

SKグループ様向け

マナビDX Quest 2023 地域企業協働プログラム 最終報告会 2024/2/9

PatchWorks



経産省主催のDX人材育成プログラム。下記2つのプログラムを通して、企業におけるDXを推進する変革の考え方やプロセスを学び、志を同じくする幅広いデジタル人材とのつながりを構築することを目的とする。

ケーススタディ教育プログラム(PBL)

- 講師による座学ではなく、参加者が情報交換して学び合い・教え合いながら、与えられた課題を解決していく**PBL (Project-based Learning:プロジェクト型学習)**を中心に据えたプログラムです。

企業の課題特定
要求/要件定義

課題解決に向けたソリューションの検討/検証

経営者/他部署に向けた
実装・組織変革の提案

地域企業協働プログラム

- 地域の中小企業の課題に取組み、**経営陣・担当者との協働作業を実施**。この経験を通じて、ケーススタディ教育プログラムで学んだ、DXを現場で推進する際の難しさやポイントを学ぶことができます。

キックオフMTG

企業とチームとの間で、2カ月で取り組む課題/ゴールを設定

プロジェクト実施

課題/ゴール達成に向けた検討の実施

最終報告

検討結果を企業に最終報告

● 主催

経済産業省
独立行政法人 情報処理推進機構 (IPA)

ITパスポート
基本情報技術者試験

● 運営

Boston Consulting Group (BCG)

● コンテンツ提供

SIGNATE

コンペティション開催

● 協力

日本ディープラーニング協会 (JDLA)

G検定
E資格



マナビDXクエスト(MDXQ)の地域企業協働プログラムにおいて、Phase1~4として**現状把握・課題設定・実現性検証・実行計画策定**を実施しましたので、報告内容をご確認ください。

また、MDXQ終了後もチームとのプロジェクト継続をご検討頂きたく、ご判断材料として、MDXQ範囲外の内容について**今後の展望**及び**プロジェクト継続ご提案**を付録しています。



ぐるっとポンとRPSデータの分析から御社課題を様々抽出し、以下タスク(A)(B)について実現性検証(PoC)実施

(A)リサイクルステーションの充填率予測

(B)古紙投入量予測によるランクアップ条件最適化(ぐるっとポンアプリ)

(A)は [REDACTED]。 (B)は [REDACTED]。また(A)(B)以外にも、PoC未実施の施策案を多く本資料中に提示。

MDXQ終了後も有償にてプロジェクト継続頂くことで、チーム側で(A)の実装を進められる。また各種施策案の深堀や実現性検証を進め、**新たなDX実装の可能性を追求**することも可能。

リサイクルステーションの充填率予測

目的

リサイクルステーションの充填率を予測することで古紙の回収量増・回収効率向上に寄与する

予測充填率の活用方法

(1)空ステーションへユーザ誘導

- 予想充填率を地図と連携してアプリに表示し、ユーザを誘導
- ポイントやガチャチケットで誘導

(2)充填率に応じて古紙回収頻度変更

- 回収推奨日を算出し、回収業者にメール自動送信

PoC結果



5年ROI(投資対効果)

Return
Investment

(1)空ステーションへユーザ誘導

- API提供のみ：
[REDACTED]
- アプリ組み込み：
[REDACTED]

(2)充填率に応じて古紙回収頻度変更

- 人件費削減効果含まず：
[REDACTED]
- 人件費削減効果含む：
[REDACTED]

MDXQ後のプロジェクト継続

有償契約によるプロジェクト継続を是非ご検討ください。

プランA (実装含む)

PoCにより実現性が高いと判断されるタスクを対象として、要件定義・設計・開発のすべてをチーム側で実施。完成後の保守・更新も実施

プランB (PoCまで)

- MDXQの延長線上で、データ可視化・分析、PoC
- 今回の内容に限らず、SK様側と自由に議論し、事業貢献する施策を検討
- PoC結果次第でプランAへの移行を検討

今後の施策案(一部抜粋)

- AIに依らない最適なランクシステム構築
- キャンペーン効果測定
- チャットボット構築
- GoogleアナリティクスへのユーザID紐づけ
- アプリ画面にユーザー名と達成内容を通知

Agenda

1. プロジェクトの背景・目的
2. MDXQの全体像、ゴールの設定
3. Phase 1: 現状把握・問題整理
4. Phase 2: 課題の設定・打ち手検討
5. リサイクルステーションの充填率予測
 1. Phase 3: 実現性検証(PoC)
 2. Phase 4: 実行計画策定
6. 古紙投入量予測によるランクアップ条件最適化
 1. Phase 3: 実現性検証(PoC)
 2. Phase 4: 実行計画策定
7. 今後の展望
8. MDXQ後の継続ご提案
9. 振り返り

01

プロジェクトの背景・目的



ぐるっとポンという自社アプリは多くのユーザーを獲得し、データも蓄積されている。
しかしデータ利活用が進んでおらず、アイデアが欲しい。

- **相談パターン**

データ分析を通じたデータ・デジタル技術の活用可能性の設計/初期的な検証

- **プロジェクト名**

地域未来牽引企業におけるデータドリブンな新規事業の加速・拡大支援

- **相談内容**

リブランディング案件。すでにいる2万件のアプリユーザー、総会員9万人に対し、これまでの活動履歴などから特定の顧客がより利用するための方法をデータをもとに仮説を立て、PDCAを回し検証する。

- **企業からのアピールポイント**

既存の価値観にとらわれない自由な発想で議論させていただき、収集したデータを活かしスピード感を持って実現していきたいと思います。忸度ない議論と実現可能性の見極めを行い、実現に向け柔軟なアイデアを持って取り組んでいただけることを期待しています。

- **受講生に求めること**

既存の価値観にとらわれない自由な発想で議論させていただき、収集したデータを活かしスピード感を持って実現していきたいと思います。忸度ない議論と実現可能性の見極めを行い、実現に向け柔軟なアイデアを持って取り組んでいただけることを期待しています。

02

MDXQの全体像、 ゴールの設定



ゴール案：“ぐるっとポン”のデータの利活用によって、RSの利用率向上、ユーザー当たりの古紙回収量増加を目指す。

スケジュール案：MDXQプログラム内で実現性の検証・将来の実行計画策定まで完了させる。開発はスコープ外。

GOAL

古紙回収事業のKSF(仮想)

- 古紙を“安価”に“大量”に回収すること



RPS“ぐるっとポン”の役割

- RS(Recycle Station)の利用率(市場規模に対するカバー率)向上
- ユーザー当たりの古紙回収量増加(離脱回避・リピート)



本プロジェクトのゴール

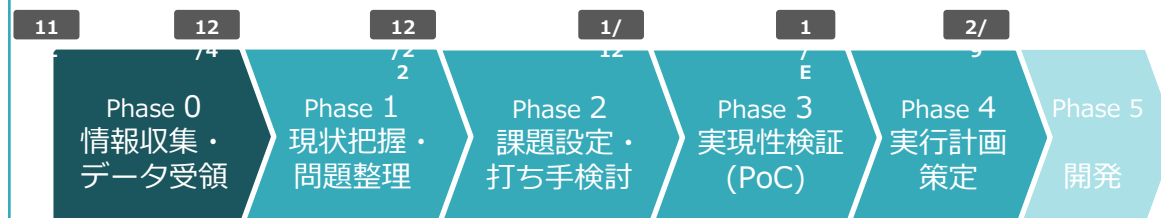
- データ利活用によってぐるっとポンの役割を強化する

SCHEDULE

当初スケジュール



実際のスケジュール



03

Phase 1: 現狀把握・問題整理





ヒアリングを通して、RPS・ぐるっとポンの未活用データを使った利用者増や回収古紙増の施策の必要性を認識。一方取得したデータの探索的分析により新たな知見を得られる可能性を認識。

ヒアリング内容の分析 (Issue-driven)

- RPS・ぐるっとポンデータの活用
- RS・ぐるっとポン利用者増
- RSへの古紙投入量・回収量増

探索的データ分析 (Data-driven)

- ぐるっとポンアプリ・施策の効果の評価
- ぐるっとポンと古紙回収の相関・因果
- ランクシステムの妥当性検証
- 顧客の属性やカテゴリー分析

①問題意識を御社からオープンにヒアリングしつつ、並行して②EA分析、③ビジネス環境分析などチーム側で能動的に調査・分析も進めることを目指した。②については、システム構成の確認に変更

現状把握・
問題整理課題の設定・
打ち手の検討データ/デジタル技術の
利活用による実現性検証

開発

ビジネス・組織の情報を把握し、問題を整理する

①御社の問題意識ヒアリング

- PatchWorks側主担当
: chef、naoya

PatchWorks側でAgendaを特定せず、御社の課題意識をオープンにヒアリングさせていただきます。



②エンタープライズアーキテクチャ分析

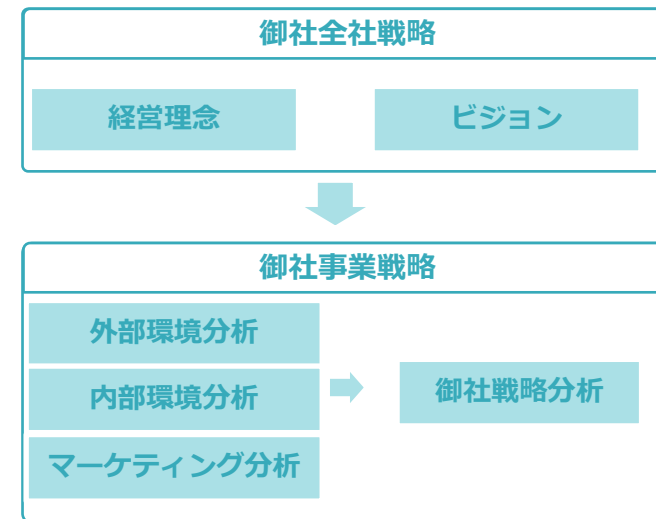
- PatchWorks側主担当
: Saito、naoya



- BA : 経営資源や業務プロセスなど
- DA : 企業が扱うデータの関連性や構造
- AA : 業務で使用するアプリの機能や構成
- TA : 企業が基盤として利用する技術全体

③ビジネス環境分析

- PatchWorks側主担当
: まっち、norosen、Tomi



ぐるっとポン及びRPSのアプリデータを使って、古紙回収量を増やしたい

RS(リサイクルステーション)の古紙回収量を増やすには

アプリデータを使って古紙回収量を増やすには、①RS利用離脱者を減らす、②古紙投入量を増やす、③古紙回収量を増やす、などの方法が考えられる。

古紙回収量

=

RS利用者数

新規を増やす
離脱者を減らす

x

1人当りの古紙投入量

投入量を増やす

古紙回収量

=

RS数

RSを増やす

x

店舗当りの古紙投入量

回収量を増やす

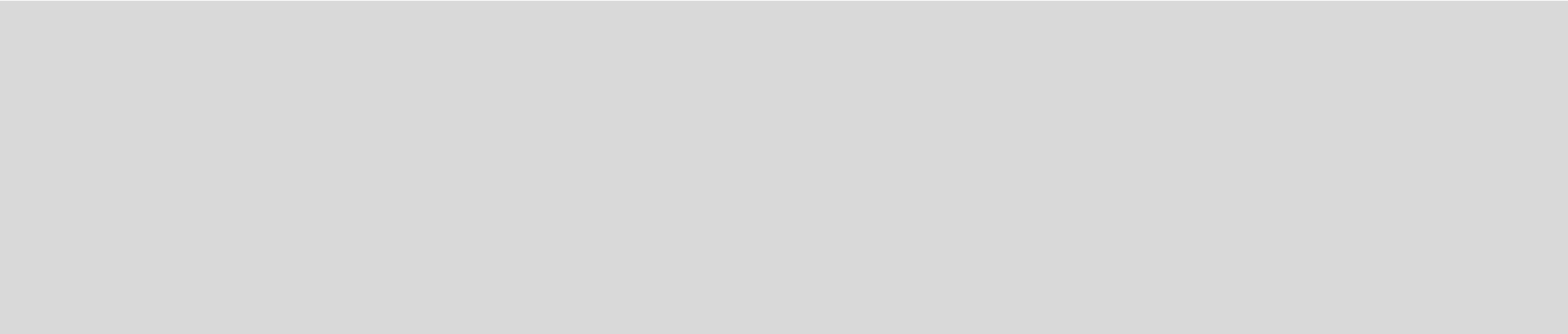
ぐるっとポンのビジネスモデル

:

:

:

:



KSF(業界で成功するための肝)は、古紙を大量に効率的に入手すること。

外部環境分析

● PESTLE

特定業界・年代のマクロ外部環境評価

古紙の国内売価は古紙の供給過剰で低位安定。
輸出価格頼みの収益構造。政治経済・法律の
影響は小さい。

P (政治): 中国の古紙禁輸による輸出価格下落
E (経済): 景気の影響を受けにくい
S (社会): ペーパーレス化による紙製品の需要減
T (技術): RPFなど製紙原料以外への古紙利用
L (法律): リサイクル関連法令に大きな変動なし
E (環境): SDGsの浸透

