

修士論文概要書

Master's Thesis Summary

Date of submission: 01/11/2022 (MM/DD/YYYY)

専攻名 (専門分野) Department	経営システム 工学専攻	氏名 Name	大久保 亮吾 Ryogo Okubo	指導 教員 Advisor	後藤 正幸 印 Seal
研究指導名 Research guidance	情報数理応用研究	学籍番号 Student ID number	CD 5221C011-1		
研究題目 Title	クレジットカードへの予測切り替え期間を用いたユーザ分析モデルに関する研究 A Study on an Analytical Model of Users' Switching Possibilities by Using Predicted Time to Credit Card Users				

1. はじめに

近年、独自のクレジット機能付きポイントカードを発行している小売企業も多くなっている。クレジット機能付きポイントカードはクレジットカードとポイントカードの機能の二面性を有しており、加盟店内外で使用することで、顧客は購買金額に応じたポイントを獲得し、ポイントに応じた様々な特典を得ることができる。クレジット機能付きポイントカードを保有するユーザ（以下、クレジットユーザ）は、ポイント専用カードを保有するユーザ（以下、ポイントユーザ）に比べて獲得できるポイントが多く、そのポイントをきっかけとした加盟店での継続的な利用が見込まれる。そのため、企業側の視点では、顧客と長期的に良好な関係を築きたいという理由から、クレジットユーザを増加させることが望ましい。

その一方で、一般のユーザにとってポイントカード作成は心理的障壁が低いのにに対して、クレジットカード作成の心理的障壁は高いため、クレジットカードの新規申込数を直接的に増やすことは容易ではない。そのため、企業はまずポイントユーザを獲得した後、クレジットユーザへ切り替えを促すことで効果的にクレジットユーザを増やす取り組みを行っている。しかしながら、多くの顧客はDM等で切り替えを促してもクレジットユーザとならないことが経験的に明らかになっている。これらの多数のユーザの中から切り替えの難易度が低いユーザを識別できれば、効果的な施策の計画の一助となると考えられる。

そこで、企業は切り替え難易度に応じた施策を講じるために、ポイントユーザがクレジットユーザへ切り替えるための難易度の見積もりを行う。ここで、クレジットユーザの中には、短期間で切り替えを行ったユーザから切り替えまで長期間を要したユーザまで幅広く存在し、長期間にわたりポイントユーザであったユーザは、それ以降も容易には切り替えの行動を行わないという傾向も示されている。そこで、早く切り替える人の方がクレジットカードへの心理的障壁が低く、遅い人の方が心理的障壁が高いとみなすことができると考え、本研究では、ポイントユーザからクレジットユーザへの切り替えまでの期間（以下、切り替え期

間）の分析を行う。このように、クレジットユーザの利用履歴から切り替え期間を予測可能であれば、ポイントユーザの切り替え難易度が推定可能となり、ターゲットごとに切り替え難易度に基づいた施策を打つことができる。

一方で、ポイントユーザの大半は、ポイントユーザのまま、クレジットユーザとならないユーザであり、切り替え期間が観測されていない。そのため、全ユーザのデータを学習データとして切替期間を予測する機械学習ベースの回帰モデルを構築することができない。そこで、本研究では分類モデルと回帰モデルを併用した切り替え期間の予測モデルを提案する。具体的には、分類器 [1] を用いてクレジットユーザへ切り替えるユーザと切り替えないユーザに分類することで、潜在的にクレジットユーザになりやすいユーザを抽出した上で、切り替えるであろうユーザに対しては回帰モデルを適用し切り替え期間の予測を行う。これにより、高い精度でポイントユーザの切り替え期間を推定し、ポイントユーザの切り替え難易度が分析可能となる。本研究では、実際の購買データを用いた分析を行い、提案手法の有効性を示す。

2. 準備

2.1. ポイントユーザとクレジットユーザ

本研究では対象事例として、小田急電鉄株式会社が発行する小田急ポイントカード（以下、OP カード）を用い、その利用履歴データの分析を行う。OP カードには、ポイント専用カードである OP ポイント専用カードとクレジット機能付きポイントカードである OP クレジットの 2 種類がある。

