

東京大学グローバルインテリジェンス寄付講座

GCI 2021 SUMMER 最終課題

—解約者の予測と解約抑止案—

GitHub ID: subaru-1012

2021/08/02

ご提案の流れ

1. 御社のデータ分析から見る課題



2. 事業提案



3. AIによる解約者予測



4. 効果検証

1. 御社のデータ分析から見る課題

御社から頂いたデータについて

	rev_Mean	mou_Mean	totmrc_Mean	da_Mean	ovrmou_Mean	ovrrev_Mean	vceovr_Mean	datovr_Mean	roam_Mean	change_mou	change_rev	drop_vce_Mean	drop_dat_Mean	blick_vce_Mean	blick_dat_Mean	unan_vce_Mean
0	23.9975	219.25	22.500	0.2475	0.00	0.0	0.0	0.0	0.0	-157.25	-18.9975	0.666667	0.666667	0.0	6.333333	
1	57.4	0.0	52.333333	0.0	42.333333	0.0	45.000000	0.0	0.000000	0.000000	0.000000	18.000000	0.000000	90.643333	0.0	
2	16.9	0.0	263.333333	0.0	69.000000	0.0	193.333333	0.0	1.666667	6.333333	5.463333	53.000000	0.333333	189.396667	0.0	
3	38.0	0.0	9.000000	0.0	0.333333	0.0	6.000000	0.0	0.000000	0.000000	0.000000	0.333333	0.000000	5.426667	0.0	
4	55.2	0.0	3.666667	0.0	1.333333	0.0	3.666667	0.0	0.000000	0.000000	0.000000	1.333333	0.000000	8.410000	0.0	
	0.0	222.333333	0.0	94.666667	0.0	137.000000	0.0	8.666667	15.000000	11.076667	66.000000	0.000000	285.233333	0.0		

● 10万件・100種類の顧客データ(一部欠損あり)

データの内容

- ▷観察日から31～60日後に発生した解約事例 (churn)
- ▷平均月次定期料金総額 (totmrc_Mean), 平均月次利用分数 (mou_Mean)
- ▷月間利用分数の対前3ヶ月平均増減率 (change_mou)
- ▷現在の端末の利用日数 (eqpdays), 通算利用月数 (months) ...etc.