内部環境分析

● VC(Value Chain)分析

付加価値を出せるのは調達フェーズ
選別・売却での差別化は困難

調達

選別

売却

● 5F分析

業界内部の競争の強さと、業界の収益性を評価。横軸は利益を確保できる構造か否か、縦軸は利益を他者から搾取されうるかを示唆する。

横軸: 買い手市場のため、
影響力の小さい売り手
(回収業者・店舗・ユー
ザー)のエリアで利益を
上げる施策が必要になる。

縦軸: 新規参入・競争は**ほぼない**ため、基本的に**自治体の回収**と古紙の奪い
合いになる。

新規参入の脅威

- ・ 新規参入はほぼ無い
- ・ 設置コストの面で障壁あり

小

売り手の交渉力

- ・ 月2回の自治体の古紙回収
に出すか、
リサイクルステーション
に出すかの**2択**

小

業界内での競争

- ・ 既存シェア内での競争は自治体の回収
 - 自治体はコスト面からやめたい
委託先はやめたくない
- ・ エリア拡大時に競争可能性あり

中

買い手の交渉力

- ・ **買い手市場**
- ・ 製紙会社の工場が主な取引
先年間の古紙買い取り枠を
決めてもらい、その中で交
渉する

大

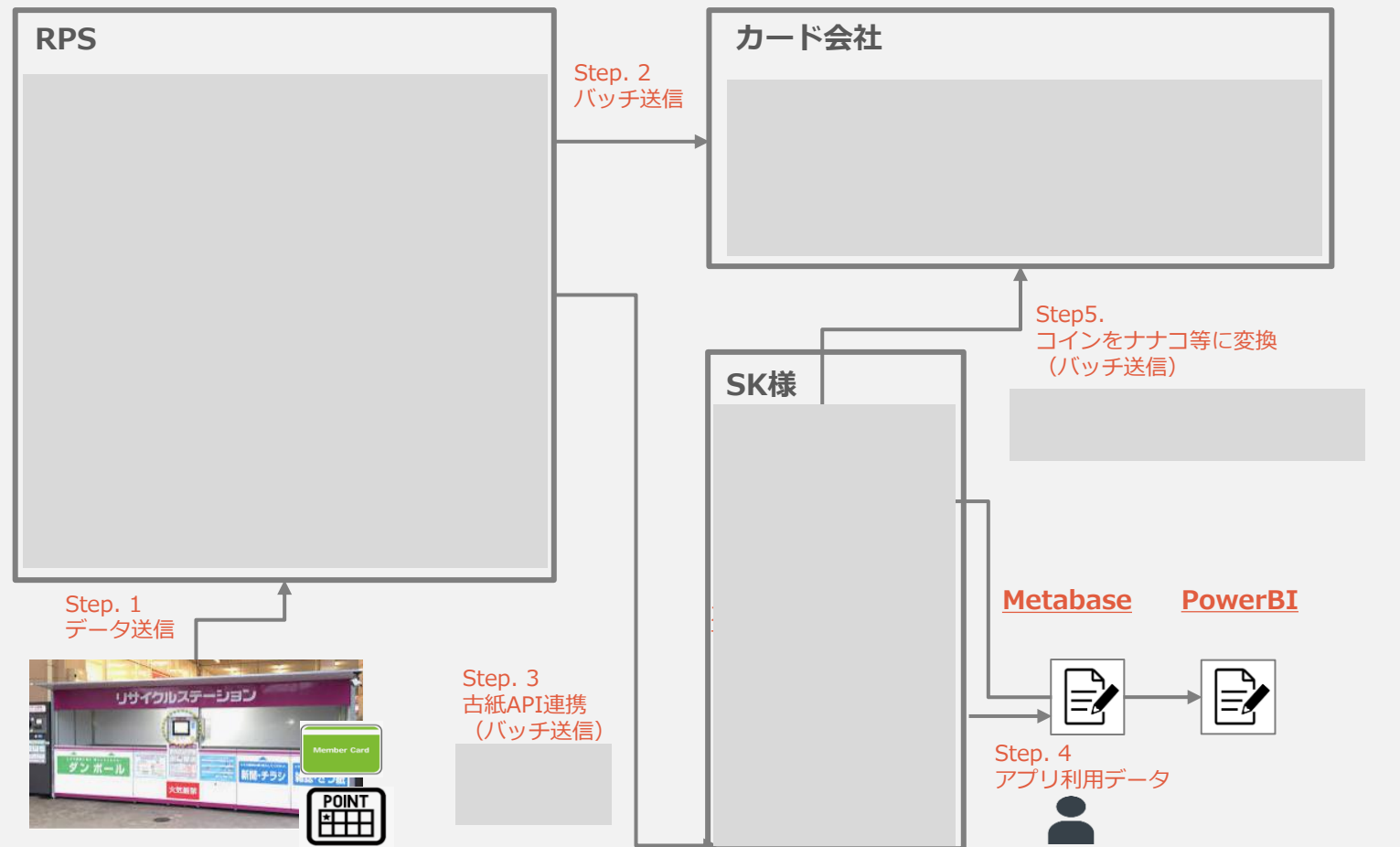
小

代替手段の脅威

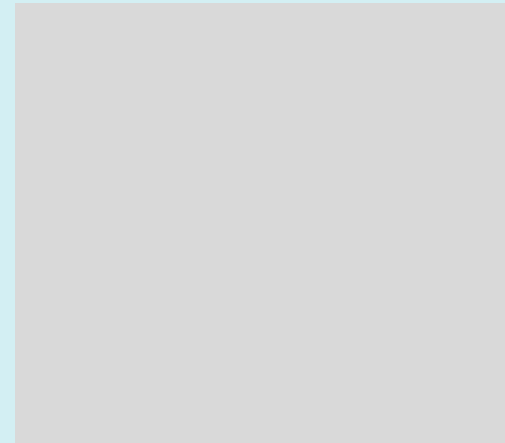
- ・ 自治体の回収 (仙台は火曜日・月2回)
- ・ 燃えるゴミに混ぜて捨てる (よくない)

RPSデータとしては [] のデータを取得可能。データとしては [] などが取得可能。
ぐるっとポンのデータとしては [] が取得可能。
ともにMetabaseでデータ分析し、BI化している。

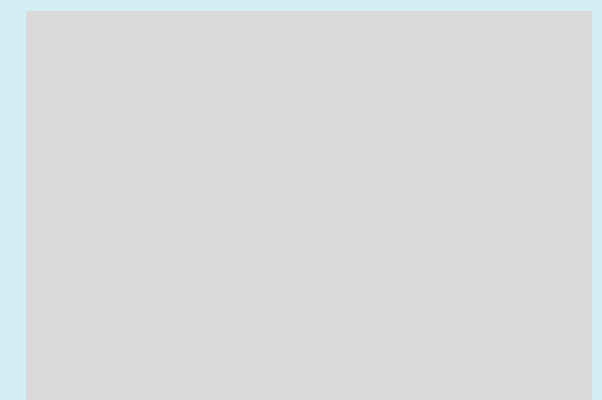
システム構成図



ぐるっとポンアプリデータ



RPSデータ



04

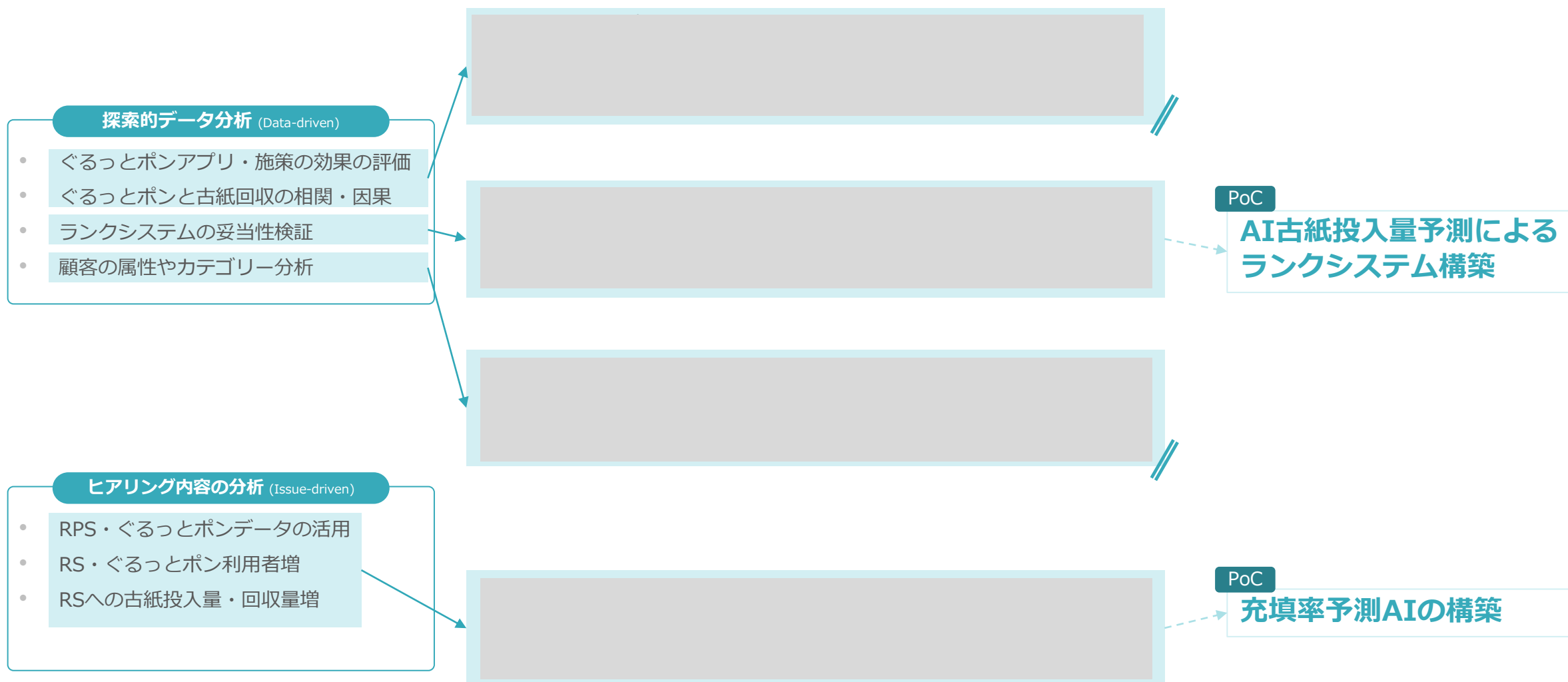
Phase 2: 課題の設定・打ち手検討





次フェーズPoC

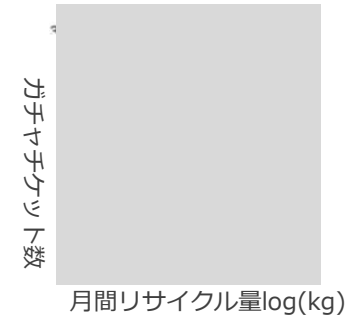
では、ランクシステム再構築による古紙回収促進、充填率予測による古紙回収効率の向上を目標とする。



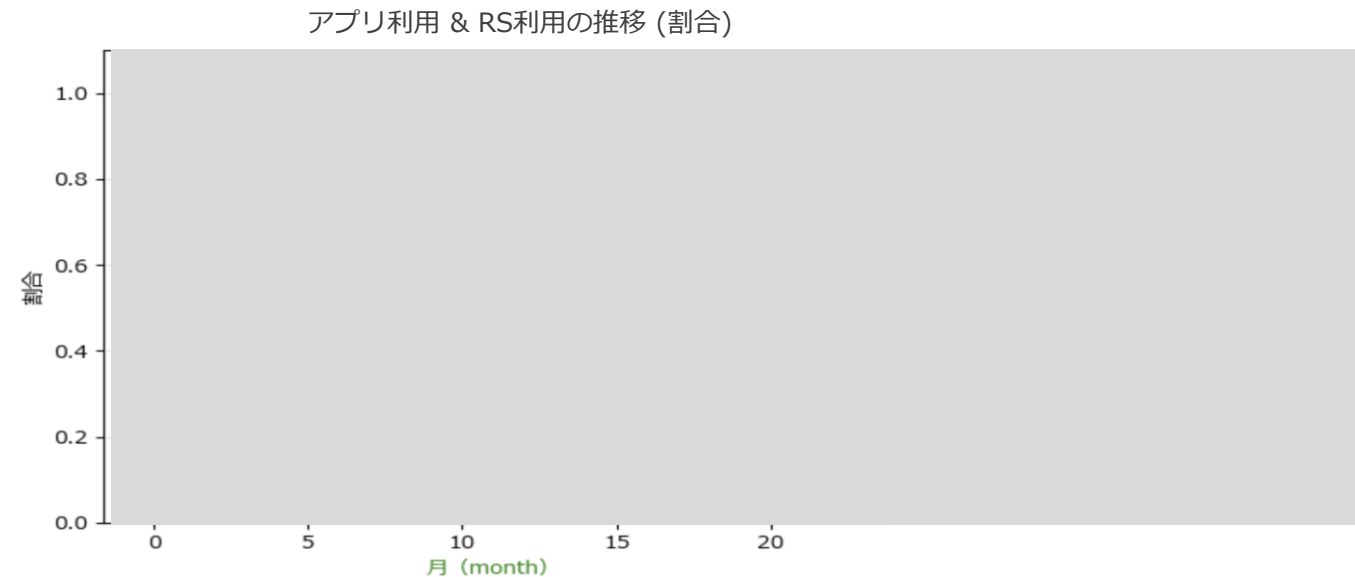
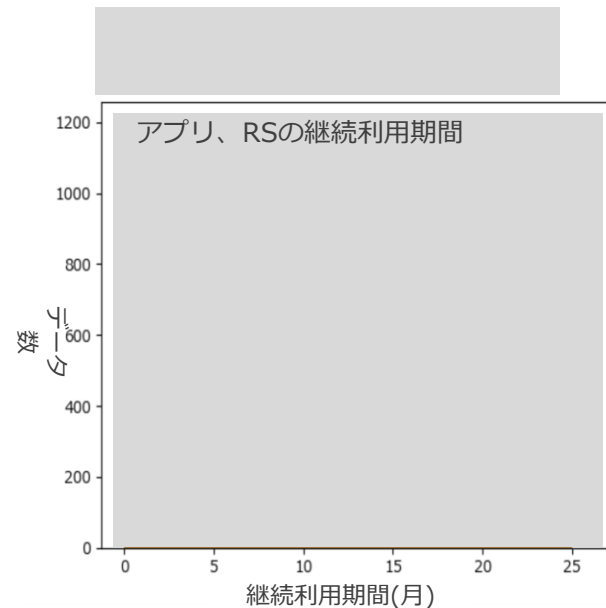
探索的データ分析 (Data-driven)

- ぐるっとポンアプリ・施策の効果の評価
- ぐるっとポンと古紙回収の相関・因果
- ランクシステムの妥当性検証
- 顧客の属性やカテゴリー分析

ぐるっとポンアプリと古紙投入量の相関



ぐるっとポンアプリの継続使用とRS継続使用の状況



探索的データ分析 (Data-driven)

- ぐるっとポンアプリ・施策の効果の評価
- ぐるっとポンと古紙回収の相関・因果
- ランクシステムの妥当性検証
- 顧客の属性やカテゴリー分析

k-means法による非階層的クラスタリング実施結果

今後の課題

ユーザ毎に古紙排出量を予測し、適切なランクアップ条件にする事でアプリ継続率UPを図る

探索的データ分析 (Data-driven)

- ぐるっとポンアプリ・施策の効果の評価
- ぐるっとポンと古紙回収の相関・因果
- ランクシステムの妥当性検証
- 顧客の属性やカテゴリー分析

AsIs :



現状のランクシステム

アクティブユーザのランク内訳
(直近3か月以内にアプリ利用履歴のあるユーザ)

ToBe : 適切なランクアップ条件の予測

- 各ユーザの**古紙排出限界を予測**
- ユーザー毎に**rank達成目標を設定
- アプリ継続率UPおよび古紙回収の増加に繋がる

過去の投入実績

次月ランクアップ条件



Aさん

投入量**少なめ**

8 kg



Bさん

投入量**多め**

15 kg

過去の投入実績に基づき、AIで予測

$$\text{古紙回収量} = \text{RS利用者数} \times \text{1人当りの古紙投入量}$$

新規を増やす

離脱者を減らす

投入量を増やす



RSの充填率を予測することで古紙の回収量増・回収効率向上に寄与する施策が検討可能

KSFを満たすために

各店舗のRS充填率を予測することで、古紙投入量・回収量の増加と、無駄な古紙回収の回避に寄与することができ、KSFの達成につながる。

$$\begin{aligned} \text{古紙回収量} &= \text{RS利用者数} \begin{cases} \text{新規を増やす} \\ \text{離脱者を減らす} \end{cases} \times \text{1人当りの古紙投入量} \begin{cases} \text{投入量を増やす} \end{cases} \\ \text{古紙回収量} &= \text{RS数} \begin{cases} \text{RSを増やす} \end{cases} \times \text{店舗当りの古紙投入量} \begin{cases} \text{回収量を増やす} \end{cases} \\ \text{回収効率} &= \text{古紙回収量} / (\text{回収工数} \text{ or } \text{回収費用}) \begin{cases} \text{無駄な回収を回避} \end{cases} \end{aligned}$$