本研究では、ユーザの利用傾向の違いからポイントユーザからクレジットユーザになるまでの切り替え期間を予測するモデルを構築する。そこで、ある一定期間の間に入会し、一定数の利用があるユーザを対象とする。その際、クレジットに切り替えるユーザと切り替えないユーザでは、その数に大きな差異があることに注意する必要がある。実際、2017 年 1 月 1 日から 2020 年 6 月 10 日にポイントカードに入会し、入会后 3 回以上利用しているユーザは、クレジットユーザが 4,857 件、ポイントユーザが 138,100 件と

なっており、ユーザ数が大きく偏っていることが分かる。

2.2. 関連研究

小田急電鉄株式会社が発行する OP カードの利用履歴データを用いた分析として清水らの研究 [2] や平野らの研究 [3] が挙げられる。清水らは、クレジットユーザの優良化を目的とし、カード所持目的分析モデルと購買店舗分析モデルから得られた分析結果を用いてクロス分析を行い、カード利用顧客の増加につながるような施策の立案を行っている [2]。

一方、平野らの研究では、ポイントユーザからクレジットユーザへの切り替え促進を目的としている [3]。クレジットユーザと類似の利用傾向を持つポイントユーザはクレジットユーザへ切り替わる可能性が高いという仮説に基づいた分析手法を提案した。これにより、クレジットユーザの利用傾向をクラスタリングし、転移学習の枠組みを援用することで、ポイントユーザがどのようなクレジットユーザへ成長する可能性があるかを示し、切り替えを促進する有用な施策の検討を可能としている [3]。

平野らは、クレジットユーザへ切り替わるか否かという点に着目していた。一方で、長期間にわたりポイントユーザであった顧客は、それ以降も容易には切り替えの行動を行わないという傾向も示されている。そこで、本研究では、ポイントカードを発行してからクレジットユーザに切り替えるまでの切り替え期間に着目する。具体的に、ポイントユーザの利用履歴データをもとに、クレジットユーザへの切り替え期間を予測するモデルを構築し、分析を行う。

3. 提案手法

3.1. 概要

本研究では、ポイントユーザがクレジットユーザへ切り替えるための切り替え難易度を評価する手法の提案を目的とし、ポイントカードに入会したユーザが、何日後にクレジットカードに切り替えるのか（切り替え期間）を予測するモデルを構築する。このとき、切り替え期間の値を保持しているクレジットユーザのみのデータを使って回帰モデルを構築した場合、以下の 2 点の問題が生じる。

1. ポイントユーザのままクレジットユーザにならなかった大多数のユーザのデータを利用しないことになる。
2. 実際に予測する対象は入会してから一定期間のポイントユーザであるが、予測モデルが「切り替えたユーザ」のみを学習したものである場合、選択バイアスが生じてしまう。

ここで、ポイント会員に入会した全てのユーザを用いて回帰モデルを学習しようとする、「ある時点で、まだクレジットカードに切り替えていないユーザ」が大多数であり、これらのユーザの切り替え日数が観測されていない打ち切りデータとして扱うことになる。これに対し、打ち切りデータに対する回帰モデルのパラメータ推定手法である

Tobit モデルを用いることが考えられる。しかし、学習データの中に切り替える可能性がないユーザが多く含まれることから予測精度が悪化してしまう。

これらの問題点を解決するために、回帰モデルの前に分類モデルを導入することで、「回帰モデルで切り替え時期を予測すべきユーザ」と「切り替える可能性がないユーザ」の 2 つに分類を行うことを考える。「回帰モデルで切り替え時期を予測すべきユーザ」であると分類されたユーザに対してのみ、回帰モデルを用いて切り替え期間の予測値を求める。「切り替える可能性がないユーザ」であると分類されたユーザに対しては、一律に設定した最大値を切り替え期間の予測値とする。この方法により、全データをモデルの学習に活用し、回帰モデルで予測すべきデータのみに対し、回帰モデルを適用することができる。

3.2. 提案モデル

提案モデルでは分類モデルと回帰モデルを併用した予測モデルを構築する。分類モデルでは、切り替え時期を予測すべきユーザであるユーザ群の抽出を行うことでユーザを絞り、回帰モデルではそのユーザ群に対して切替期間の予測を行う。

3.2.1. 分類と回帰を併用した予測モデル

Step1) 全ユーザを対象に分類器を用いて、「ポイントユーザに留まると予測されるユーザ」と「クレジットユーザに切り替わると予測されるクレジットユーザ」を分類する。

Step2) Step1 で分類された「クレジットユーザに切り替わると予測されるクレジットユーザ」に対して回帰モデルにより切り替え期間の予測を行う。Step1 で「ポイントユーザに留まると予測されるユーザ」に対しては、事前に設定された最大値を予測値とする。

