

東京大学グローバル消費インテリジェンス寄付講座

GCI 2020 Winter | 最終課題

 [tom-uchida](https://github.com/tom-uchida)

2021/02/09

本課題における状況設定(モデル化)

対象とする市場・商品

日本・携帯電話

A社のシェア率

A社：1割（NTTドコモ：4割，au：3割，ソフトバンク：2割）

A社の契約台数

185,228,700台(日本における携帯電話の総契約台数^[1]) × 0.1(A社のシェア率) = 18,522,870台
→ **1800万台**とする

[1] 一般社団法人 電気通信事業者協会, “事業者別契約数”, <https://www.tca.or.jp/database/>

A社の契約者数

1人につき1.5台を契約しているとする：
1800万台 × 2/3 = **1200万人**

A社から頂いたデータ

提供して頂いた10万件のデータ^[2]は，契約者1200万人から**ランダムサンプリング**されたものとする

[2] kaggle, “Telecom customer”, <https://www.kaggle.com/abhinav89/telecom-customer>

はじめに

事業の概要

まず、弊社の機械学習モデルによって、**解約者を予測**します。
その後、解約すると予測された顧客に対して、弊社の提案事業をご提供することで、**解約者数を減少**させ、**解約による損失額を大幅に減少**させます。

事業のご提案までの流れ

1. データ分析

ご提供して頂いた10万件・100種類のデータを分析することで、各データの重要度を確認します



2. 特徴抽出

各データの重要度を参考にしてさらに分析を進め、解約者の特徴を記述するデータを抽出します



3. 仮説構築

抽出した特徴をもとに、解約者に対する仮説を構築します



4. 事業提案

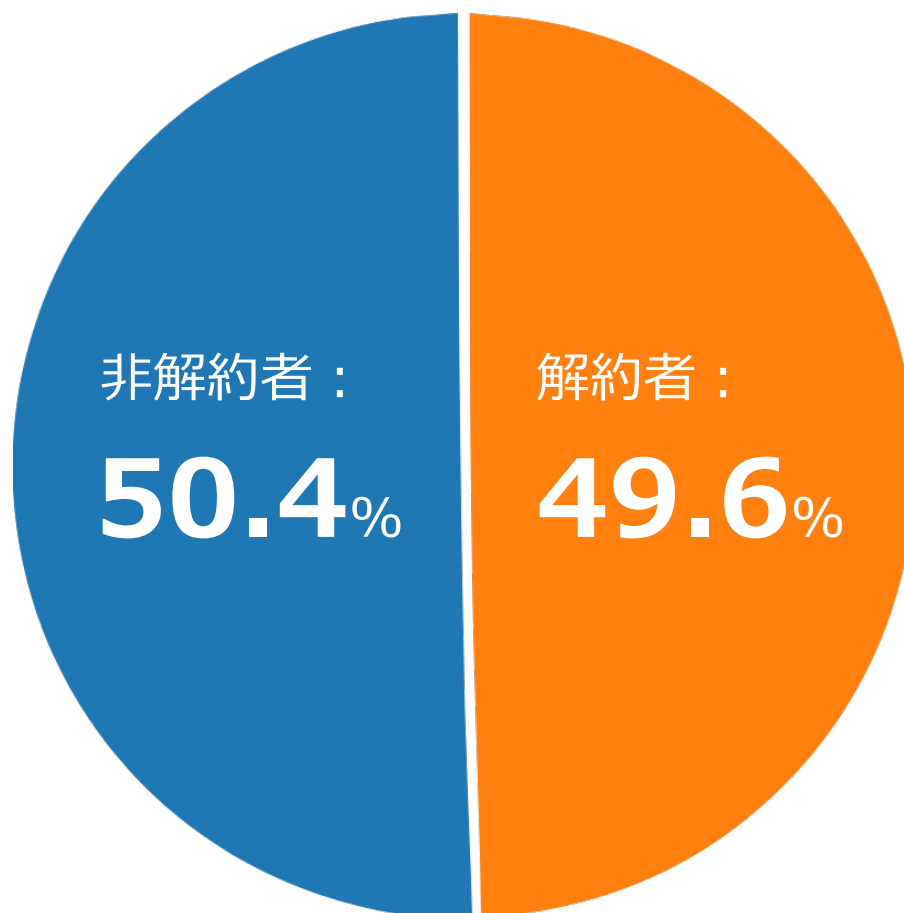