ヒアリング内容の分析 (Issue-driven)

- RPS・ぐるっとポンデータの活用
- RS・ぐるっとポン利用者増
- RSへの古紙投入量・回収量増

充填率の判定

ぐるっとポン・RPSの施策

ユーザの誘導

- **ぐるっとポンで空きのあるリサイクルステーションにユーザを誘導する**
例) 予想充填率をアプリに表示する
ポイントやガチャチケットキャンペーンで誘導する

回収業者の誘導

AsIs

-
-
-

ToBe

- リサイクルステーションの充填率に応じて**回収ルートを変更**する
- **追加回収をする／回収業者にインセンティブを付与する**

05

リサイクルステーションの 充填率予測



RSの充填率を予測することで古紙の回収量増・回収効率向上に寄与する施策が検討可能

KSFを満たすために

各店舗のRS充填率を予測することで、古紙投入量・回収量の増加と、無駄な古紙回収の回避に寄与することができ、KSFの達成につながる。

$$\begin{aligned} \text{古紙回収量} &= \text{RS利用者数} \begin{cases} \text{新規を増やす} \\ \text{離脱者を減らす} \end{cases} \times \text{1人当りの古紙投入量} \begin{cases} \text{投入量を増やす} \end{cases} \\ \text{古紙回収量} &= \text{RS数} \begin{cases} \text{RSを増やす} \end{cases} \times \text{店舗当りの古紙投入量} \begin{cases} \text{回収量を増やす} \end{cases} \\ \text{回収効率} &= \text{古紙回収量} \div (\text{回収工数} \text{ or } \text{回収費用}) \begin{cases} \text{無駄な回収を回避} \end{cases} \end{aligned}$$

ヒアリング内容の分析 (Issue-driven)

- RPS・ぐるっとポンデータの活用
- RS・ぐるっとポン利用者増
- RSへの古紙投入量・回収量増

充填率の判定

ぐるっとポン・RPSの施策

ユーザの誘導

- **ぐるっとポンで空きのあるリサイクルステーションにユーザを誘導する**
例) 予想充填率をアプリに表示する
ポイントやガチャチケットキャンペーンで誘導する

回収業者の誘導

AsIs

-
-
-

ToBe

- リサイクルステーションの充填率に応じて**回収ルートを変更**する
- **追加回収をする／回収業者にインセンティブを付与する**

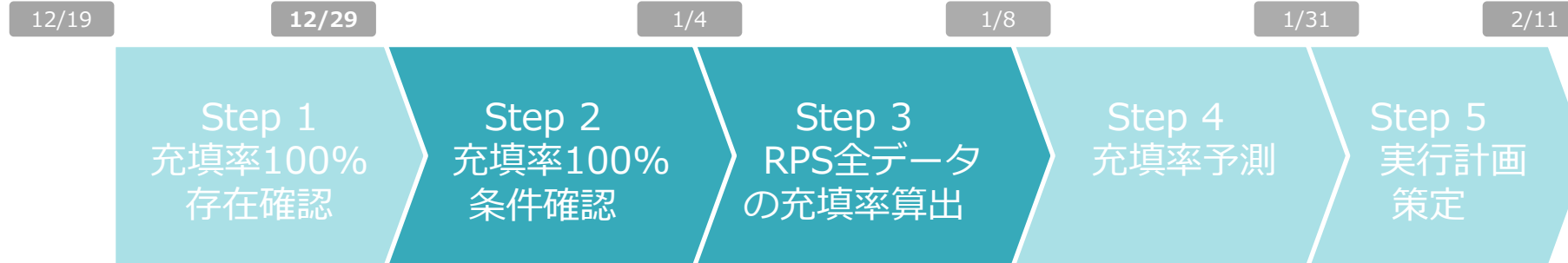


各店舗においてRPSの前回からの利用間隔を算出

-
-

➤ 充填率100%が存在

ヨークベニマル 仙台愛子店

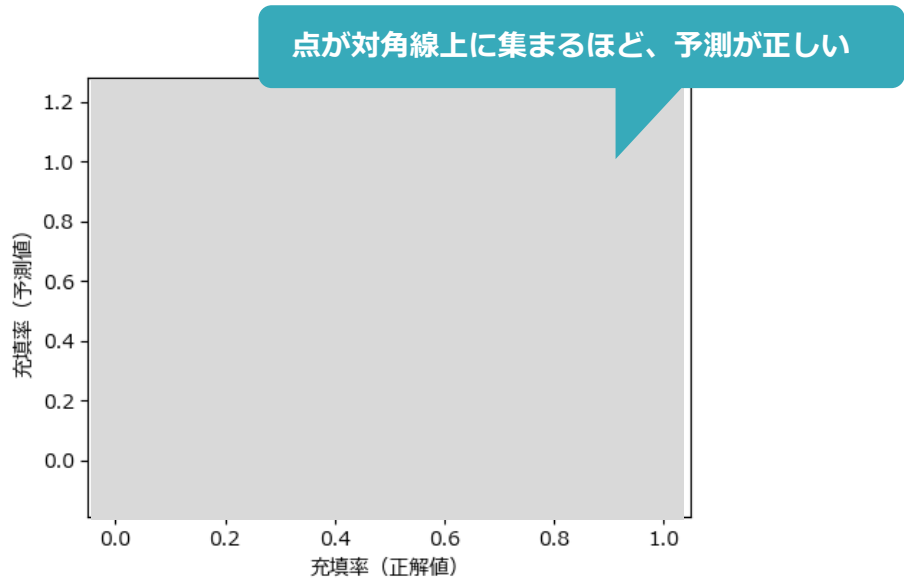


いずれかの条件を満たす場合、1日の最終利用時が充填率100%とした

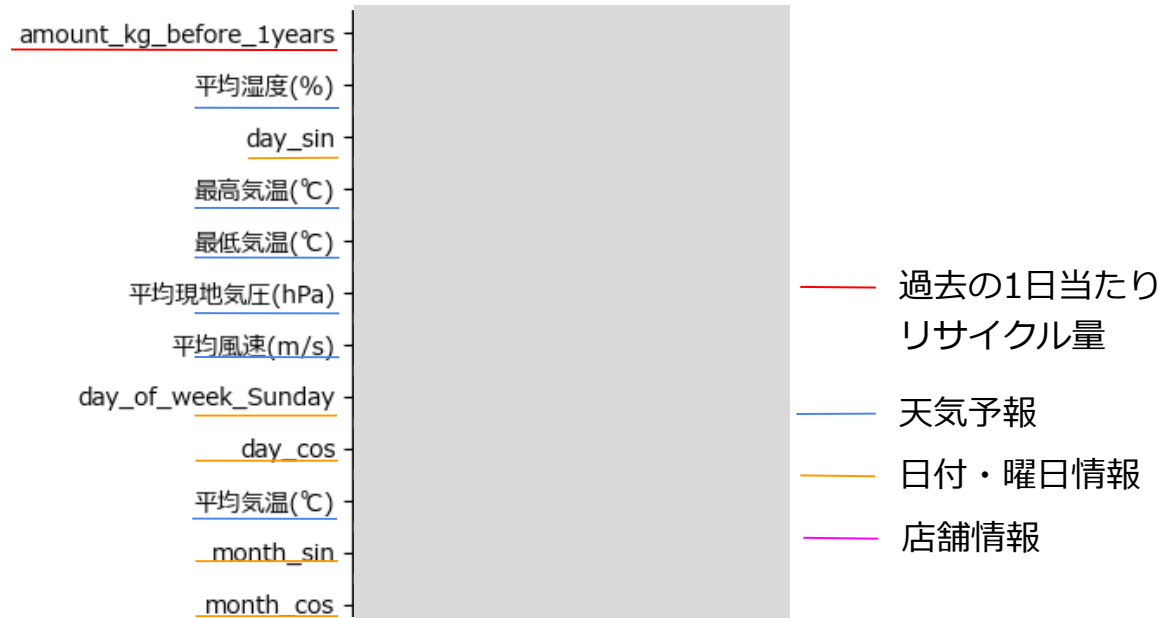
ヨークベニマル 仙台愛子店



- 5 日後までの充填率を予測するAIを構築
- RMSE =



正解値 vs 予測値グラフ



学習データの重要度グラフ

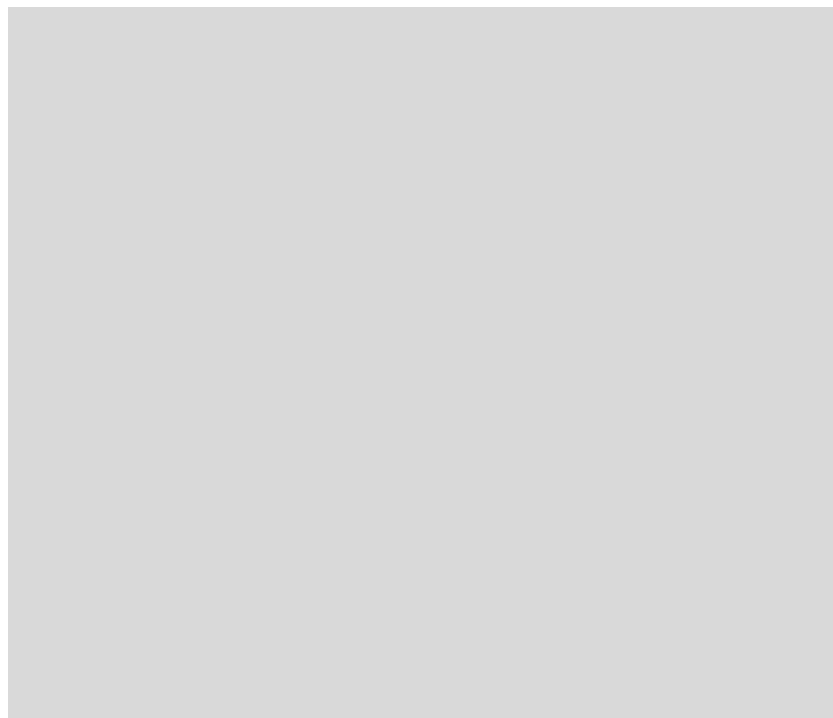
実施期間	方法 1 (空ステーションへユーザ誘導)		方法 2 (充填率に応じて回収数変更)																									
4か月	<div>充填率改善・AIモデル改善</div> <div>2回収収日を網羅するため、2回収収日を予測するAIモデルを構築</div> <div>充填率の正当性を確認するため、一部店舗で充填率を計測（例：カメラを配置し監視）</div> <div>AIモデルを改善し、精度を向上</div>																											
3～6か月	<div>ぐるっとポン実装</div> <div><ul style="list-style-type: none">予想充填率を地図と連携してアプリに表示し、ユーザを誘導ポイントやガチャチケットキャンペーンで誘導する機能を追加</div> <div></div>	<div>回収推奨日提示</div> <div><ul style="list-style-type: none">回収推奨日を算出し、メールを自動で送信するシステムを構築</div>	<div>回収予定表 (〇〇様、〇日週)</div> <table><tr><th>日時</th><th>〇〇店</th><th>〇〇店</th><th>〇〇店</th></tr><tr><td>〇/〇</td><td>2回</td><td>1回</td><td>1回</td></tr><tr><td>〇/〇</td><td>1回</td><td>0回</td><td>0回</td></tr><tr><td>〇/〇</td><td>2回</td><td>1回</td><td>1回</td></tr><tr><td>〇/〇</td><td>2回</td><td>1回</td><td>1回</td></tr><tr><td>〇/〇</td><td>2回</td><td>1回</td><td>1回</td></tr></table>		日時	〇〇店	〇〇店	〇〇店	〇/〇	2回	1回	1回	〇/〇	1回	0回	0回	〇/〇	2回	1回	1回	〇/〇	2回	1回	1回	〇/〇	2回	1回	1回
日時	〇〇店	〇〇店	〇〇店																									
〇/〇	2回	1回	1回																									
〇/〇	1回	0回	0回																									
〇/〇	2回	1回	1回																									
〇/〇	2回	1回	1回																									
〇/〇	2回	1回	1回																									

単位：万円

方法1 (空ステーションへユーザ誘導)

現ぐるっとポユーザー(約1割)が満杯時に別店舗を利用した場合、 の追加利益
5年後はぐるっとポユーザーの増加見込みが

月平均の古紙回収利益



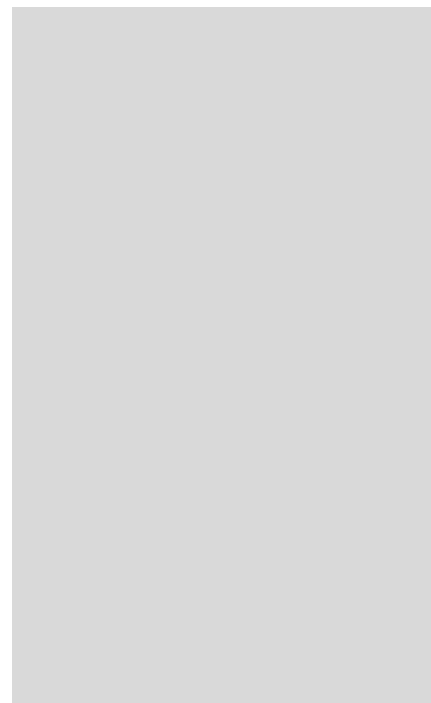
提案前

提案後

方法2 (充填率に応じて回収数変更)

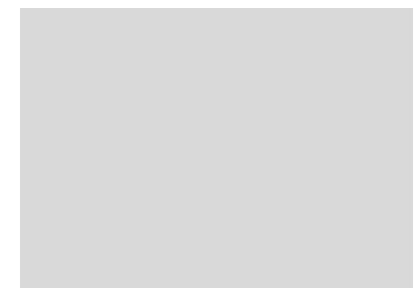
全ユーザが常時古紙を投入できる場合、回収コストの増加を考慮すると、 の追加利益
回収業者が回収頻度を効率よく減らした場合、 のコスト削減

月平均の古紙回収利益



提案前

提案後

月平均の古紙回収コスト
(2023年,12月除く)

