

貴社課題分析と 機械学習モデルのご提案

顧客解約率に対するアプローチ

2023年1月31日

本提案の目的

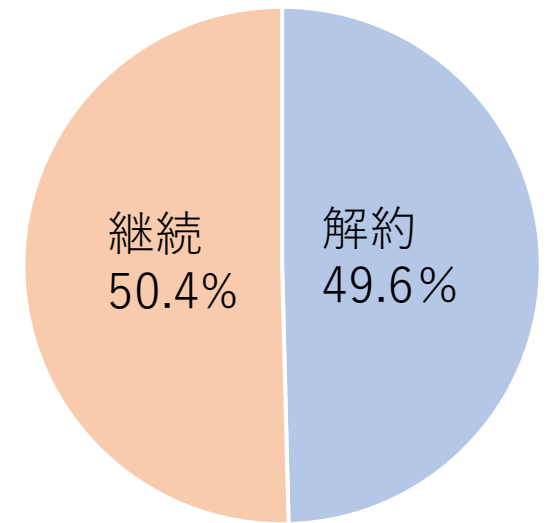
貴社の、データを分析した結果、churn(：観測日から1ヶ月～2ヶ月の間に解約した顧客)の割合は、以下のようになっている。

継続利用者：50.4%(50438人)

解約者：49.6%(49562人)

携帯電話業界では、一般的に年間解約率は1-2%程度※とされているため、50%近くであることは異常である。

よって、今回は解約率を下げることを目的とする。



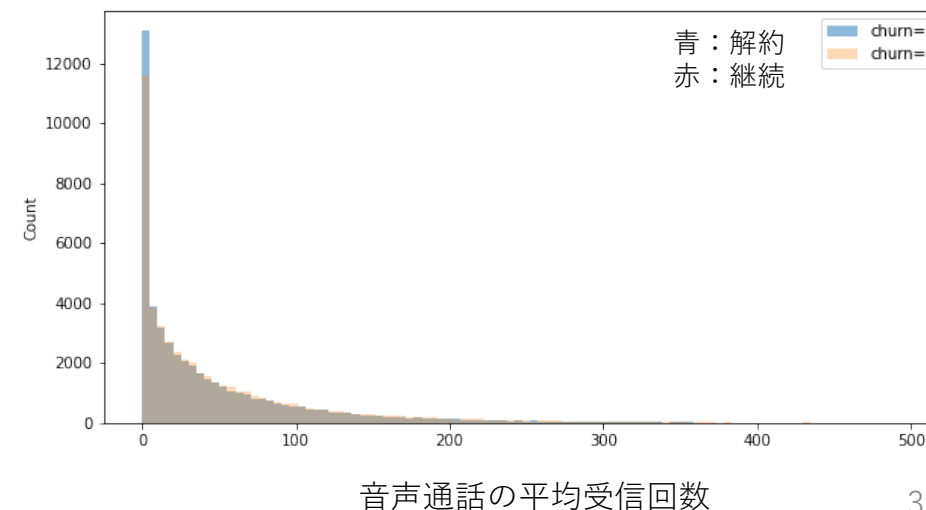
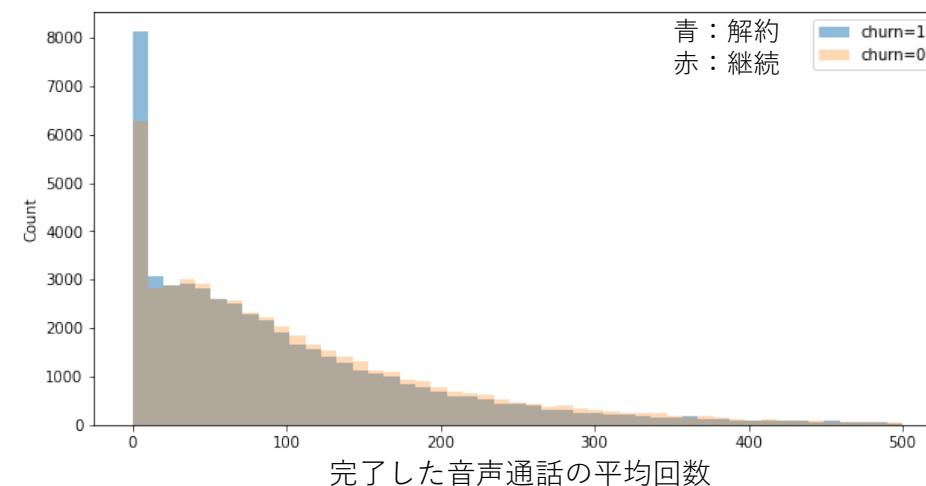
※総務省 各種資料 令和2年8月27日