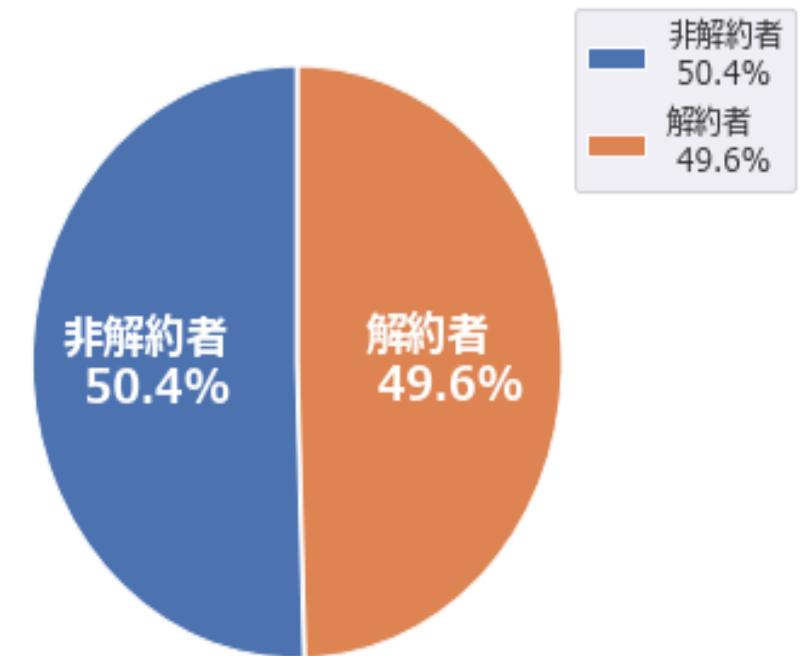
解約者の割合

ご提供いただいた10万人のデータから、**解約者**と**非解約者**の割合を調査したところ、以下の結果が得られました。

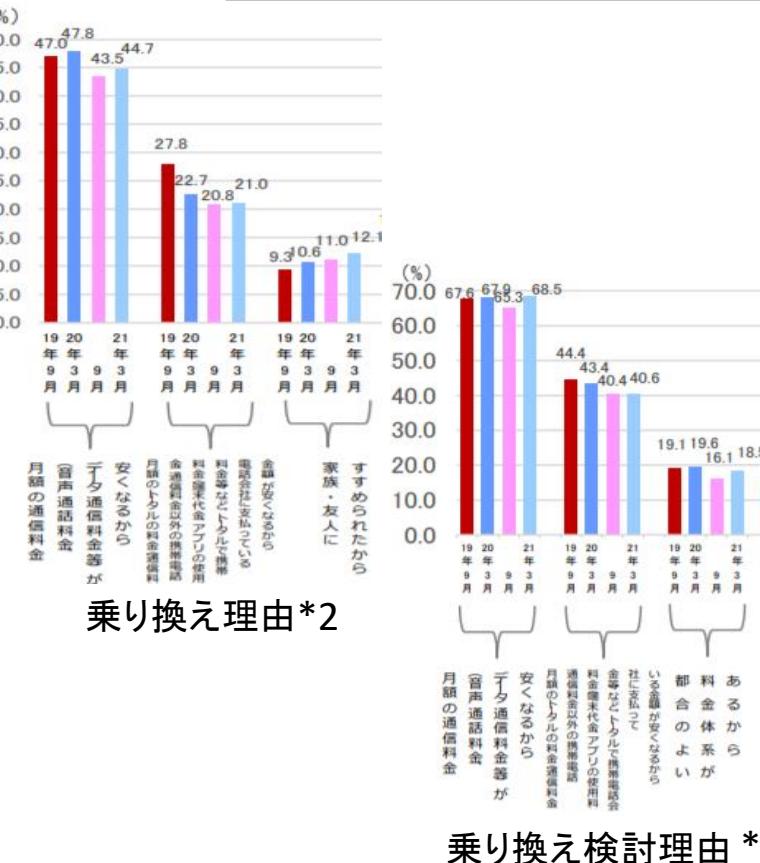
解約者数: 49,562人 (49.6 %)

非解約者数: 50,438人 (50.4 %)

➡ 解約者数が**49.6 %**と高い割合を占めており、**解約数を抑えることで、収益の増加が見込める**と考えられます。



解約の要因に関する仮説



2019～2021年度に総務省が行った「携帯電話利用者の意識調査」*1では、**乗り換え理由及び乗り換え検討理由**として、以下の上位2項目が共通して挙げられています。

1. 月額の**通信料金が安くなるから**(約45.8 %*2, 約67.3 %*3)
2. 月額の**トータル料金が安くなるから**(約23.1 %*2, 約44.2 %*3)

この結果から、「**現在の料金に対する不満**」が解約の要因として考えられます。

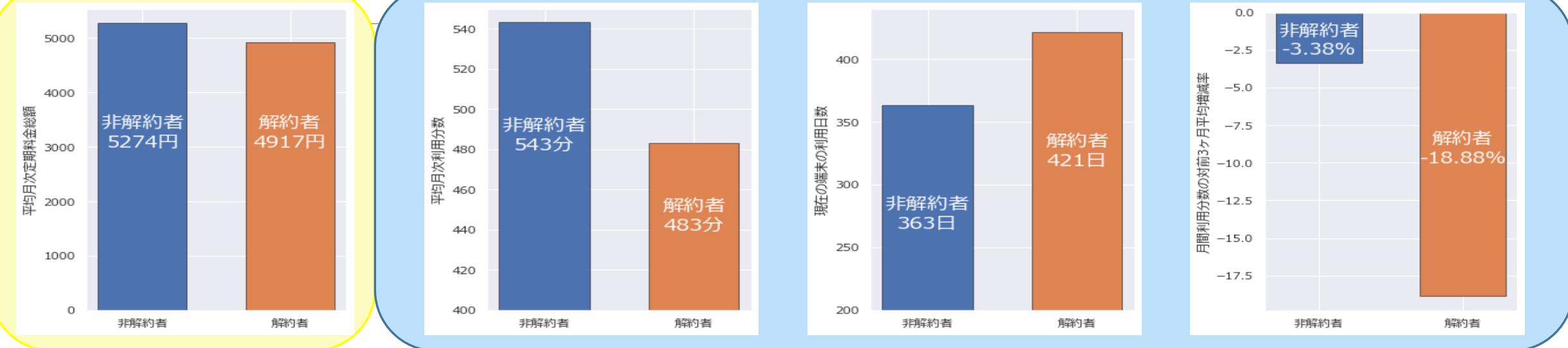
*1 総務省、「携帯電話の料金等に関する利用者の意識調査」、2019～2021年度にかけて計4回実施
https://www.soumu.go.jp/main_content/000752805.pdf