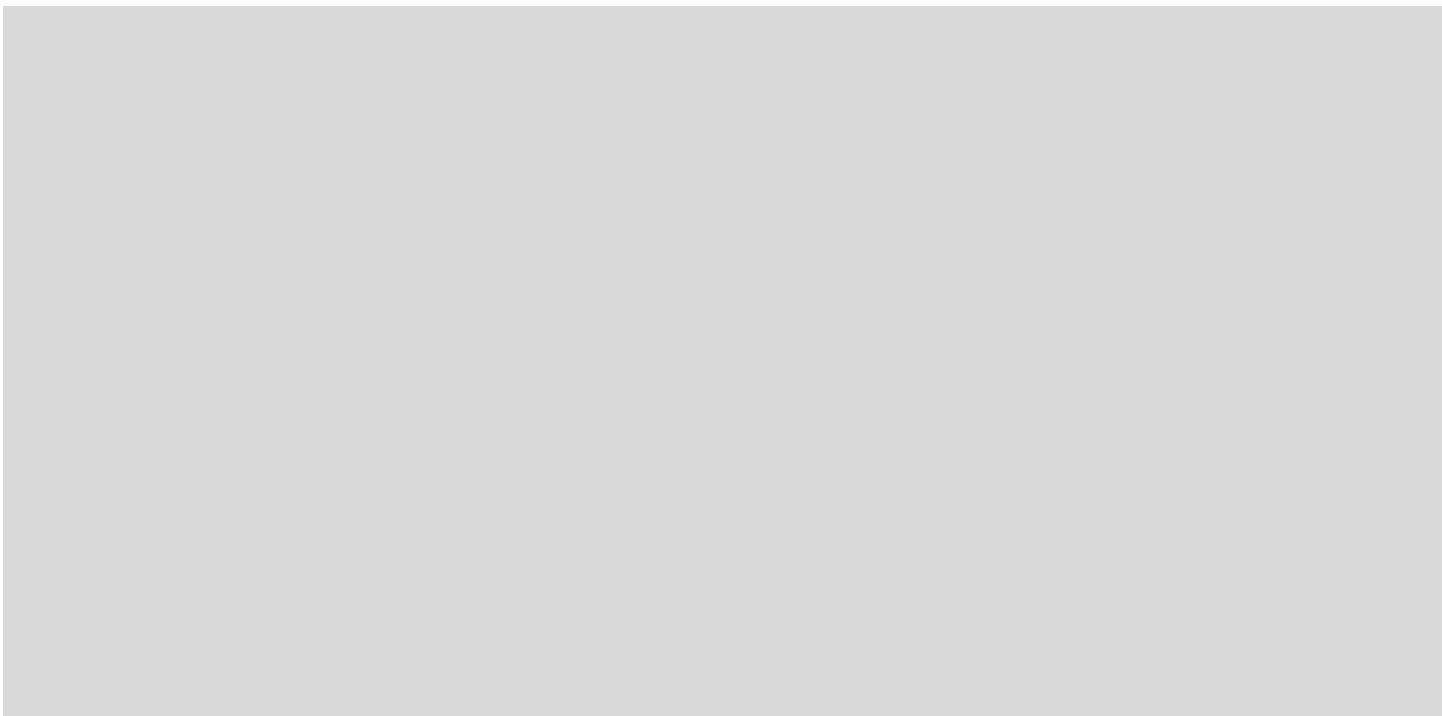
提案前

提案後

	方法 1（空ステーションへユーザ誘導）		方法 2 （充填率に応じて回収数変更）
	API提供のみ	アプリ組み込み	
PoC継続費			
システム構築費			
システム運用費			
※AWS利用時の概算			
保守費			
合計投資額（5年後）			
システム導入による 利益（5年後）			
ROI（5年後）			

今後の展望（充填率の再算出）

2回収日は充填率の予測に大きく影響するが、すべての2回収日をデータとして把握できていない
→把握できている1回収、2回収日から、把握できていない日の回収回数を予測する



06

古紙投入量予測による ランクアップ条件最適化





探索的データ分析 (Data-driven)

- ぐるっとポンアプリ・施策の効果の評価
- ぐるっとポンと古紙回収の相関・因果
- ランクシステムの妥当性検証
- 顧客の属性やカテゴリー分析

ユーザ毎に古紙排出量を予測し、適切なランクアップ条件にする事でアプリ継続率UPを図る

AsIs : ランクアップ基準が高すぎる



現状のランクシステム

アクティブユーザのランク内訳
(直近3か月以内でアプリ利用履歴のあるユーザ)

ToBe : 適切なランクアップ条件の予測

- 各ユーザの**古紙排出限界を予測**
- ユーザー毎に**rank達成目標を設定
- アプリ継続率UPおよび古紙回収の増加に繋がる

過去の投入実績

次月ランクアップ条件



Aさん

投入量**少なめ**

8 kg



Bさん

投入量**多め**

15 kg

過去の投入実績に基づき、AIで予測

古紙回収量 = RS利用者数 x 1人当りの古紙投入量

新規を増やす

離脱者を減らす

投入量を増やす

PoC前に目指していたランクアップ条件の具体例は以下の通りである。メタベースの膨大なユーザ情報で学習したAIを用いて、古紙投入量可能性を予測してランクアップ条件に反映。

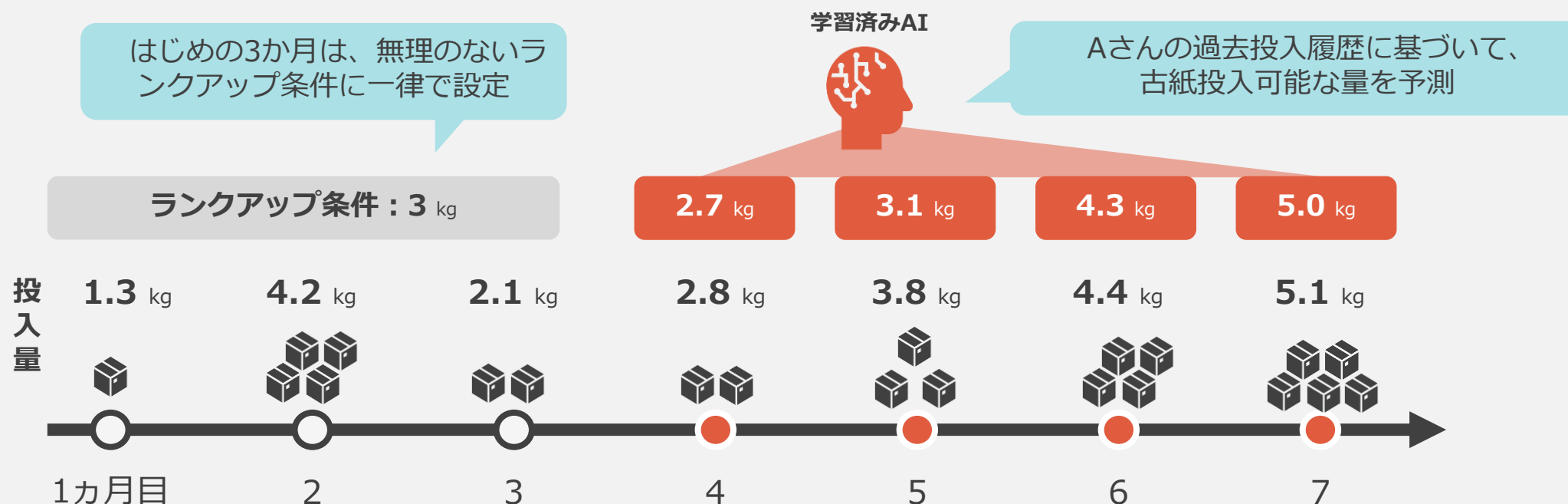
ペ
ル
ソ
ナ






- ・毎月約5 kgの古紙が家庭で発生
- ・RPSの他、自治体回収にも古紙を出している

➡ **ランクアップ条件を約5kgにして、すべてRPSに投入させたい**


運
用
例




受領した「.csv」と「.csv ※」を、AIの学習に適した形式にデータ整理した。
その結果、今回対象となるユーザの数はとなった。



不正なデータを削除




RPS利用開始日が2022-12-31以前のユーザに限定



1ヵ月で20kg以上投入のユーザは削除



・ 2023/1～2023/11投入量がすべて0kg
・ 2023/11が投入量が0kg } を削除



AIの学習に用いるデータと、予測するデータは以下の通りである。**2023/01から2023/10の古紙投入量データとユーザ固有情報を学習して、2023/11の古紙投入量を予測**する。

学習するデータ	備考
2023/01～2023/10 各月の古紙投入量	—
ユーザ居住地の緯度	—
〃 経度	—
総古紙投入量	—
月平均古紙投入量	—
月平均古紙投入回数	—
平均ランク	ブロンズ0、プラチナ4とした ときの平均ランク値
サービス利用開始からの 経過日数	2023/10/31 - RPS開始日
RPS継続利用月数	RPS最終利用日 - RPS開始日
ランク維持月数	ランクを何ヵ月維持しているか

学習に用いるデータの詳細

古紙投入量データ

※一部省略して表示

ユーザ固有情報

予測するデータ

実際に扱うデータの例

NaN(古紙投入記録なし)は家庭に古紙が無かったのではなく、**別の手段(自治体など)で廃棄**してしまったと仮定。単に0で補うのではなく、平均値で補った。しかし**データ数が少ない場合**に限っては、実態と乖離した値を補ってしまっている可能性あり。

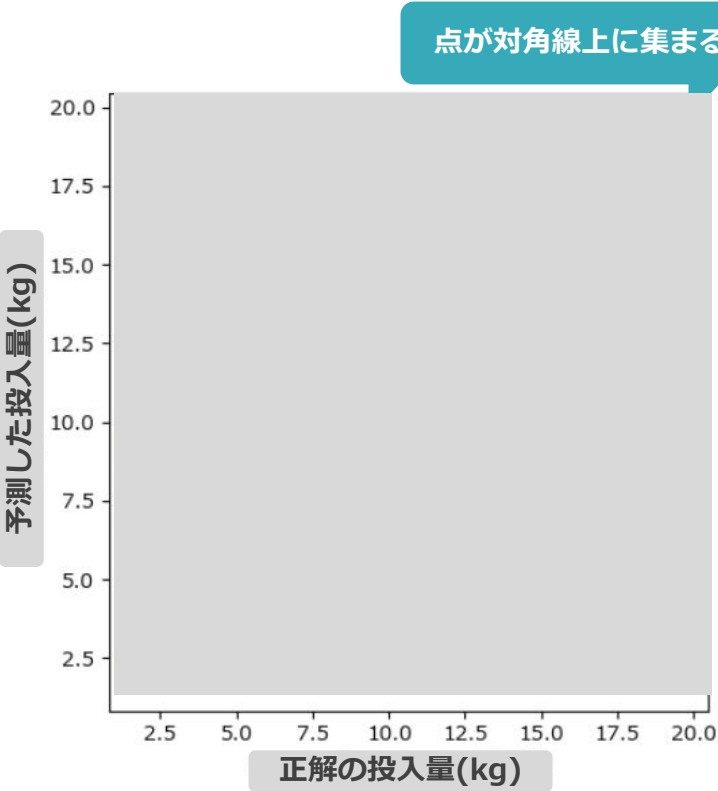
データ処理前

データ処理後

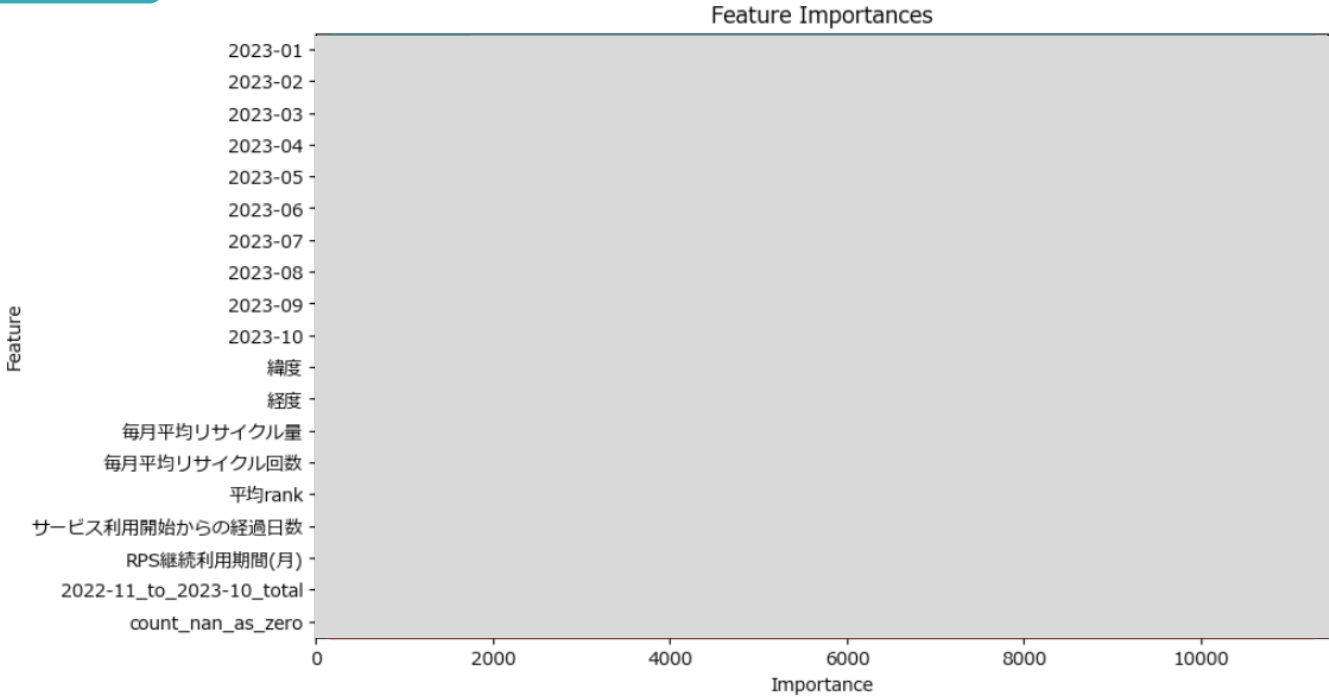
AIのモデルには、高速で効率的な**LightGBM**を使用。（詳細はAppendixを参照）

