に出てきているので精度が劣化する可能性あり

※発生原因

とあるジャンルの商品が前例にない販売傾向を示しているため。

今後の活動

運用

①モデルの精度劣化を検知

②精度劣化の原因とビジネス背景を分析

③モデルの再構築

①②③を誰でもできるように標準化
(ルール・手順の整理)を実施する