*2 「乗り換え理由」の計4回の平均、複数回答可

*3 「乗り換え検討理由」の計4回の平均、複数回答可

1. 御社のデータ分析から見る課題

解約者の分析

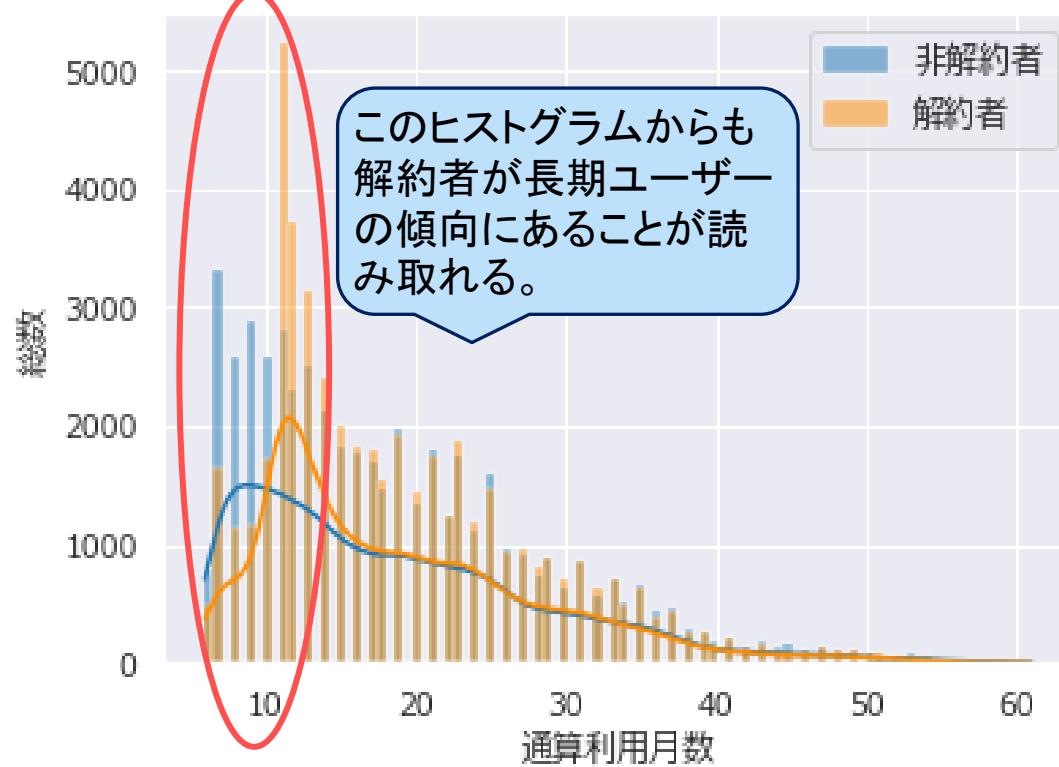


平均月次定期料金は非解約者の平均が5274円^{*4}、解約者の平均が4917円^{*4}と料金の差はほぼ見られませんでした。しかしながら、平均月次利用分数は解約者の方がおよそ60分ほど短く、現在の端末の利用日数の平均データから、解約者の方が58日ほど長く使っていることが分かりました。加えて、解約者は非解約者に比べ、利用時間の減少率が高い傾向にある^{*5}と言えます。

*4 [totmrc_Mean]のデータは\$であると判断し、1\$=110.38円(2021年7月26日 19:02 UTC)で換算

*5 右図「月間利用分数の対前3ヶ月平均増減率」、[change_mou]に外れ値が確認されたため、 $\mu \pm 2\sigma = (-566..., 538...)$ の範囲で分析を実施