RMSE(平方根平均絶対誤差)※は

※ AIの精度評価指標の一つ。詳細はAppendix参照



正解値 vs 予測値グラフ

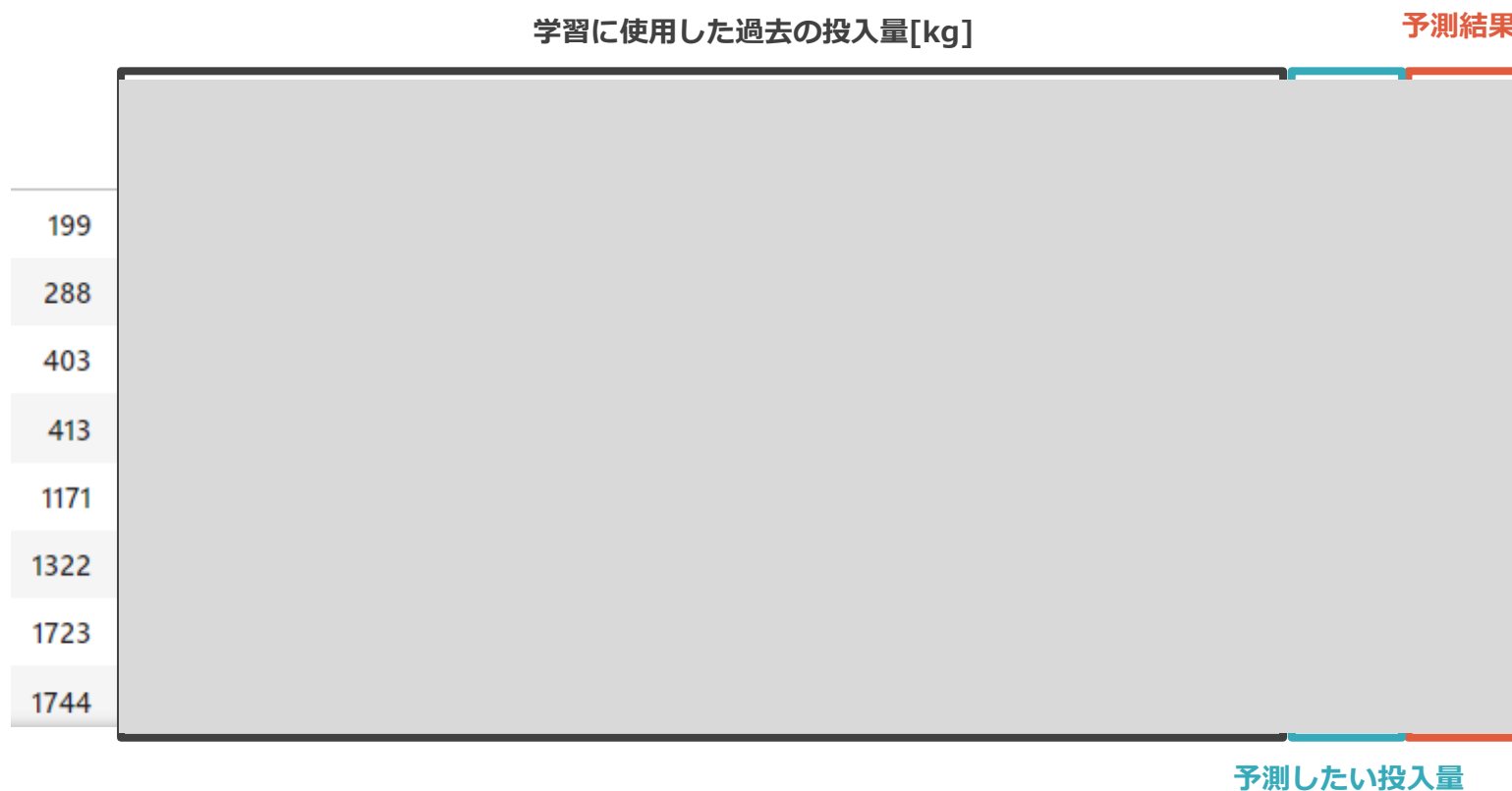


学習データの重要度グラフ

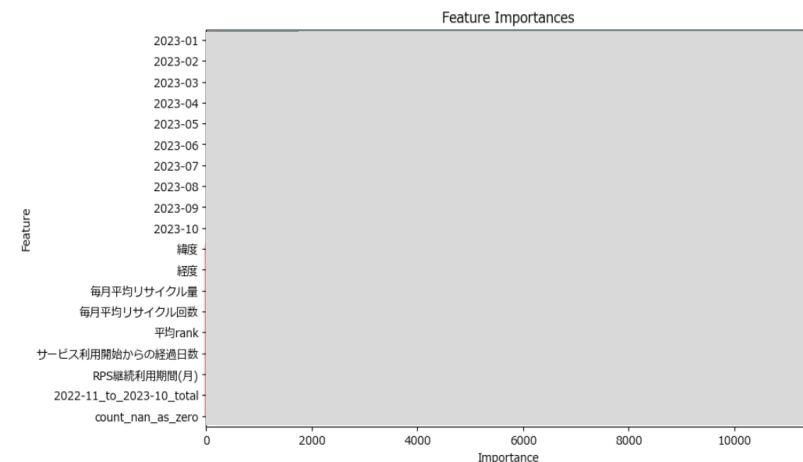
学習データと見比べてみると、前月までの傾向とは違う、突発的な投入量の変化に対応することができていない。

これは、月別投入量の重要度が高く、ユーザ特有情報の重要度が低いことに起因する。

→**重要度の高いユーザ固有情報**を新たに追加することで、精度向上の可能性あり。



学習データと予測結果の一部



学習データの重要度グラフ

以下のようなユーザ固有情報を新たに加えることで、精度向上する可能性がある。メタベースにはない情報があるため、**アンケート等で別途収集**する必要がある。また充填率予測のように**グルーピング**することで、**投入量の突発的な変化を緩和**して、精度の良い予測ができる可能性あり。

追加する説明変数の候補

説明変数	備考
新聞購読の有無	有無で大きく増減
ECサイトの利用頻度	頻度に応じて増減
車の所持台数	先行研究より投入量と相関あり
家族構成・世帯数	人数に応じて増減
投入した古紙の種類	古紙の種類によってモデル変更
年収	年収が購買行動に影響の可能性
出身地	地元かどうかで古紙増減の可能性
スーパーの利用頻度	スーパーに行くついでに古紙投入
SNS使用の有無	相関の可能性

グルーピング例

- ・地域ごとに古紙投入量を予測することで、RSを新たに設置する地域の選定根拠に
- ・ etc...



もしSK様側で「こういうグループの古紙投入量を予測できればこういう使い方ができるな」などのアイデアがあれば、気兼ねなくお申し付けください

Agenda

1. プロジェクトの背景・目的
 2. MDXQの全体像、ゴールの設定
 3. Phase 1: 現状把握・問題整理
 4. Phase 2: 課題の設定・打ち手検討
 5. リサイクルステーションの充填率予測
 1. Phase 3: 実現性検証(PoC)
 2. Phase 4: 実行計画策定
 6. 古紙投入量予測によるランクアップ条件最適化
 1. Phase 3: 実現性検証(PoC)
 2. Phase 4: 実行計画策定
- ▶ 7. 今後の展望
8. MDXQ後の継続ご提案
 9. 振り返り

07

今後の展望



GoogleアナリティクスへのユーザーIDの紐付け導入の提案

ユーザー分類やリサイクル量の予測などのユーザー別の予測に役立つ特徴量が増え、より高度な分析ができます

1. アプリ使用頻度（ログイン率）
2. アプリ使用時間帯
3. 平均セッション時間
4. アプリ使用デバイスやOS
5. 地理的なアプリ利用状況
6. コンバージョン（達成）率：リサイクル回数 / セッション数
7. 各ページ閲覧数



上記のデータはユーザーID付きでcsvダウンロード可能になります。



GoogleアナリティクスへのユーザーIDの紐付け導入②

実装期間 : 6h x 1人（ページの基本部分は共通部品化されていることを想定）

ユーザーIDを紐付けてより詳細な追跡を行う旨の**プライバシーポリシーの更新が必要**となります。

→SK様作業想定

プライバシーポリシーには、GoogleアナリティクスでユーザーIDを紐付けて個人を特定できない範囲でデータ分析を行う旨を明記し、データの匿名化や擬似匿名化の方法を具体的に明記します。

※EUや、カリフォルニア州でもぐるっとポンが使える場合は現地のデータ保護法が適用されてしまうためプライバシーポリシーの変更は専門家のアドバイスが必要になってきます。





リリース後、アプリ内で全ユーザーに通知を行います。

変更の要点と変更されたプライバシーポリシーへのリンクを提供します。

GoogleアナリティクスへのユーザーIDの紐付け導入③

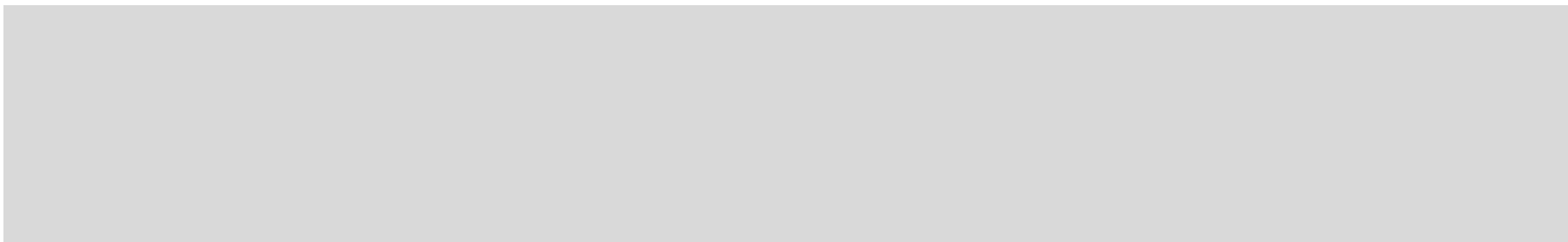
アナリティクス導入による属性追加と、報酬アンケートによる属性追加（まっち提案資料）からデータが溜まった1年後にはクラスタリングによるパーソナライズされたキャンペーンを実施する。

翌年から  人に対して毎年パーソナライズされたキャンペーンを実施すると仮定し期間中  %の古紙投入量アップの差分があるとする。

利益 :



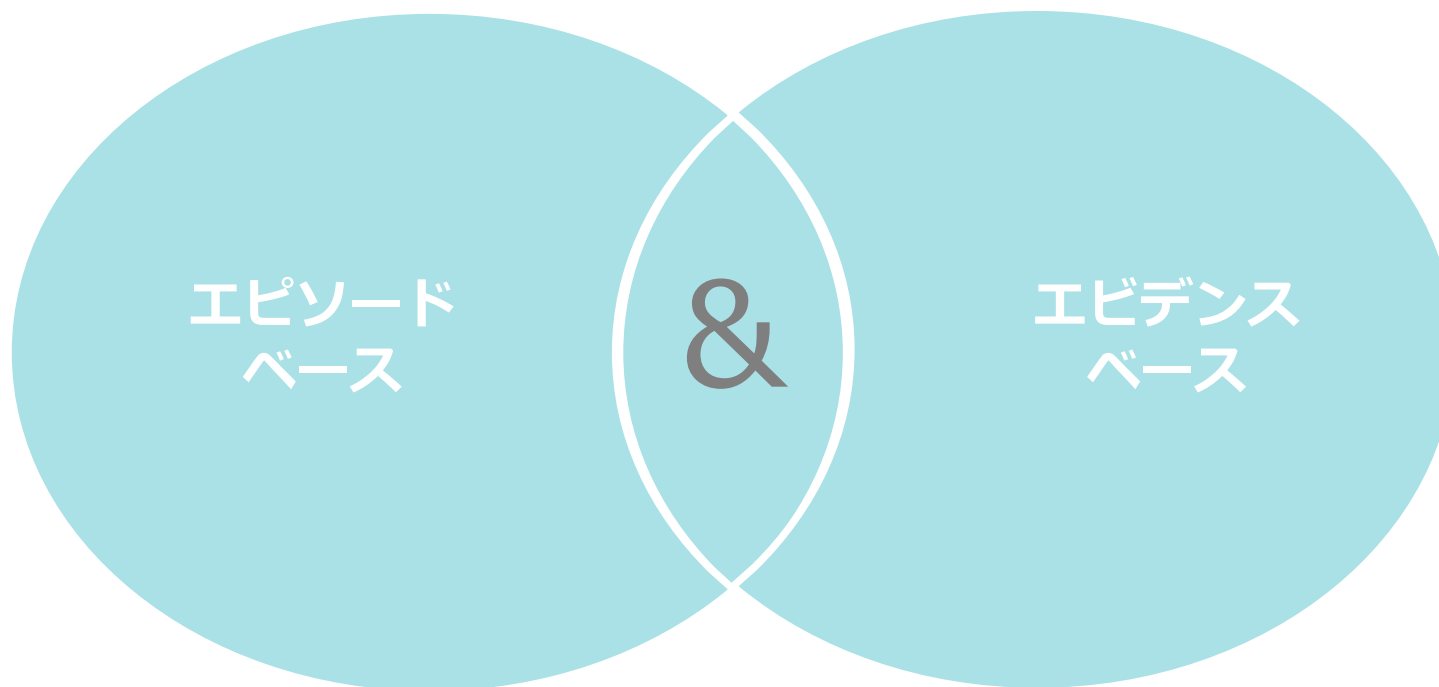
5年後の純利益 =



キャンペーンの効果測定に関するご提案

現状の課題：さまざまなキャンペーンや施策を行っているものの、定量的な効果検証ができていない（エピソードベース）
→最適なマーケティング戦略実施かつ迅速な意思決定につながっていない（エビデンスベース）

理想的な意思決定のプロセス



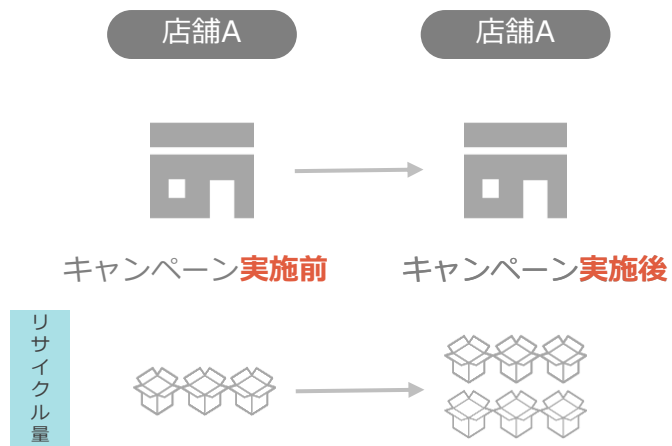


・キャンペーンの効果測定

日次バッチ（AWS LambdaとAmazon EventBridgeで実現でもOK）でキャンペーンの終了日があるかを確認し、終了したキャンペーンに対しての効果測定を行う。例えば統計検定（t検定・U検定・分散分析など）では仮説が統計的に正しいかを検証することができる。分析により効果的なキャンペーンを知ることは、**データドリブン経営を推進**する上で非常に重要だと考える。