構築した仮説にもとづいて、事業のご提案をさせていただきます



5. 提案事業の効果検証

ご提案する事業を導入することによって得られる効果を、定量的に検証します

非解約者と解約者の割合



ご提供して頂いた10万件のデータを用いて、
非解約者と**解約者**の人数を調べたところ、
次のような結果が得られました：

非解約者数 : 50,438人

解約者数 : 49,562人

この結果より、

「解約率が約49.6%」

であることがわかりました。

解約による損失額のお見積もり

ご提供して頂いた10万人の顧客データを分析したところ、
A社様の月平均の損失額^{*1}は、次のように見積もることができます：

$$\begin{aligned}\text{損失額} &= \text{解約者に対する請求額の合計} \\ &= 2,870,796\$ \\ &= 298,103,457\text{円}^{*2}\end{aligned}$$

契約者全体(1200万人)に見積もりを拡大すると：

$$\begin{aligned}\text{損失額} &= 298,103,457\text{円} \times 120 \\ &\approx \text{約360億円}\end{aligned}$$

以上の見積もり結果より、解約者数を1%減らすことができれば：

解約による約 **3.6億** 円の
損失を防ぐ ことができます

^{*1}: 損失額は「解約者に対して請求するはずだった金額」とし、「解約者における rev_Mean(単位：\$) の合計」としてしています

^{*2}: 1\$ = 103.84円 (2021年1月17日 時点)

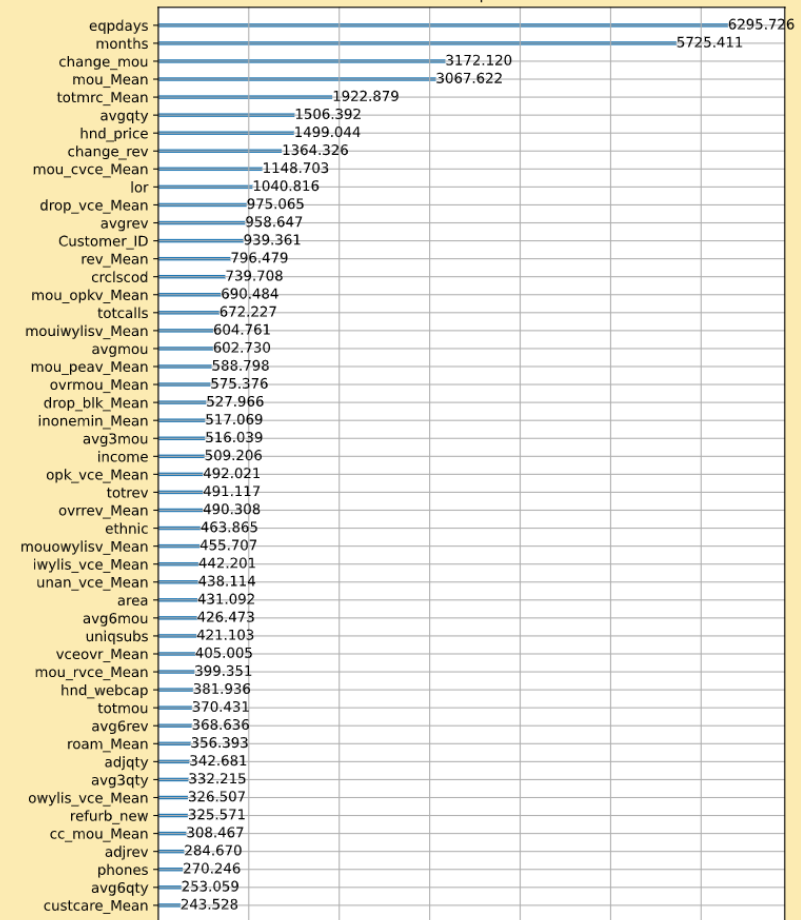
重要度が高いデータの抽出(1/2)

100種類あるデータに対して、LightGBM^[3]という機械学習モデルを用いて解約者の予測をおこなうことで、**各データの重要度の分析**をおこないました(右図)。