解約者の分析



非解約者と比較した解約者の特徴

- ・利用時間が短い
- ・長期ユーザーの傾向
- ・利用時間の減少率が高い



解約の原因

▷長期ユーザーかつ利用時間が短いにもかかわらず、料金プランは定額のため、安くならないという不満。

⇒利用状況に合った料金プランでない可能性

解約抑止案

携帯電話各社の料金プラン*6	大容量 (無制限)	中容量 (20ギガまで)	小容量	※月額、税別、割引は含まず
docomo	ドコモ	6650円	2700円	3150~6150円
au	KDDI	6580円	2480円	3150~6150円
SoftBank	ソフトバンク	6580円	2480円	2980~4980円
Rakuten Mobile	楽天モバイル	2980円	1980円	無料~980円

総務省の携帯電話料金の引き下げ検討指示により、2021年3月から大手キャリア(NTTdocomo, au, SoftBank)を中心に、**中容量の格安新料金プランが提供開始されました。***6

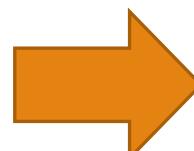
利用時間の短い(通信量の少ない)傾向にある解約者に対して、御社も同様の格安新料金プランを訴求できれば、**解約抑止につながると考えられます。**

*6 産経新聞「選択肢増える携帯料金プラン」 2021/3/28 <https://www.sankei.com/article/20210328-C2NHWRLYJPHDBVZ5KFCY3PL4/>

AIによる
解約者予測



解約予測者に対する
料金診断・格安新料金
プランの提案



解約者数の減少
↓
収益の増加

AIによる解約者予測と解約抑止案

御社の新料金プランの設定

大手キャリア3社(NTTdocomo, au, SoftBank)の中容量新料金プランをもとに仮定いたしました。

- ▶月額料金 : 2,553円(3社の平均, 税別)
- ▶データ量 : 20GB

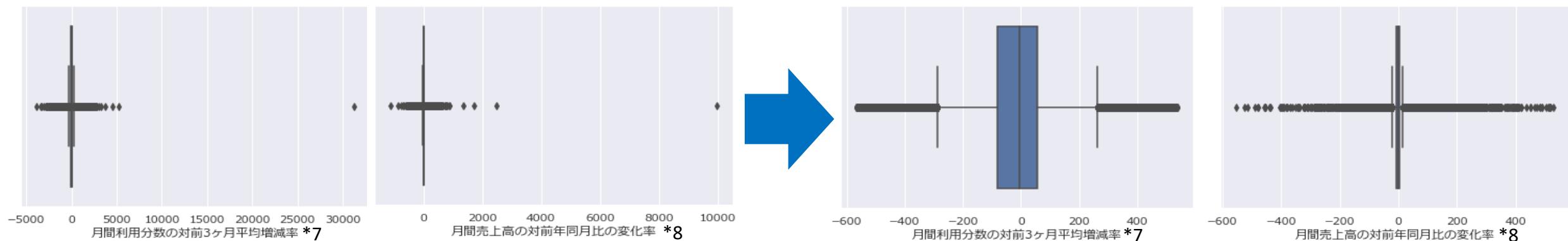
ご提案内容

- ▷弊社開発の**機械学習モデルを用いて解約者の予測**を行い、解約予測された顧客に対して、**料金診断を無料で実施**します。
- ▷この診断により、現在の利用状況に見合ったプランでないと判断された場合、御社の**格安新料金プランを訴求**いたします。

前処理

1. 外れ値の除去

「月間利用分数の対前3ヶ月平均増減率*₇」と「過去3ヶ月平均の月間売上高の変化率*₈」に記入ミスと思われる外れ値が確認されたため、それぞれ $\mu \pm 2\sigma$ 、 $\mu \pm 4\sigma$ の範囲を用いました。