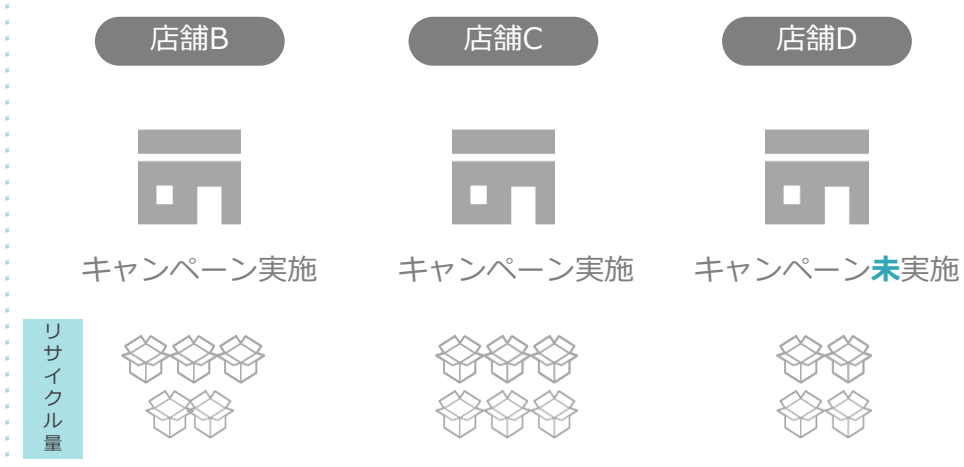
Case1

同店舗での効果検証

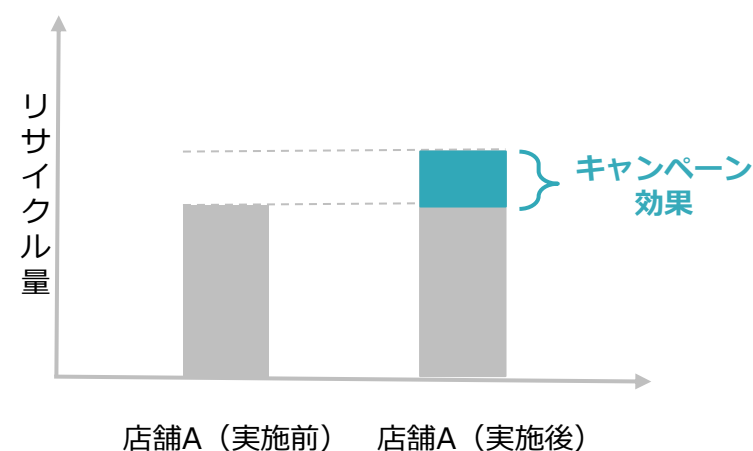


Case2

別店舗・複数店舗との比較



Case1



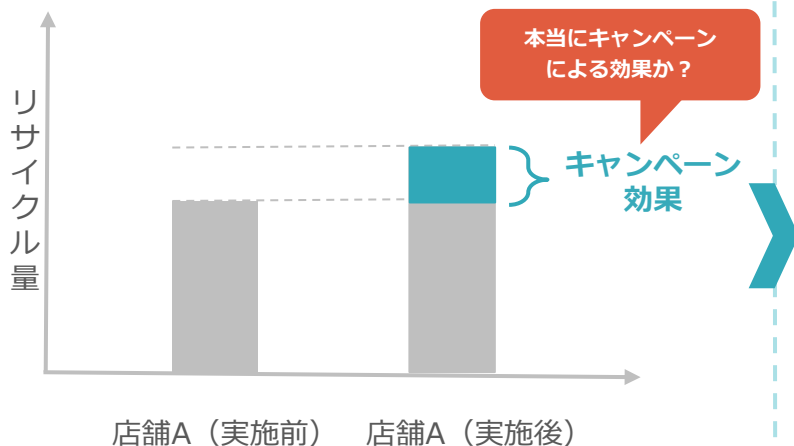
キャンペーンは本当に効果があったのか？
目的とする結果（リサイクル増加、顧客流入増加など）を達成したか定量的に評価

・キャンペーンの効果測定（その他の手法）

統計的検定ではキャンペーン実施前後のデータや異なる群同士の比較により、平均値や中央値の差異が統計的に有意かどうかを評価できる（差があるかどうか分析できる）。ただしその差は本当にキャンペーンによるものなのか、それともほかの要因によるものなのかの真実是不明。より詳細かつ正確な効果測定には例えば統計的因果推論の適用が考えられる。

※コストや目的に応じて、どの程度の正確性を求める必要があるかは事前に考慮が必要。

Case1



因果推論を用いた効果検証

リサイクル量増加（減少）はキャンペーンによるもの以外に、例えば

- ・季節性要因
- ・曜日や時間帯
- ・その他のイベントや外部環境
- ・経済状況
- ・競合の活動
- ・顧客層の変化 など

これらの影響を因果推論の手法（傾向スコアマッチングや差分の差法など）で解決

結果のご提供（一例）

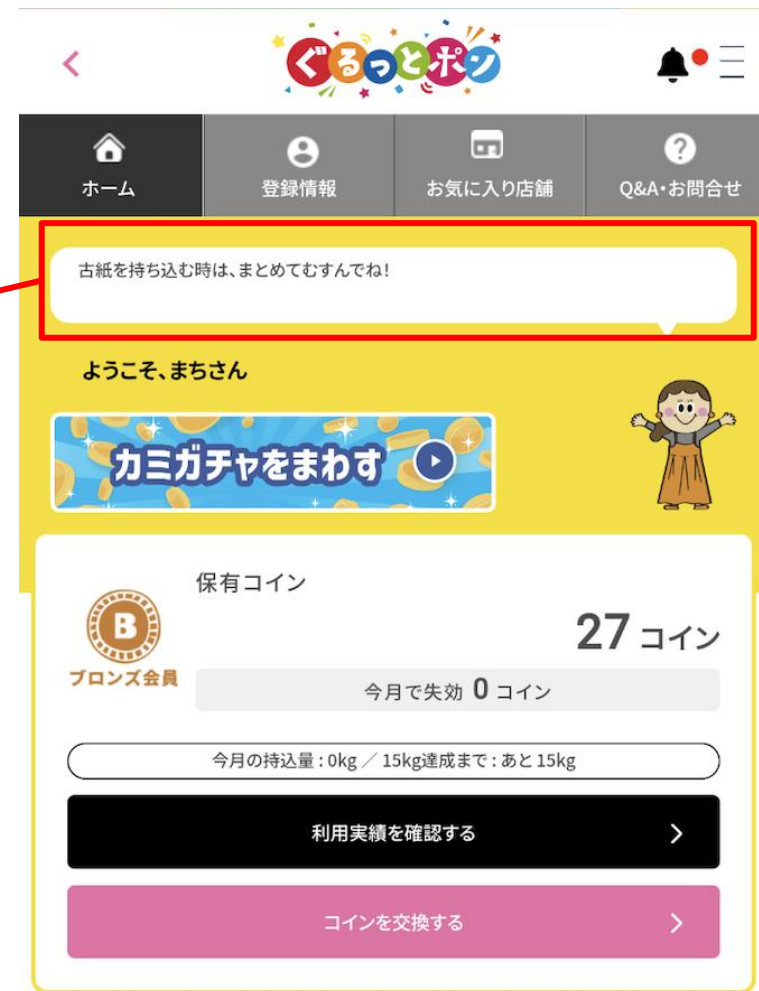
- ・ダッシュボード上での視覚的な報告
- ・ファイルストレージへの保存
- ・メールでの送信

例えばキャンペーン実施以外のリサイクル量に影響を与えうる、未観測変数を考慮して正確な効果検証を行う
統計的に有意な差があるかどうか＋効果量の大きさまで評価が可能



アプリ画面上部にユーザー名と達成内容を通知する

シルバー以上のランクアップや、ユーザー近隣店舗での大量投入、寄付時にニュースとして通知する
大量投入、寄付の内部的動機づけとなる



実装期間：20h

※実装には設計・UT・IT・ST含むものとする。



その他の提案もご用意しております。

・ユーザーへの新たなキャンペーンの提案

①アプリ使用経過日数が長いユーザーには「アプリ忠誠度報酬」を、新しいユーザーには「初心者ブースト」のような特別なイベントやボーナスを提供し、リサイクル機会の捻出を図る

②地域ランキング、コンペティション

店舗、もしくは地域ごとのランキングやコンペティションを設定し上位のユーザーに報酬を提供

アプリ内上部のエリアにお気に入り登録店舗のランキング一位のユーザーを告知してもいい、内部動機付けになる。

コンペは地域対抗リサイクルバトルなどで、リサイクル量上昇率一位の地域のユーザーにボーナスポイント付与することでリサイクルの活性化を図る。

その他の提案もご用意しております。

・ユーザーへの統計情報の可視化

ユーザー自身へリサイクル活動やアプリ使用状況に関する洞察を提供。

例えば、統計情報として累計リサイクル量、月平均リサイクル量やガチャ使用状況など

図として月のリサイクル量の遷移のグラフを表示することで自身のリサイクルライフを客観的に見ることができ、リサイクルに対する意識が高まる。

・アプリ通知によるリマインド

最近リサイクルしていない人にリサイクル活動のリマインド(アプリ通知)をする。

こういう傾向がある人は離脱するという調査をした上で**一年に一度効果的なタイミングで通知してあげる**ことで離脱したユーザーをリサイクル活動に戻り、リサイクル量、アプリ利用率の向上が見込める。

自然言語モデルによる施策のご提案

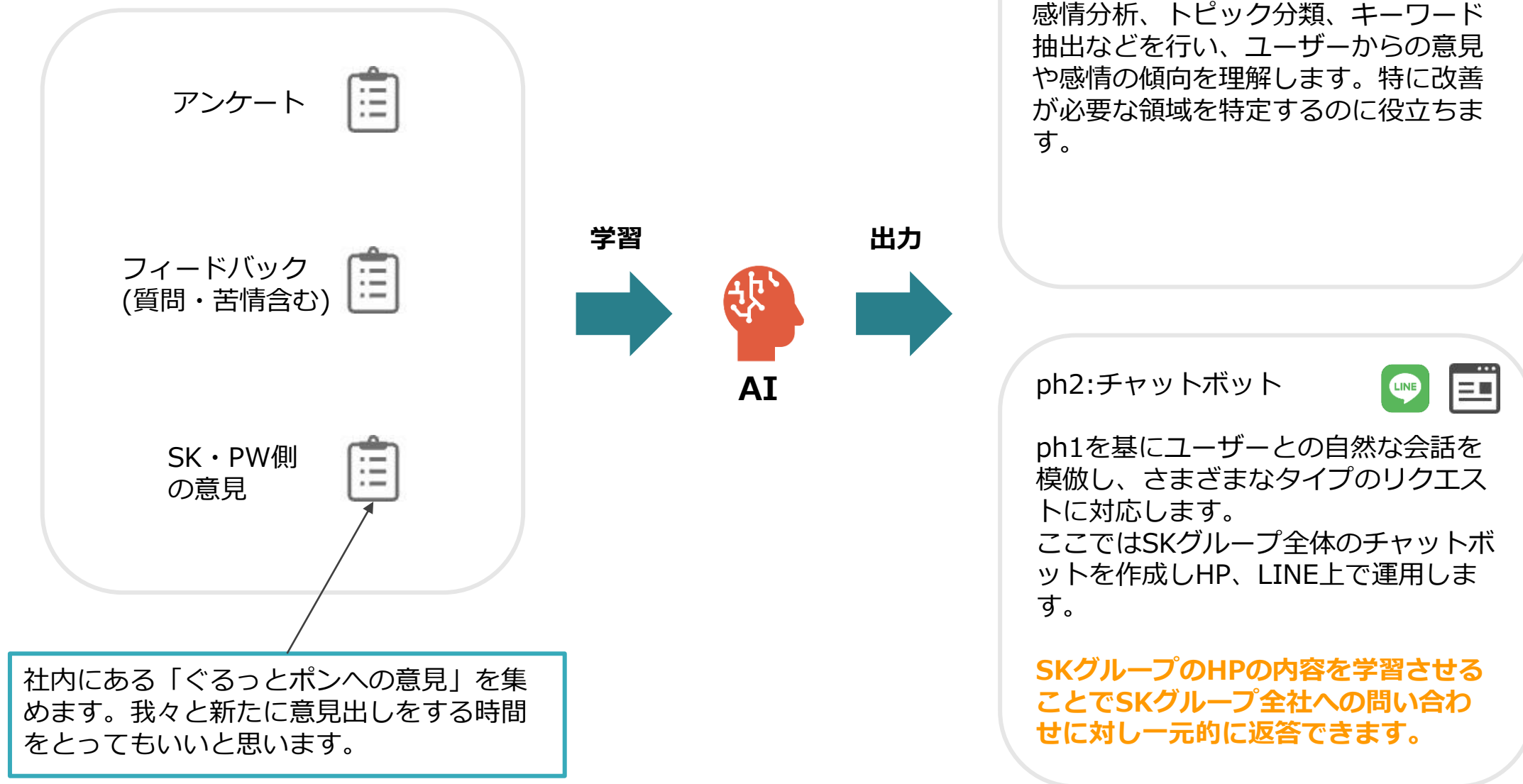
現状の課題：

アンケートやフィードバックが溜まっているが分析できていない。

SKグループ全体として情報が溜まっているが活かせていない。

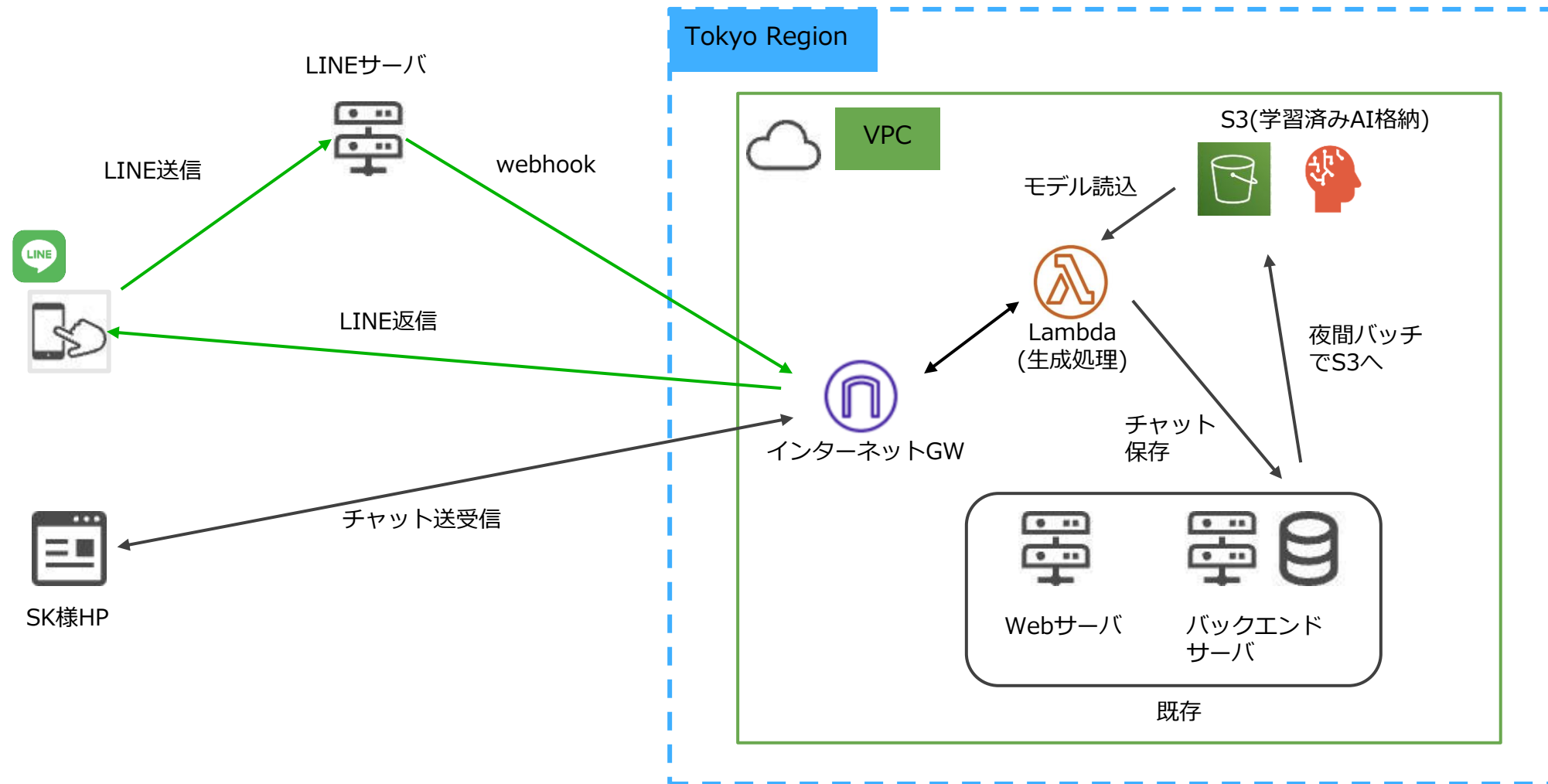


自然言語モデルによる施策のご提案





自然言語モデルによる施策のご提案：アーキテクチャ図(簡易版)