右図では、上位50種類のデータを表示していますが、これだけでは、「**どのデータまで参照するべきか**」という指標がありません。

そこで、次のスライドでは、パレートの法則^[4]：
「20%の要素が、全体の80%を生み出している」
の成立を仮定して、結果の80%を説明できるであろう、
「**重要度が上位20%のデータ**」を抽出^{*1}しました。

各データの重要度(上位50種類)

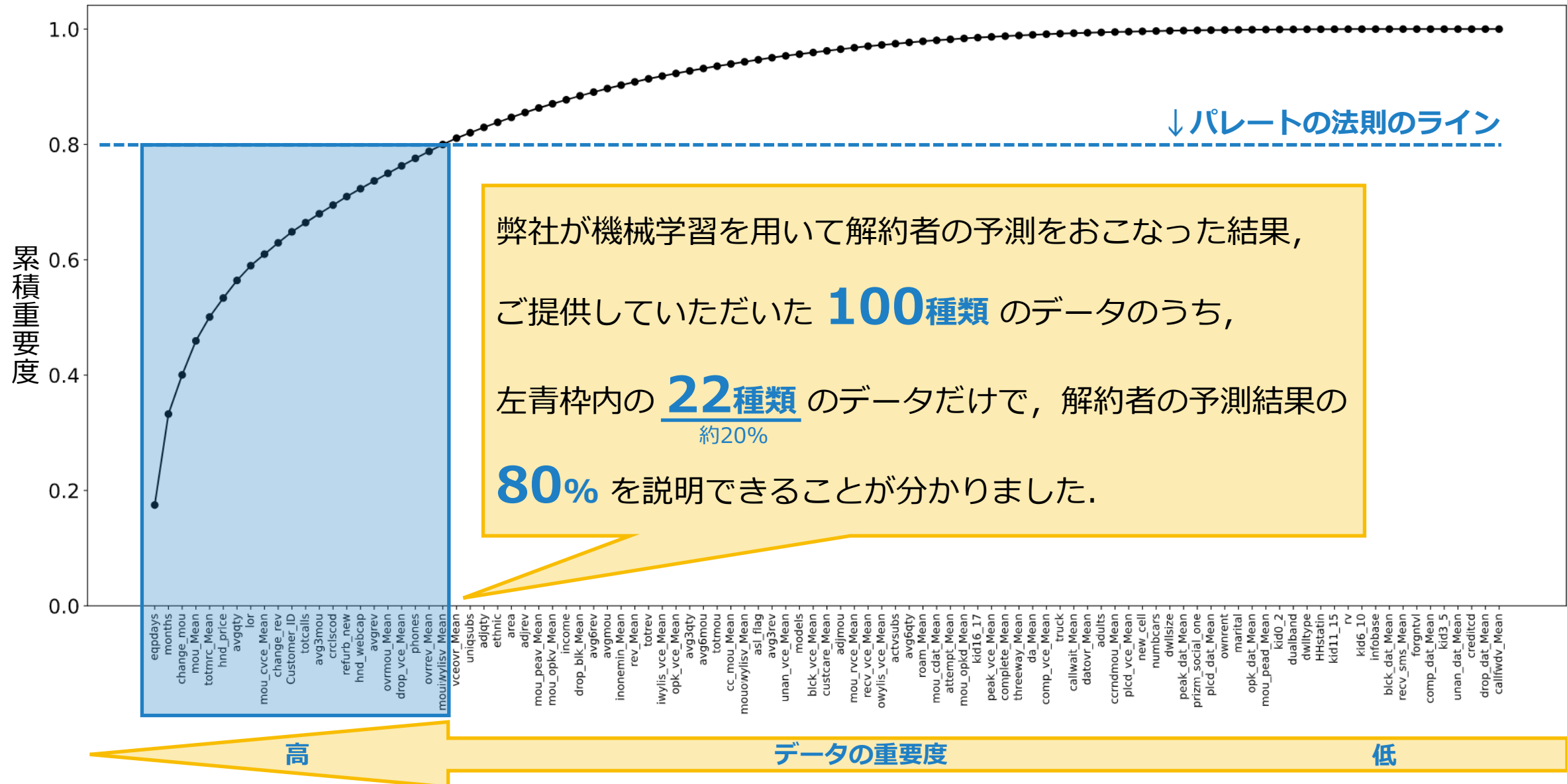


[3] Microsoft Corporation, "LightGBM", <https://lightgbm.readthedocs.io/en/latest/>