2. 欠損値の処理

欠損が1/3(33333件)以上であるデータ*₉は削除しました。また数値データの欠損は各データにおける解約者平均と非解約者平均で補完しました。

*₇ [change_mou], 下図参照, ◆マークが外れ値を示す

*₈ [change_rev], 下図参照, ◆マークが外れ値を示す

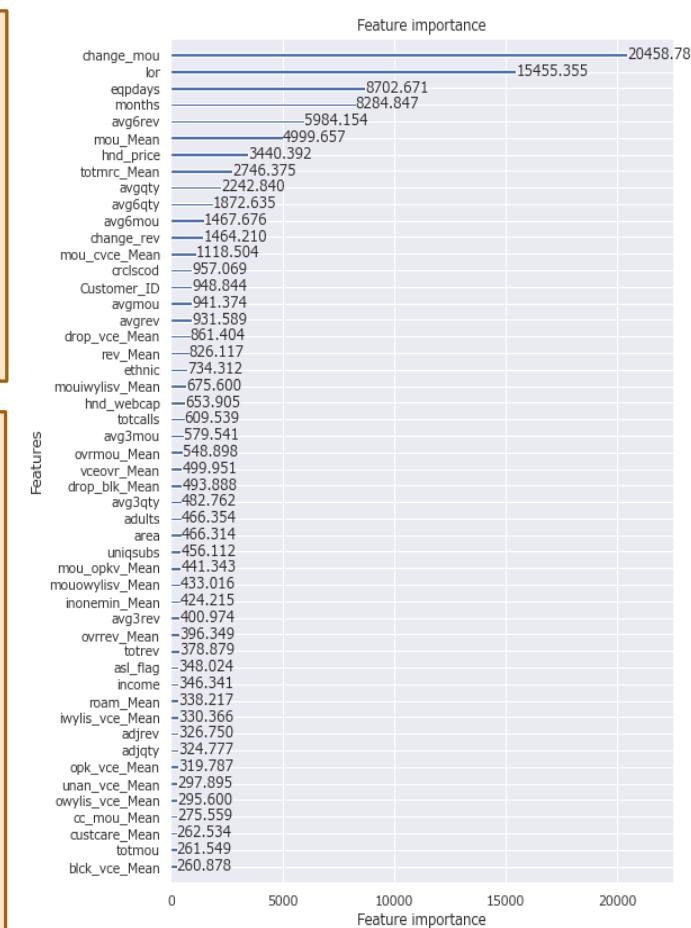
*₉ [ownrent], [numbcars], [Hhstatin], [dwllsize]

各データの重要度分析

3. 重要度分析

欠損のあるカテゴリカルデータを有する顧客のみ削除し、LightGBM^{*10}という機械学習モデルを利用して**各データの重要度の分析**を行いました。右図は重要度の上位50種類のデータを示しています。

*10 Microsoft Corporation, 「Welcome to LightGBM's documentation!」, <https://lightgbm.readthedocs.io/en/latest/>



4. 機械学習に用いるデータの抽出

「2.欠損値の処理」後、この上位50種類に関して、10万人分の顧客情報を抽出し、欠損のあるカテゴリカルデータ^{*11}を削除して数値データに変換しました。その後、**計47種類の10万人分の顧客情報をもとに、複数の機械学習モデルを作成しました**^{*12}。

*11 [area], [ethnic], [hnd_webcap]

*12 機械学習のアルゴリズム(手法)は多種多様であり、ここでは「LightGBM」、「ランダムフォレスト」、「ロジスティック回帰」、「多層パーセプトロン」、「XGBoost^{*13}」を用いて、比較・検討を行いました。

*13 「XGBoost Documentation」, <https://xgboost.readthedocs.io/en/latest/>

機械学習モデルの作成



- ① 7万人分の顧客情報(過去の顧客情報)をもとに、AIが学習します。
- ② 学習したAIに新規顧客情報を与えます。なお、ここでは予測の精度を測定するため、3万人分の顧客情報(過去の顧客情報)を与えました。
- ③ 学習したデータをもとに、AIが予測した結果(0 or 1)を出力します。

モデルの比較と解約者予測

▷「AUC^{*14}」という評価指標を用いて、計6個の機械学習モデルの比較検討を行いました。

▷右図より、「LightGBM」はAUC=0.8683で、最も精度が高かったため、このモデルを解約者の予測に使用しました。

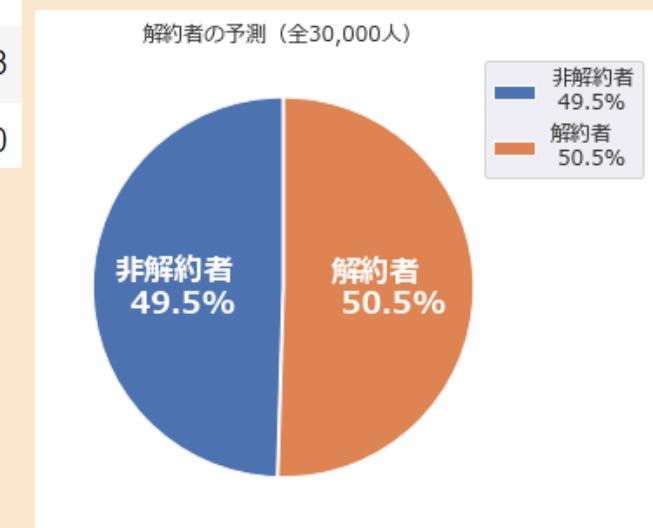
▷3万人のデータから、解約者と非解約者の割合を予測したところ、以下の結果が得られました。