顧客通信状況データの分析

顧客通信状況のデータを解約者と継続者に分けて分析した結果、下記の特徴量において傾向が見つかった。

- comp_vce_Mean : 完了した音声通話の平均回数
- change_mou : 月間使用時間の対前3ヶ月平均の変化率
- recv_vce_Mean : 音声通話の平均受信回数
- inonemin_Mean : 1分未満の着信通話数の平均値
- peak_vce_Mean : ピーク時の平均音声通話回数
- mou_peav_Mean : ピーク時の音声通話の平均使用時間
- attempt_Mean : 平均通話試行回数
- complete_Mean : 平均通話完了数

通話をする回数や時間が少なくなるほど、解約する傾向にある。



顧客属性等データと解約率の紐づけ

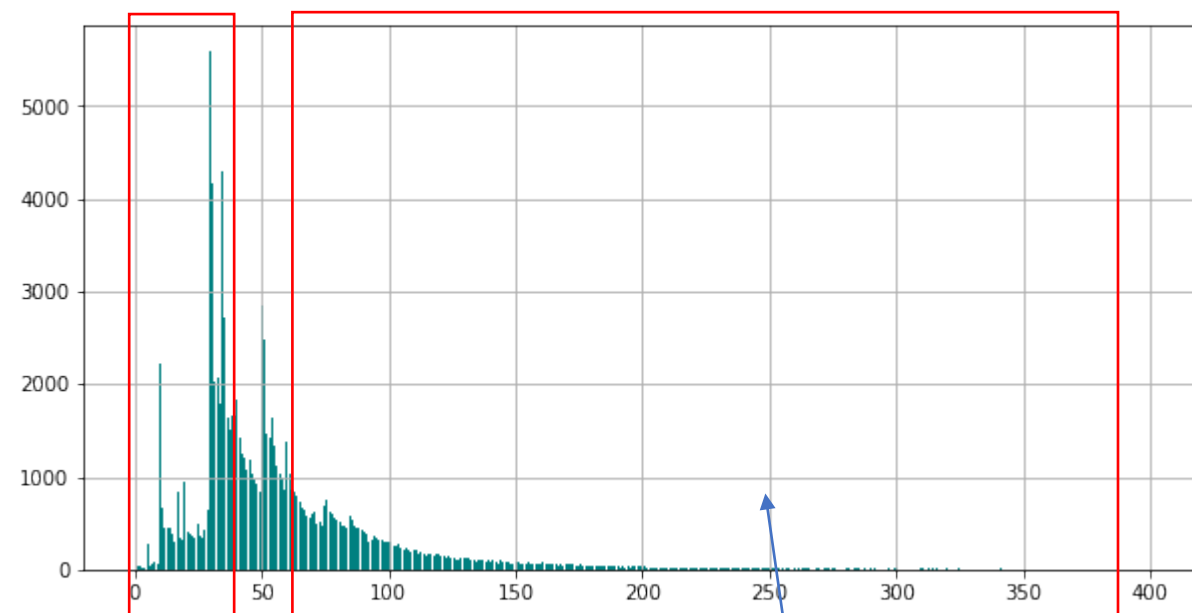