[4] 野村総合研究所(NRI), "パレートの法則", https://www.nri.com/jp/knowledge/glossary/lst/ha/pareto_princ

*1 : 100種類のデータの重要度を正規化した後に、重要度の累積値を計算することによって抽出

重要度が高いデータの抽出(2/2)



解約者の特徴分析(1/2)

前スライドで抽出した、重要度が高い22種類のデータを対象として、**解約者の特徴を表す手がかりとなるデータの抽出**をおこないました。

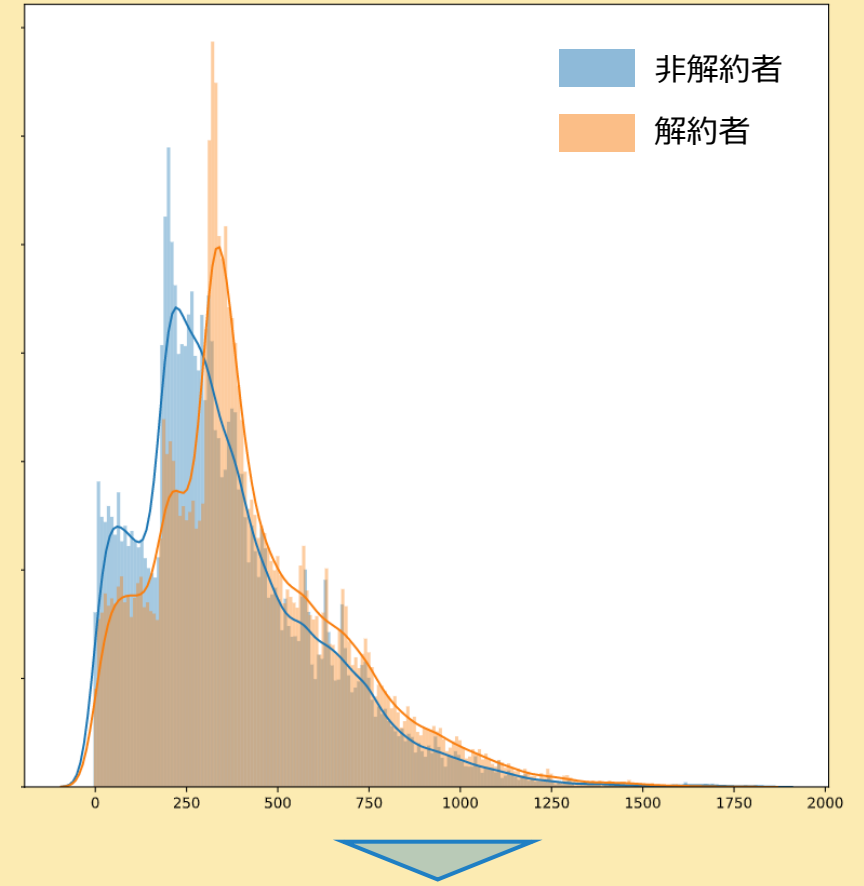
具体的には、非解約者と解約者ごとに各データの平均値と標準偏差を計算し、**解約者と非解約者で差が顕著であるデータを精査**しました。

その結果、弊社は、次の3種類のデータに着目しました：

1. **端末価格**(重要度：7位)
2. **端末の使用日数**(重要度：1位)
3. **端末の月平均使用時間の変化率**(重要度：3位)