解約者数：15,143人（50.5 %）

非解約者数：14,857人（49.5 %）

*14 0～1の値を取り、1にどれほど近いかでモデルの良し悪しを判断する。

機械学習モデル	AUC
LightGBM	0.8683
XGBoost	0.8646
ランダムフォレスト	0.8591
決定木	0.8389
ロジスティック回帰	0.6178
多層パーセプトロン	0.5000



訴求効果の仮定

解約予測者に対する訴求効果の仮定

▷解約予測者に対して、**無料の料金診断と格安新料金プラン(月額料金：2,553円, データ量：20GB)**の訴求を行い、新料金プランに変更するものと仮定する。



1. 解約予測者の**平均月次定期料金総額を2,553円^{*15}とする。**

*15 1\$=110.38円(2021年7月26日 19:02 UTC)より、解約予測者の[totmrc_Mean]を23.13\$とした。

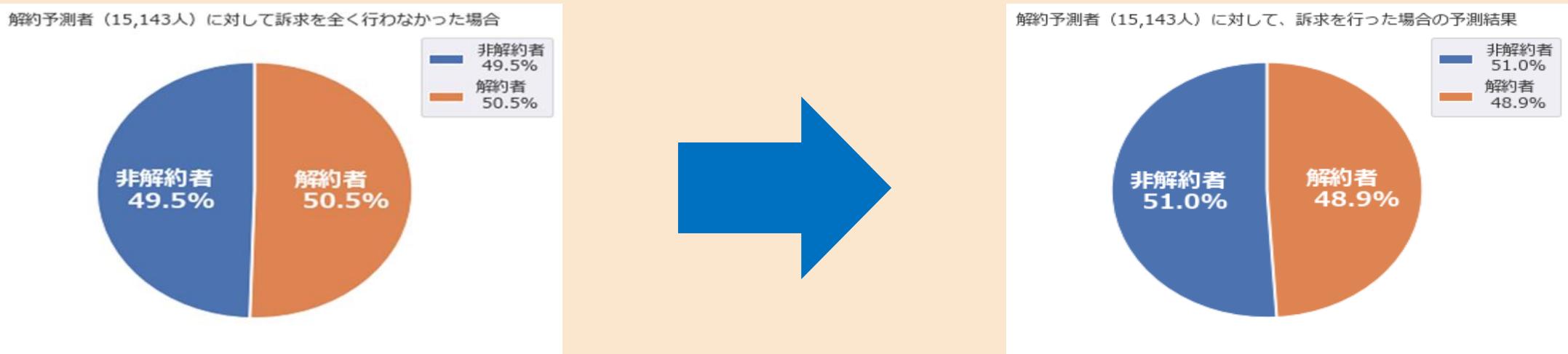
2. 解約予測者の**平均月次収益を2,180円分^{*16}減らす。**

*16 解約予測者の平均月次定期料金総額(42.88\$)を23.13\$したことによる減額、 $42.88\$ - 23.13\$ = 19.75\$ = 2,180$ 円



訴求効果の仮定を反映し、弊社の機械学習モデルを用いて、**再度解約者の予測**を行います。

事業の導入効果



解約阻止者数・解約阻止による収益・訴求しなかった場合の損失額

$30,000 \text{人} \times (0.505 - 0.489) \text{(解約阻止割合)} = \text{480人(解約阻止者数)}$

480人(解約阻止者数) $\times 2,553 \text{円(新料金プラン)} = \text{1,225,440円(解約阻止による収益)}$

480人(解約阻止者数) $\times 6,425 \text{円}^{*18} \text{(解約者平均月次請求額)} = \text{3,084,000円(訴求せずに解約してしまった場合の損失額}^{*20})$

➡ 30,000人規模で、解約損失額を39.7%^{*19}削減し、収益の増加に貢献！

^{*18} [rev_Mean]の解約者平均を1\$=110.38円(2021年7月26日 19:02 UTC)で換算 ^{*19} $(1,225,440 \text{円} / 3,084,000 \text{円}) \times 100\% = 39.7\%$ ^{*20} 訴求により、解約が阻止可能な者を対象