自然言語モデルによる施策のご提案：コスト

工数

①フィードバック分析（ph1）データ前処理、モデルの選定・ファインチューニング費用

40h x 4名

②チャットボット（ph2）PoC・AI開発費・システム組み込み費

240h x 4名

ランニングコスト

AIサーバーは置かずにストレージに訓練済みモデルを格納しサーバーレス関数で運用することでコストはほぼかからない

- ・ストレージコスト：約\$0.0115/月
- ・Lambda実行コスト：\$0.05/月
- ・データ転送コスト：\$0（同一リージョン内）

年1回、モデル更新のコストのみ発生する

30h x 2名

自然言語モデルによる施策のご提案：利益

ここでは主に③チャットボットまでを実装した場合、下記が利益として期待できます。

①人件費の削減

グループ全体での問い合わせ回数の減少による人件費の削減。

②顧客満足度向上による利益

顧客満足度向上によるグループ全体の売上向上。

③ブランドイメージの向上

チャットボットを通じて提供される迅速かつ一貫した情報提供は、ブランドイメージを向上させる効果があります。長期的な顧客ロイヤルティやブランドの信頼性向上に寄与します。

08

MDXQ後の継続ご提案





継続をご希望いただける場合、**プランA**もしくは**プランB**にて有償契約締結のご検討をお願いします。

プラン詳細

プランA

作業内容

- PoCにより実現性が高いと判断されるタスクが対象として、要件定義・設計・開発のすべてをチーム側で実施
- 完成後の保守・更新も実施

方針

- 月稼働想定**40**時間。大きく超過する場合には、プロジェクトの納期を長めにする。
- 一部メンバーはプランBで参加の可能性あり
- チーム側は会社設立を検討

参加人数

- 最大4名

保守

-
-

プランB

作業内容

- MDXQの延長線上で、データ可視化・分析、PoC
- MDXQの内容に捕らわれず、SK様側と自由に議論し、SK様の事業に貢献する施策を検討
- PoC結果次第でプランAへの移行を検討

方針

- 月稼働上限は**20**時間
- 稼働時間変更の場合、比例して報酬額を変更

参加人数

- 最大6名 (ただし、うち1人は無償契約にて開始)

共通事項

-
- 上述の報酬は各プラン参加者1人当たりの報酬
- 活動による成果物(ソフトウェア等)を含む金額
- 現地での調査・打ち合わせなどに係る出張費用は含まれない
- 開発に必要な備品など購入費用は報酬金額には含まれない

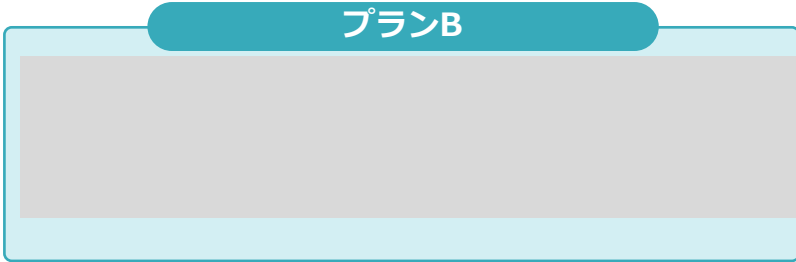
個別条件

-
-



継続をご希望いただける場合、**プランA**もしくは**プランB**にて有償契約締結のご検討をお願いします。

プラン



期間

固定契約期間 + 1か月ごと契約更新

- 固定契約期間を2か月とし、その後は毎月更新
- 更新は双方の合意が必要
- 解除はSK様側の判断で可能（報酬に見合わない場合は解除頂けます）



運営

隔週定例会の設置

- 隔週の定例会を設置し、進捗モニター、稼働時間報告

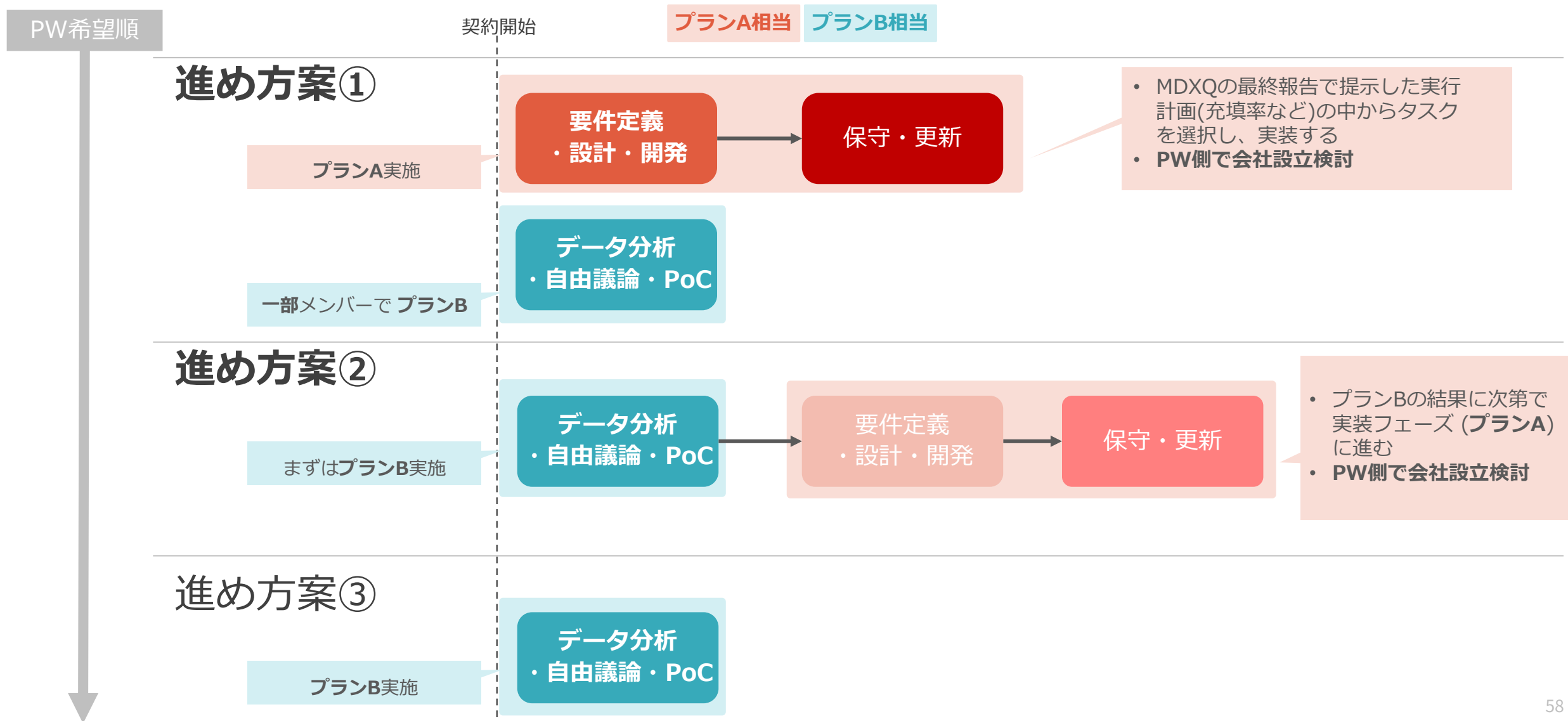
契約

契約書の締結

- MDXQ参加時の誓約書をベースに、プロジェクトのスコープ、期間、報酬、運営方法なども盛り込んだ契約書を締結する。



MDXQのPoC結果にて投資ご判断頂ける場合は、すぐに**プランA**を開始、そうでなければ**プランB**にて再度課題の設定や実現性、投資対効果を議論させてください。



09

振り返り





概ね皆様にとって有意義で、一定の満足感の得られるMDXQ企業協働プログラムとなった。
プログラム内で着地できたのはよかったが、MDXQに囚われずもっとお互いの利益になるゴール設定もあり得る。

MDXQ企業協働全体の満足度 (5点満点)

平均 4.4

*全投入者、5点もしくは4点

総評・コメント (SK様)

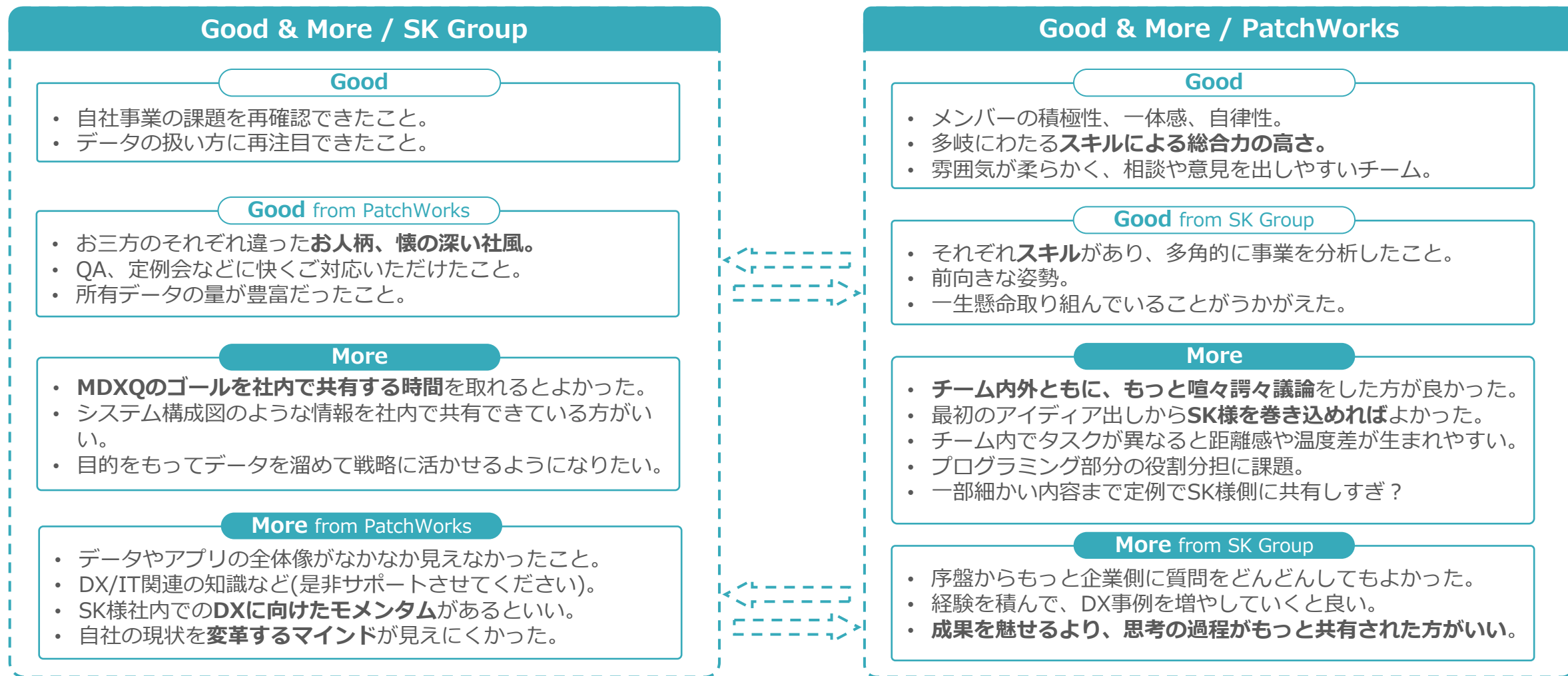
- ・ **DXスキル・経験のあるチームとマッチング**してよかった。
- ・ 内容の割に時間が限られる。
- ・ **時間をかけられれば相互のニーズを捉えた活動ができ、より実現可能性の高いことにチャレンジする時間が作れたのではないか。**
- ・ 今回の成果物は、SK社内で大切な指針となりえる。
- ・ チーム側は短納期、またデータやり取りがうまく進捗しない中でも、尽力頂いた。

総評・コメント (PatchWorks)

- ・ **実データ**に触れる貴重な機会。
- ・ **DXの具体的なプロセスを経験**できる。
- ・ 楽しく苦しい3ヶ月だった。
- ・ **有意義**だったが、まだまだやれることは多い。
- ・ SK側学生さんとも連携取って何かできそう。
- ・ 短期間でDXプロジェクトの一連のプロセスを回し、無事着地できたことはよかった。
- ・ 双方にとって無意味ではなかったと思う。



SK様の社風と皆様のお人柄、PatchWorksのスキルとチームワークにより、円滑なプロジェクト運営ができた。
今後は双方が目標や過程・熱量を共有し、一体感を持って自由闊達に交流することで、より良いプロジェクトとなり得る。

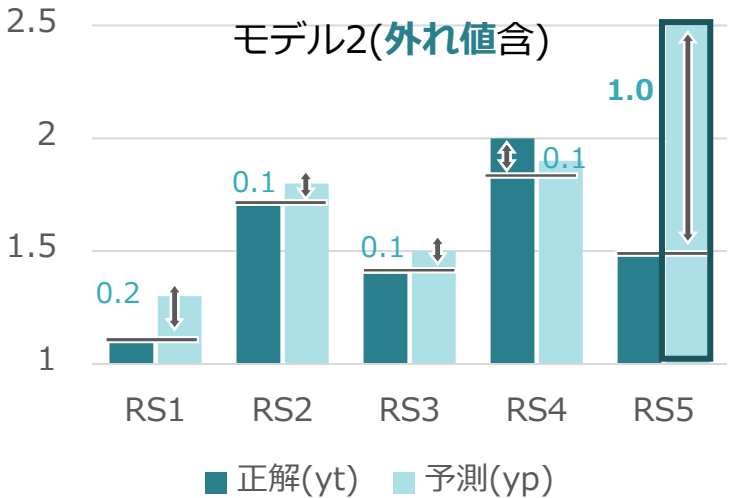
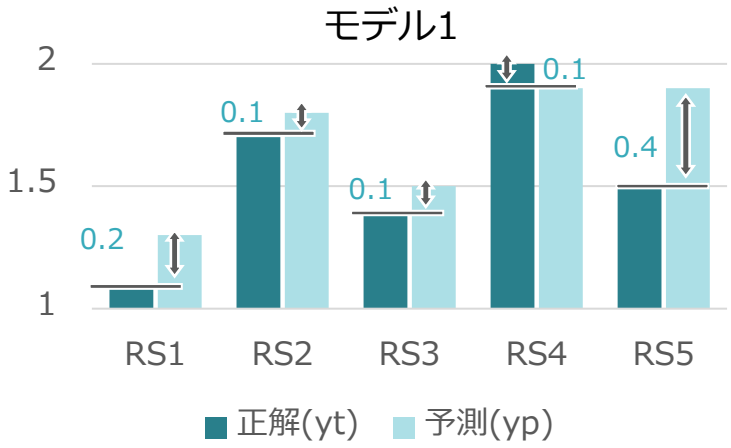