顧客通信状況データと顧客属性等データを結びつけるために

過去3ヵ月間の平均月間収益が32pt以下の顧客を電話の利用率が低い層としてrev_miniのデータとした。

32pt以上の顧客を電話の利用率が高い層としてランダムにサンプリングし、rev_randとした。

データ数はそれぞれ21579件である。

rev_miniデータ→利用率低い→解約
rev_randデータ→利用率高い→継続



過去3ヵ月間の平均月間収益

rev_mini
32pt以下、全体の2割

||
解約

rev_rand
32pt以上でランダムに
サンプリング

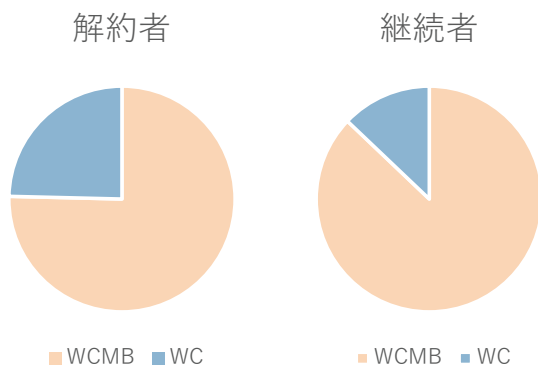
||
継続

解約率と顧客属性等データの傾向 1

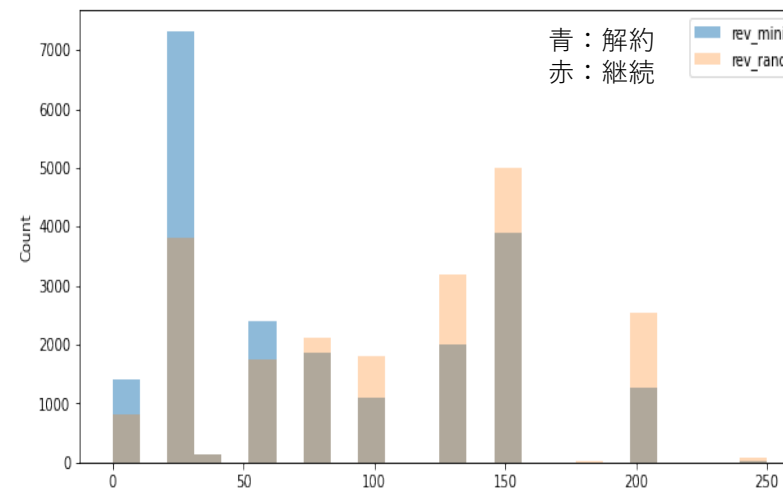
解約する傾向を携帯電話端末に注目して分析した結果が以下のようなものである。