次のスライドでは、上記の3種類のデータの分析結果から、**解約者に見られる3つの特徴を抽出**します。

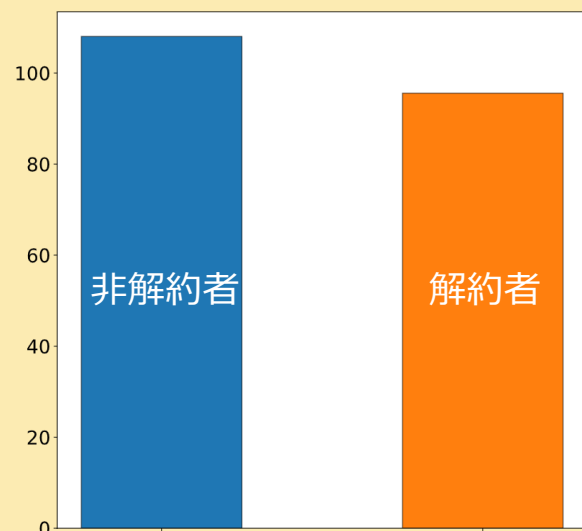
非解約者と解約者における端末の使用日数のヒストグラム



非解約者と解約者で差があるということは、**解約者に特有の特徴**なのではないか？

解約者の特徴分析(2/2)

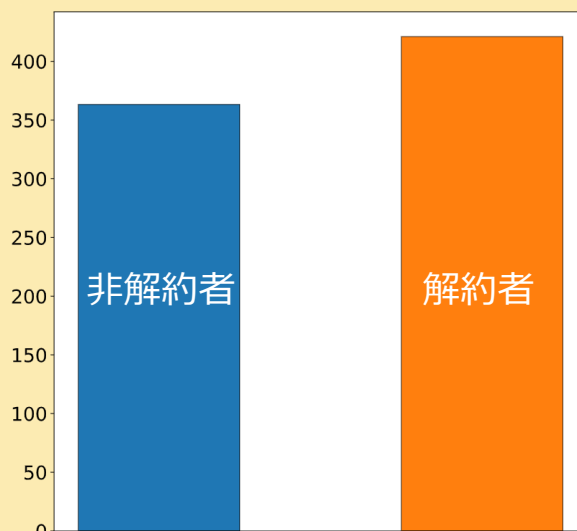
1. 端末価格の平均値



解約者の端末は安価

特徴1: 機種が古い

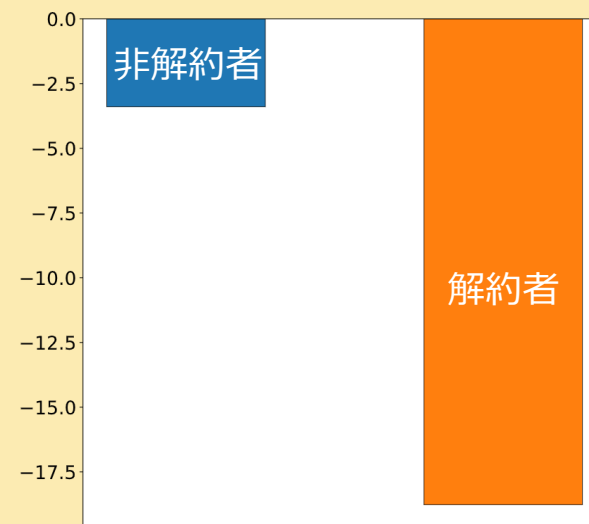
2. 端末の使用日数の平均値



解約者は使用日数が多い

特徴2: 端末が劣化

3. 端末の月平均使用時間の変化率の平均値



解約者は使用時間が減少

特徴3: 使用機会が減少

解約者に対する仮説・事業のご提案

仮説

解約者は、端末の **機種が古い** 上に **端末が劣化** してきており、
解約者の特徴 1 解約者の特徴 2
端末自体に魅力を感じなくなった結果、 **使用機会が減少** し、
解約者の特徴 3
解約に至ってしまっているのではないかと。