Appendices





AIの精度は、予測したい値(正解値)とAIの出力値(予測値)がどのくらい異なるか(誤差)によって評価します。この時、誤差の測り方には様々な種類があります。今回はRMSEを採用しています。



MAE (平均絶対誤差)

特徴

- 絶対値を平均するため、予測値や正解値に少数の外れ値があっても、それ以外の誤差が小さければMAEの誤差も小さくなる(外れ値に鈍感)
- 二乗していないので、誤差の大きさを直感的に理解しやすい

Calculation

モデル1 = $(0.2 + 0.1 + 0.1 + 0.1 + 0.4)/5 = 0.18$

↓ 167%

モデル2 = $(0.2 + 0.1 + 0.1 + 0.1 + 1.0)/5 = 0.30$

MSE (平均二乗誤差)

特徴

- 誤差の二乗を取るため、予測値もしくは正解値に少数の外れ値があると、それ以外の誤差が小さくてもMSEの誤差が大きくなりやすい。(外れ値に敏感)
- 二乗するので、誤差の大きさを直感的に理解しにくい

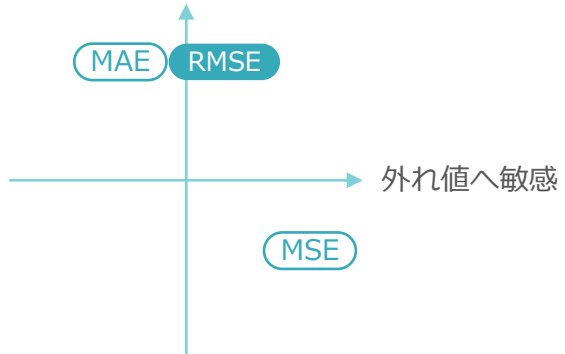
Calculation

モデル1 = $(0.04 + 0.01 + 0.01 + 0.01 + 0.16)/5 = 0.046$

↓ 465%

モデル2 = $(0.04 + 0.01 + 0.01 + 0.01 + 1.00)/5 = 0.214$

直感的に理解しやすい



RMSE (平方根平均絶対誤差)

特徴

- 誤差の二乗を取るため、MAEより外れ値に敏感だが、平方根を取るためMSEより鈍感。
- 二乗後に平方根を取るため、誤差の大きさを直感的に理解しやすい

Calculation

モデル1 = $\text{MSE}(0.046) \text{の平方根} = 0.21$

↓ 219%

モデル2 = $\text{MSE}(0.214) \text{の平方根} = 0.46$



AI精度指標の一つとして決定係数が登場します。相関関係を表す相関係数から派生した指標です。

相関係数

$$= \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \times \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}}$$

相関係数とは何か

- **相関係数とは**：相関関係の強さを表す指標。因果を示すわけではない。
- **直感的な理解**：相関関係とは、2つの変数(x,y)について、一方の変数が増えた時に、もう片方の変数も増えるような関係性
- **計算方法**：(x,yの共分散) / {(xの標準偏差)x(yの標準偏差)}=>[詳細](#)
- **計算結果**：最小-1、最大+1 (結果の解釈及び例は下記参照)

因果関係との違い

- **因果関係とは**：2つの事象について、片方が原因で、もう片方が結果である関係。どちらが原因でどちらが結果かわかっている必要がある。
- **例**：海の溺死者数とアイスの販売数量が強い相関を示したとしても、どちらかが原因(結果)であるとは言えない。
- **測定方法**：代表的なものとしてランダム化比較試験が挙げられる。

決定係数R^2

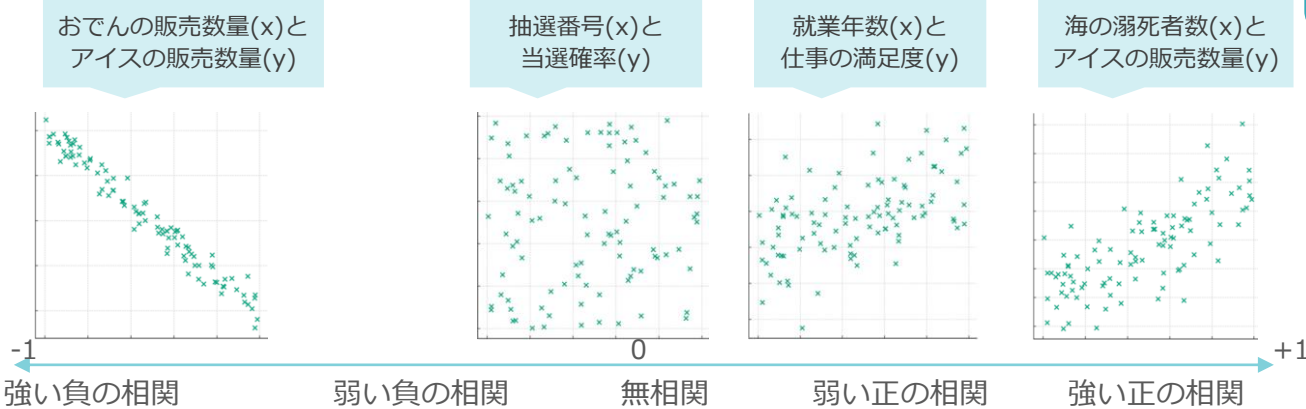
$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

決定係数とは何か

- **決定係数とは**：モデルが複数説明変数が目的変数の変動をどれだけ説明しているかを表す指標で、0~1の間の値を取る。
- **計算方法**：1 - {残差平方和(RSS) / 偏差平方和(TSS)}
- **直感的な理解**：RSSは誤差の大きさ=モデルの性能、TSSはそもそもデータがどれくらいばらけているか。RSS/TSSの値は、モデルの性能を、データのばらけ具合で調整した指標と言える。

名前の由来

R^2という名前は、元々相関係数から発展したことに由来。1変数の線形回帰(Y=aX+b)においては、“相関係数の2乗値”というのは、“手元のデータから仮に最適なパラメータを見つけたとすると、モデルの予測値と正解値の誤差がどのくらいになるか”を示し、この値が上記の決定係数の結果と一致する。





今回は、世界的に広く使われている、軽くて速い”LightGBM”という機械学習モデルを採用しています。

LightGBM

LightGBM(Light Gradient Boosting Machine)のモデル概要

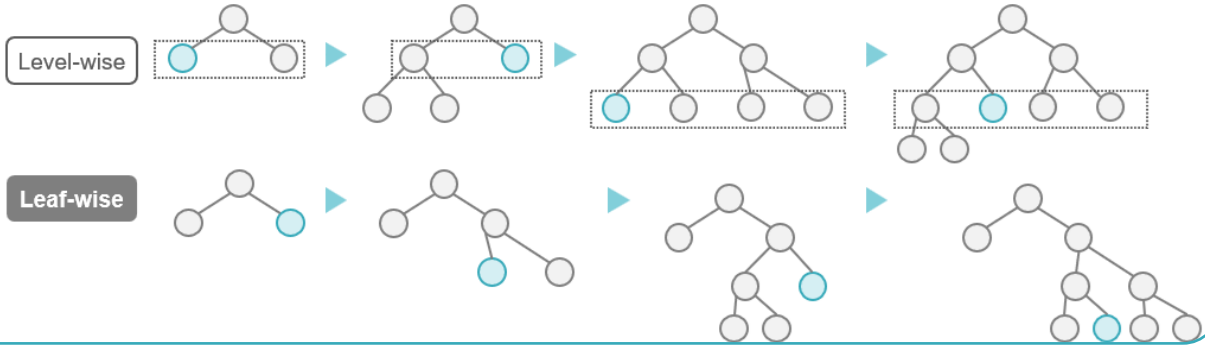
Microsoftによって開発された、勾配ブースティングによる機械学習モデル。2016年に発表されて以降、国際的なAIモデル構築のコンペティションで最も使われているモデルの1つ。

勾配ブースティング (Gradient Boosting)

ブースティング(右記参照)の一種。前の弱学習器の残差(正解値と予測値の差)を学習する手法。具体的には、最初の弱学習器は入力データから正解値を予測するが、その残差を次の弱学習器が学習対象とする。残差とは例えば、予測したい古紙の量と、モデルが出力した古紙の量の差。なお”勾配”というのは、モデルをほんの少し変化させた時に、残差がどのくらい変わるか、という数学的な情報。この”勾配”を使って残差が少なくなるようにブースティングの手法でモデルを学習させるので、”勾配”ブースティングという。

(参考)LightGBM登場の背景

2000年代以降ランダムフォレストというモデルが多用されていた。これはTree=木(決定木)をたくさん備えたForest(森)という言葉通り、決定木を用いたバギングベースのアンサンブル学習モデル。2014年に勾配ブースティングベースのXGBoostが登場し、高精度を誇ったが、学習に時間を要した。そこで”Lightな”(軽い)勾配ブースティング手法として登場したのがLightGBM。具体的には、XGBoostが当初採用した同じレイヤー(level)の分岐を作るLevel-wiseではなく、性能に寄与する分岐を作ることを優先したLeaf-wiseを採用。



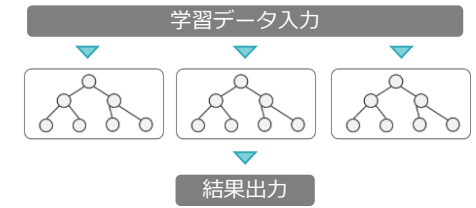
アンサンブル学習

概要

複数の単純な機械学習モデル(弱学習器)を組み合わせ、一つの強力なモデル(強学習器)を作る技術の総称。単一のモデルよりも精度が高く、予測の信頼性が向上するのが特徴。弱学習器には決定木が用いられることが多い。弱学習器の組み合わせ方によって、バギング、ブースティングなど手法が細分化される。

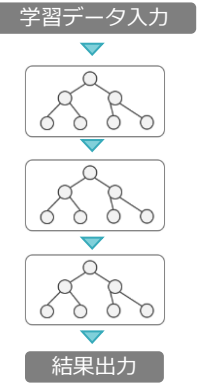
バギング

弱学習器を並列に学習させる。最終的には多数決などで出力。計算が速いが学習不足になりがち。



ブースティング

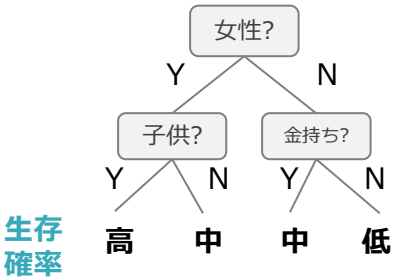
弱学習器を直列に学習させる。学習結果を次の弱学習器に渡すことで精度が高くなりやすいが、計算が遅い。



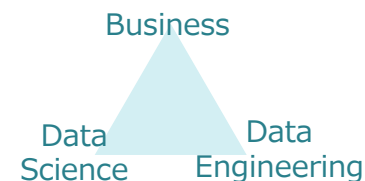
決定木

アルゴリズム概要

データ分類の単純なアルゴリズム。質問を用意して回答によってデータを分岐させ、最終的にカテゴリ分類などの出力を得る。右はタイタニック号の生存者の例。



メンバーの多くがAI・ITに関する知識・経験を持っています！



chef (リーダー)

B D S

- 精密機器メーカーで海外営業・マーケティング等10年
- データサイエンス/プログラミングは独学で少々 (G検定、E資格、VBA、Python)
- 昨年も企業協働での実績あり
- chef = オーケストラの指揮者

リーダーとして、企業様・チームの**双方にとって良いプログラムだったと言って終われること**を目標に頑張ります！

naoya (サブリーダー)

D S D E

- 日系SIerと中華系SIerを6年経て個人事業主
- サイバー大学AIテクノロジーコース在学中
- 顧客折衝から要件定義、開発まで一通り経験
- こちらからも積極的に提案させていただきます

課題解決をしたい企業様とAIの実務経験を積みたい我々の**双方にとって価値のある成果を出すことが大切**と考えております。

まっち

D S D E

- メーカー研究所の機械系開発職
- 大学院では制御の研究
- 業務で画像分類系のアプリ開発経験あり
- 自社でデータ駆動型研究所の実現を目指す

貴重な機会をいただきありがとうございます！
試行錯誤の末に、お互い納得のいく結果に至れるよう、**尽力させていただきます！**

norosen

D S D E

- メーカーで基礎研究
- Pythonでのデータ分析や数値最適化が主業務
- 研究でのAIモデル活用に向けて試行錯誤中

このまたと無いチャンスを絶対に良い経験にしたいと思っています！お話をきちんと伺って、**企業様にとって価値あるものを一緒に作っていきたい**です。

Saito

D S D E

- 大学院卒業後大手SIerに勤務 (2年目)
- 社内ワークショップ運営や技術相談に対応
- AI関連知識保有 (E資格取得、コンペ参加)
- 各種ITスキル保有 (情報処理安全確保支援士合格、クラウド、アプリ開発経験あり)

企業様の**潜在的な課題**も見逃さず、**現実的かつ革新的な提案**をしたいと考えております。

Tomi

B D S

- 現在大学4年生
- 地方創生×データサイエンスの実証研究中
- 院卒業後データサイエンティストを目指す

大学での学びが実務にどう活かせるか、この貴重な経験を通して学ぶ機会を頂けることに、心より感謝申し上げます！企業様と協力し、**プログラムを充実させていききたい**です！

“PatchWorks”は、IT・AI専門知識を有するメンバーが、
和気藹々とコミュニケーションを取りながら、
DXの実現を目指して目指して活動します。

What

<“PatchWorks”とは？>

パッチワークとは**小さい布切れをつなぎ合わせて一枚の布にする**手芸の技法です。

我々のミッションはDXを実現することですが、それは則ち**体系化されていないアナログ作業をAI・ITのスキルでつなぎ合わせ**、システムとして調和させることであると言えます。

DXのプロセスも、DXを目指すために集まった我々も、さながらパッチワークのようです。私たちは、このパッチワークの精神をもって、DXにより新たな価値を創造できるようにと願いを込めました。



How

当グループは以下の点を重視して結成されました。企業様との協業においても、初志貫徹で活動します。

<活動方針>

1. “個人の時間をすべて捧げる”という意味での**フルコミットを求めない**。それぞれの専門性とペースで最大の成果を追求する。
2. 企業様とチームがビジョン・ゴールを共有して取り組めるように**良好なコミュニケーション**を心がける。同時に、**和気藹々**とした雰囲気的大事にする。
3. プロジェクトの**完遂**に向け、持続可能なコミットメントにより確実な成果を目指す。

<活動方法>

企業やチーム内の重要な定例mtgを除き、全員参加を強制しない。自主性を尊重し、柔軟な参加を推奨する。

Why

<Mission / Vision / Purpose>

私たちは、**DXを実現したい**という強い想いを抱いています。

単なるデジタル化やシステム化に留まらず、AI等を用いた**より高度な課題解決、ビジネスモデルの構築**など、付加価値の高い成果を出すことを目指します。