- 携帯電話価格が低価格
- 端末を長期間使っている
- スマートフォン以外である

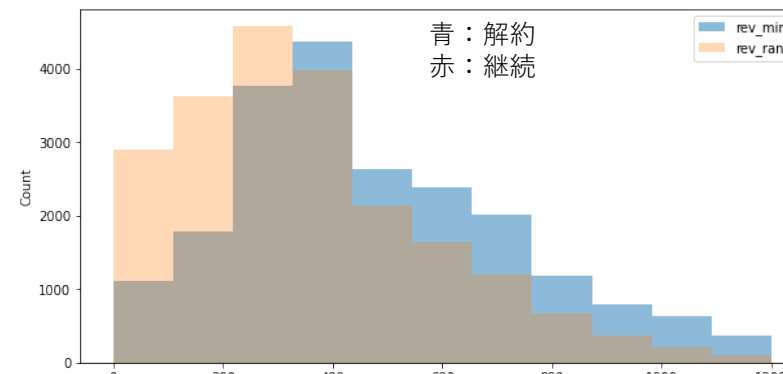
以上の結果から、**携帯電話を乗り換えるタイミング**で解約をするといえる



※WCMB:スマホにおけるWebサービス
WC:スマホ以外(PCなど)のWebサービス



現在の携帯電話機の価格



現在の端末の使用日数

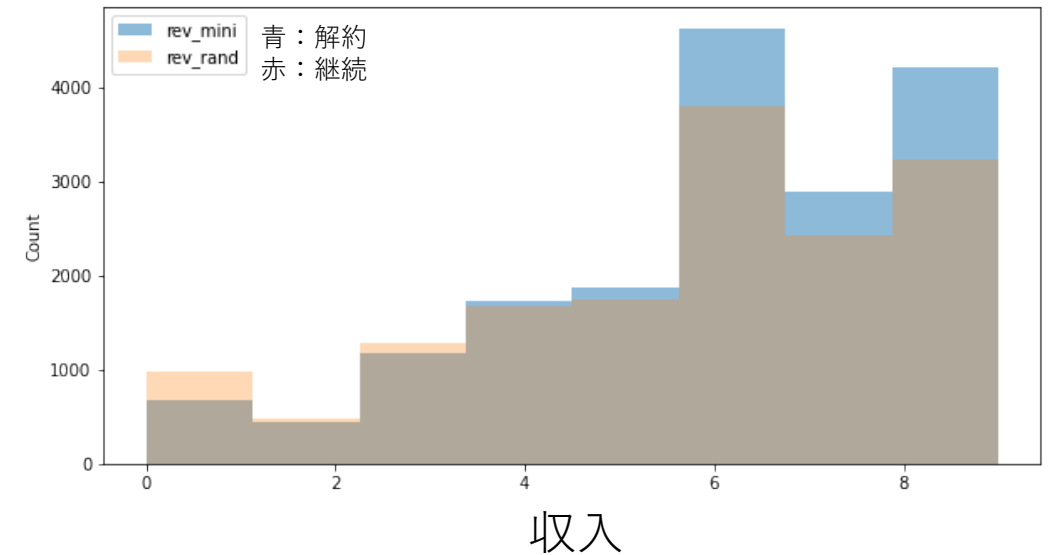
解約率と顧客属性等データの傾向 2

低収入者よりも高収入者のほうが平均月間収益が少ない(つまり解約傾向にある)というデータが見れた。

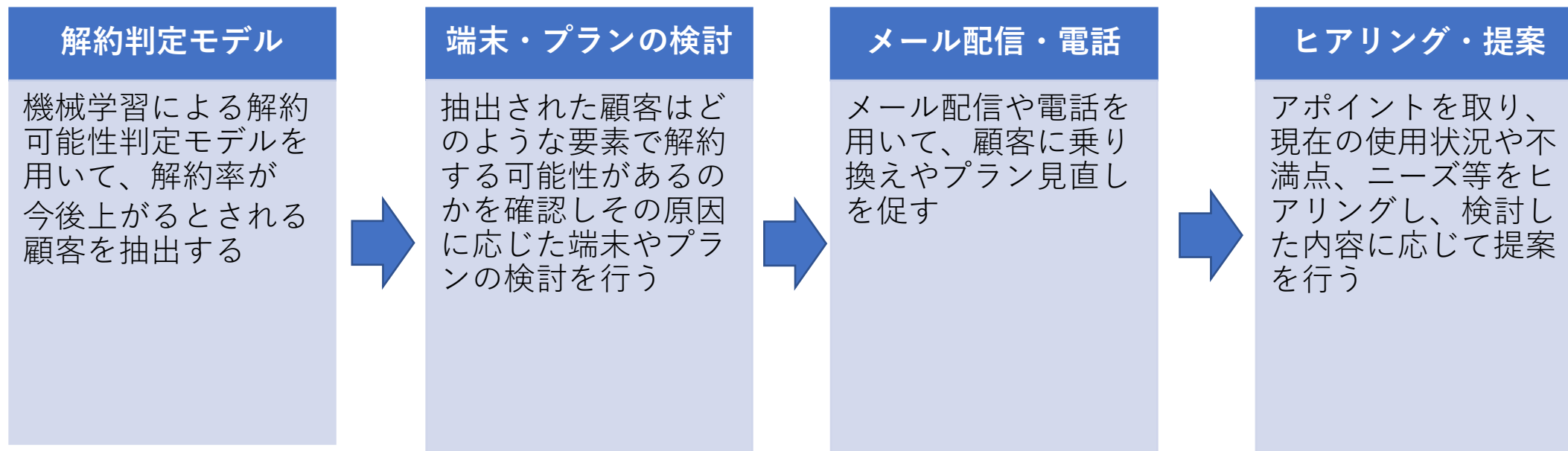
自分の気に入ったところ以外はできる限り儉約をするという価値観が、高所得者にはより強くあると考えられる。