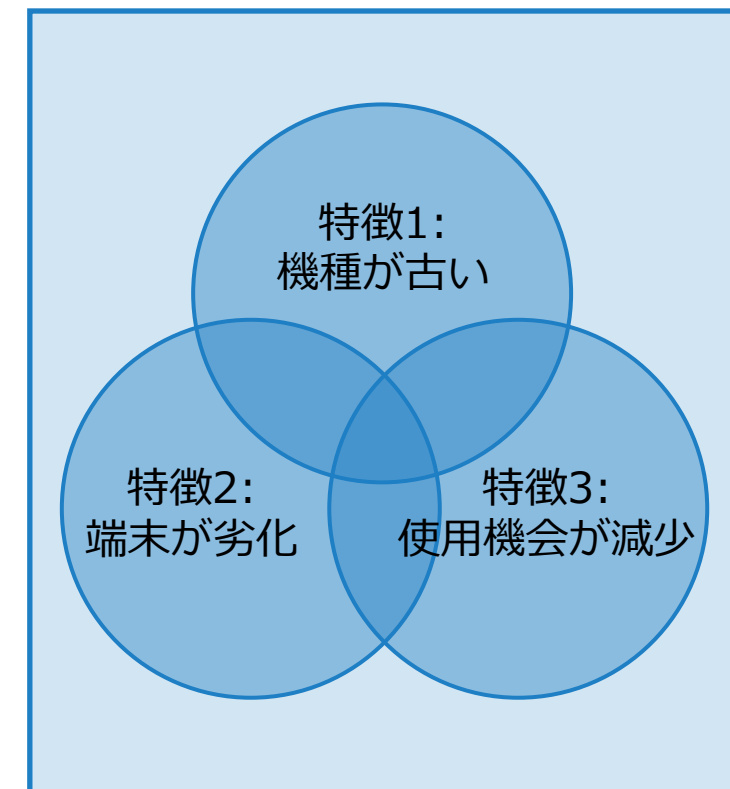


A社様への事業のご提案

弊社の機械学習モデルを用いて **解約者の予測** をおこない、
解約すると予測された顧客に対して、 **最新機種端末** を、
0円でご提供^{*1} することで、解約者の人数を減少させます。

*1: ただし、2ヶ月間は月々の請求額を1.5倍に増額し、その間は解約できないという条件を付けます

解約者の特徴



提案事業の設定(1/2)

提案事業によって期待される効果の反映

提案事業によって得られる効果をデータに反映させるために、解約すると予測された顧客に対して、以下の仕様で5種類のデータを更新します：

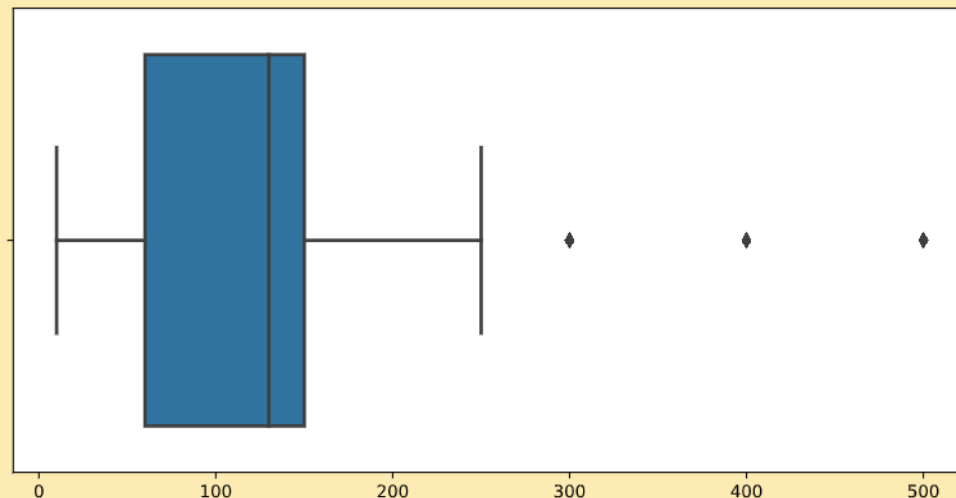
端末の使用日数	0日
毎月の請求額	×1.5 \$
月平均の請求額の変化率	+50 %
端末価格	150 \$ ^{*1}
月平均使用時間の変化率	+70 % ^{*2}

以上の効果を反映後に、再度、弊社の機械学習モデルで学習・予測をおこない、何も手を打たない場合の解約者の予測人数と比較します。

^{*1,*2}：次スライドで、この値に設定した根拠をご説明します

提案事業の設定(2/2)

非解約者の端末価格の分布

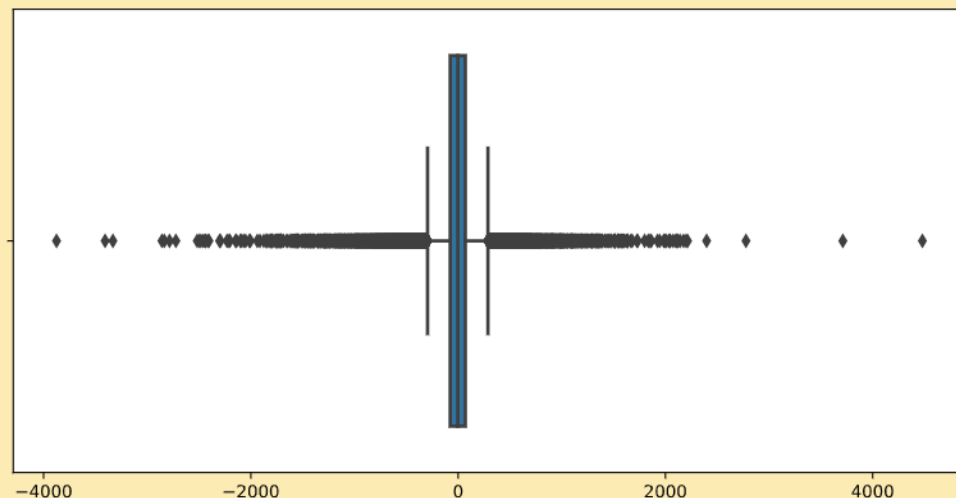


非解約者総数	50,438 人
平均値	107 \$
標準偏差	62 \$
最小値	0 \$
第二四分位数	60 \$
中央値	130 \$
第三四分位数	150 \$
最大値	500 \$

無料提供する最新機種端末の価格を、
非解約者における第三四分位数に設定。

最新機種端末の価格：
150 \$

非解約者の月平均使用時間の変化率の分布



非解約者総数	50,438 人
平均値	-5 %
標準偏差	250 %
最小値	-3875 %
第二四分位数	-77 %
中央値	-3 %
第三四分位数	70 %
最大値	4480 %

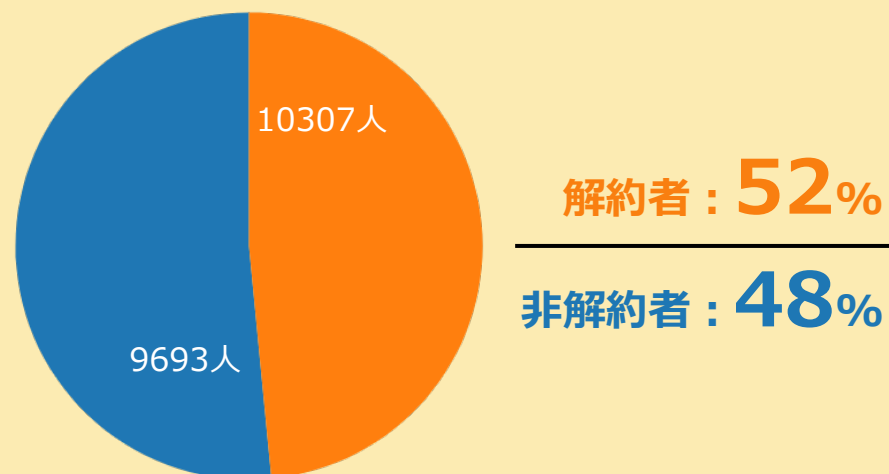
最新機種端末の無料提供によって
増加するであろう使用時間の変化率を、
非解約者における第三四分位数に設定。