今回、電話の使用頻度が少ない人が解約する傾向にあると分かったため、使用しないにも関わらず、所持し続けることを無駄であり、その金額を支払うほどの価値はないと判断したと読み取れる。

→ **高所得者は**名義者本人は使わずとも、世帯者に**高額プランで契約する可能性があるため**、キープポイントとなる。



ご提案サマリー



以上の実施により、他社に顧客を流すことなく、
解約予測者を自社内の変更でとどめることができる

解約可能性が高い顧客の主な特徴と改善案に関して

◆低価格端末・スマートフォン以外を利用している人

→現在使用中の端末の機能不足や時代の流れにより、

機種変更を望んでいると考え、キャンペーン情報や最新機種の情報を発信

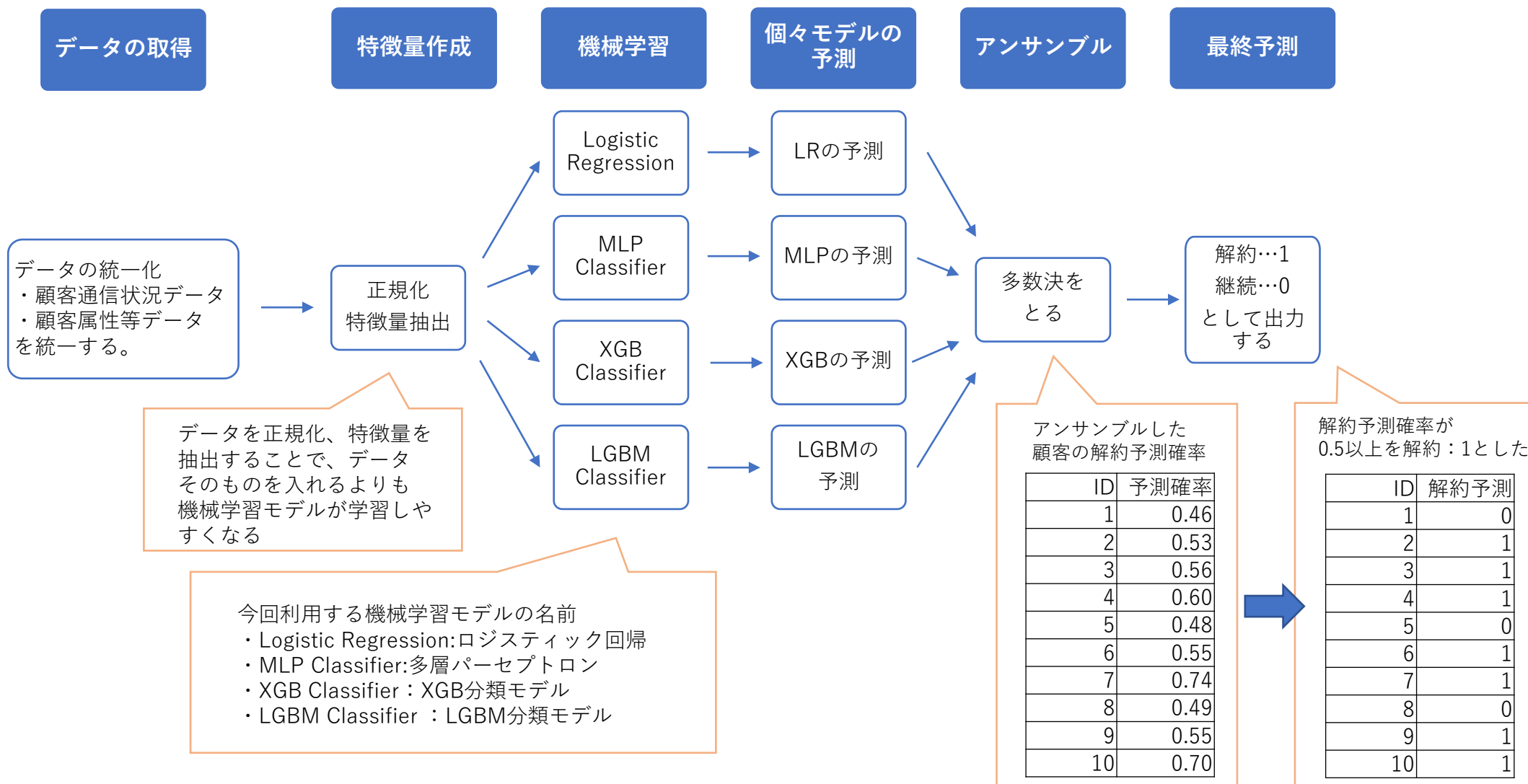
◆電話使用頻度が極めて少ない人

→所持していても使う機会がないので解約するというロジックだと捉え、

使用した分料金が上乗せされるプランや、利用制限が低く設定されている分料金も抑えられるプランなどへの変更を提案する

※前述のように、高収入者ほど使用状況を考えて解約判断をするのでアンテナを張るべきである。

解約可能性予測までのフロー（技術資料）



解約可能性予測までのフロー（技術資料）

データの取得	特徴量の作成	機械学習	個々モデルの予測	アンサンブル	最終予測
データを統一、一本化し欠損値を削除する。	データを正規化し、モデルが学習しやすくとともに、解約に相関の強い特徴量を抽出することで、解約者の判定精度を上げる。	ロジスティック回帰、多層パーセプトロン、XGB分類モデル、LGBM分類モデルの4つのモデルでそれぞれ学習させる。	4つのモデルそれぞれで顧客の解約確率がわかる。	4つのモデルがだした確率をすべて足し、4で割り平均をとることで、解約可能性の誤判定を防ぐ。	アンサンブルした表は確率で表されているため、確率が50%を超えるものを「解約：1」、50%以下を「継続：0」として最終予測の表に書き出す。※

※ここで、確率50%を基準としたが、それ以外の確立にすることも可能である。

解約可能性判定モデルの実装

顧客通信状況データのみでモデルを作成した結果以下のスコア（判定率）がでた。

Train Score: 0.67

Valid Score: 0.65

つまり、65%の精度で解約判定が可能であるということだ。

これを顧客属性等データと紐づけ、データを一本化することによりスコアは70%から80%まで改善するとみて、モデル導入後の収益を算出した。また、解約可能性が高い顧客へアプローチして、顧客として残る割合を、2割, 3割, 4割とした。

見込まれる効果、収益

avg6rev(：過去6か月の平均月間収益)の全データの平均値をとると、58.7ドルと分かった。つまり、**一人の解約を防ぐことができれば、58ドルの収益**が見込める。

1ドル=130.15円とすると、58ドル=7500円である。

前ページと上記のことから

判定モデルの予測精度：70%, 75%, 80%

顧客として残る割合：2割, 3割, 4割

1人当たりの収益：7500円

この前提の下で、見込まれる収益を次のページに算出した。

見込まれる効果、収益

解約者に対してアプローチをすることで、
現収益よりも **少なくとも5300万円の収益**
を見込めることがわかる。

今後、**モデルを導入し、データの一本化**
と蓄積により 予測精度が向上することで、
顧客の解約可能性と原因が分かり、解約者
を減らし、**さらに収益の増加が見込める。**

判定モデルの予測精度：70%, 75%, 80%
顧客として残る割合：2割, 3割, 4割
1人当たりの収益：7500円

モデルの予測精度	アプローチして 継続する人の割合	見込まれる収益
70%(35,000人)	2割(7000人)	5300万円
	3割(10500人)	7900万円
	4割(14000人)	1億500万円
75%(37,500人)	2割(7500人)	5700万円
	3割(11250人)	8500万円
	4割(15000人)	1億1300万円
80%(40,000人)	2割(8000人)	6000万円
	3割(12000人)	9100万円
	4割(16000人)	1億2100万円