使用時間の増加率：
+70 %

提案事業の導入効果

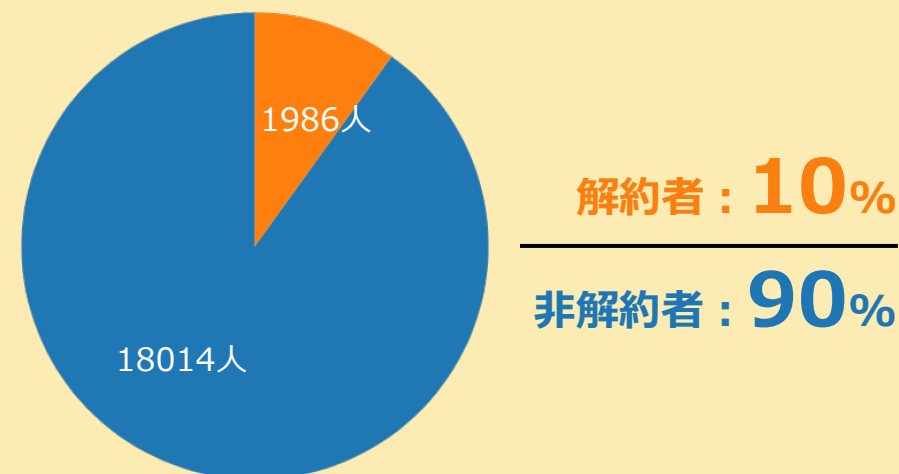
Before

何も対策を打たない場合の解約者数の予測結果*1,2：



After

提案事業を反映させた場合の解約者数の予測結果：



*1：学習と予測は、全10万件のデータを、学習データ：8万件、テストデータ：2万件に分割しておこないました

*2：予測精度：0.685(AUCと呼ばれる評価指標を用いています。AUCは、0~1の値を取り、1に近いほど予測精度が高いことを意味します)

弊社の提案事業を導入していただくと、**解約率** を約 **40%減少** させることができ、
解約による約 **150億円***3 の **損失を防ぐ** ことができます。

*3：3.6億円*4 × 解約率の減少率(52%–10%) ≈ 約150億円

*4：解約者数を1%減少させることで防ぐことができる、解約による損失額(スライド5を参照)

無料提供の条件の妥当性

最新機種端末を無料提供することによる損失額を A_{loss} とすると：

$$A_{\text{loss}} = 150\$ (\text{最新機種端末の価格}) \times \text{約} 50,000 \text{人} (\text{最初の解約者の予測人数}) \\ \approx \text{8億円}$$

一方、毎月の請求額を1.5倍にすることによる増収額を A_{profit} とし、増額前の毎月の請求額を R_{before} 、増額後の毎月の請求額を R_{after} とすると：

$$R_{\text{before}} = 60\$ (\text{非解約者の毎月の請求額の平均値}) \times \text{約} 50,000 \text{人} (\text{非解約者数}) \\ \approx 3 \text{億円}$$

$$R_{\text{after}} = 90\$ (60\$ \times 1.5) \times \text{約} 40,000 \text{人}^{*1} (\text{解約を防いだ人数}) + R_{\text{before}} \\ \approx 7 \text{億円}$$

$$A_{\text{profit}} = R_{\text{after}} - R_{\text{before}} \approx \text{4億円}$$

以上より、「 $A_{\text{profit}} \times \text{2ヶ月} - A_{\text{loss}} = 0$ 」となるため、最新機種端末を無料提供することによる損失額は、毎月の請求額の1.5倍増によって、**2ヶ月で回収可能**となります。
したがって、「2ヶ月間：請求額1.5倍かつ解約できない」という条件は妥当^{*2}だといえます。

*1：本提案の導入効果によって解約率が約40%減少します。最初の解約率が約50%(約5万人)であったため、導入後の解約率は約10%(約1万人)となります。したがって、本提案の導入によって解約を防いだ人数は、5万人 - 1万人 = 4万人 となります

*2：最新機種端末の価格・増額倍率・増額期間の設定や、他の新たなオプションの設定次第で、相殺ではなく利益を生み出すことも可能です

まとめ

A社様の現状

解約率：約50% | 解約による損失額：約360億円

解約者の3つの特徴

特徴1: **機種が古い** | 特徴2: **端末が劣化**している | 特徴3: **使用機会が減少**している

弊社が考える仮説

解約者は、端末の**機種が古い**上に**端末が劣化**してきており、端末自体に魅力を感じなくなった結果、**使用機会が減少**し、解約に至ってしまっているのではないかと考えられます。

弊社の提案事業

弊社の機械学習モデルを用いて**解約者の予測**をおこない、解約すると予測された顧客に対して、**最新機種端末を0円でご提供**^{*1}することで、**解約者の人数を減少させる**。

*1: ただし、2ヶ月間は月々の請求額を1.5倍に増額し、その間は解約できないという条件付き

提案事業の導入効果

解約率を約40%減少させることができるため、**解約による約150億円の損失を防ぐ**ことができる。