3.2.2. 提案モデルの学習方法

Step1) ポイントユーザに対し、クレジットユーザの数との同数になるようアンダーサンプリングを行う。

Step2) ポイントユーザとクレジットユーザの学習データから機械学習の分類器を用いて「ポイントユーザに留まると予測されるユーザ」と「クレジットユーザに切り替わると予測されるクレジットユーザ」に分類する分類モデルを学習する。

Step3) 切り替え期間の予測のために、「クレジットユーザのうち機械学習の分類器でクレジットユーザと判定されたユーザ」と対象ユーザを絞り、回帰モデルを用いて切り替え期間の予測モデルの学習を行う。

これにより、切り替え期間を予測すべきユーザのみを対象に回帰モデルを学習することで、対象を適切に抽出した予測モデルの構築を行うことができる。

3.2.3. 切り替え期間の予測に関する分析

妥当性が確認された分類と回帰を併用した予測モデルに対して、実際のポイントユーザを対象に分類器で「今後クレジットユーザに切り替わると予測されるポイントユーザ」に対して切り替え期間の予測を行うことで、ポイントユーザの切り替え難易度の指標化を行う。

4. 実データ分析

小田急電鉄株式会社が保有する OP カードの利用履歴データに対して、提案手法を適用し、分析を行うことで提案手法の有用性を示す。

4.1. 対象データ

小田急電鉄株式会社が発行する OP カードの利用履歴データである。2017年1月1日から2020年6月10日にポイントユーザになり、入会後3回以上利用しているユーザを対象とし、2020年6月10日時点でクレジットユーザに切り替わった4,857件、ポイントカードに留まったユーザ138,100件を用いた。また、ポイントユーザの利用期間を2017年1月1日以降でOPカードを保持したポイントユーザの2020年6月10日までの経過日で示し、最大日数は1,257日とする。本分析で用いる変数を表1に示す。

表 1: モデルに使用した特徴量

	特徴量名	定義
説明変数	利用金額平均	利用金額の合計値/利用回数
	利用回数	期間内にカードを提示した回数
	付与 P 合計	期間内に付与されたポイント数
	還元 P 合計	期間内のポイント支払いの合計
	付与 P 回数	期間内にポイント付与された回数
	還元 P 回数	ポイント支払いを行った回数
	利用業種数	期間内に利用した業種数
	業種別利用割合	ある業種の利用回数/利用回数
目的変数	切り替え期間	クレジットユーザに切り替わるまでの期間

4.2. 分析条件

分類モデルにおいてクレジットユーザ数が4,857件であることから、ポイントユーザ138,100件をクレジットユーザと同数の4,857件にアンダーサンプリングを行う。そして、ポイントカードに留まるユーザとクレジットカードに切り替えるユーザの分類には、SVM(Support Vector Machine)を適用する。SVMは2クラス問題の線形識別関数構成法であり、マージン最大化という分類のための明確な基準を保持するため、正例をクレジットユーザ、負例をポイントユーザとして2値分類器の構築に用いる。回帰モデル構築のための学習データとしては、「クレジットユーザ全体を学習する方法」と「クレジットユーザの中でSVMにより切り替えると検出されたユーザ集合を学習データとする方法(学習データをSVMで絞り込む方法)」の2パターンを実験する。そして、切り替え期間を予測するための回帰モデルには、線形重回帰分析を用いる。また、回帰モデルを構築するために「分類モデルで抽出したユーザのみ」を学習に用いる提案手法の有効性を検証するため、「ク

レジットユーザ全体」を学習データとする方法を比較手法として評価実験を行う。それぞれに対して、テストデータに対して線形回帰モデルを適用し、予測精度の比較を行う。また評価指標としては、テストデータのクレジットユーザのうちSVMでクレジットユーザと判定されたユーザに対するMAE(Mean Absolute Error)による誤差を用いる。提案手法の評価として、ポイントユーザとクレジットユーザが同数になるように、5分割交差検証を用い、学習データとテストデータに分割し評価する。

4.3. 分析結果

4.3.1. 提案手法と比較手法の予測精度の比較

提案手法と比較手法の2つの学習に対する学習データ、テストデータにおけるMAEを図1に示す。横軸は、SVMの分類境界を表し、0が正例と負例の閾値となっており、負の方向に行くほどクレジットユーザの再現率は向上する。

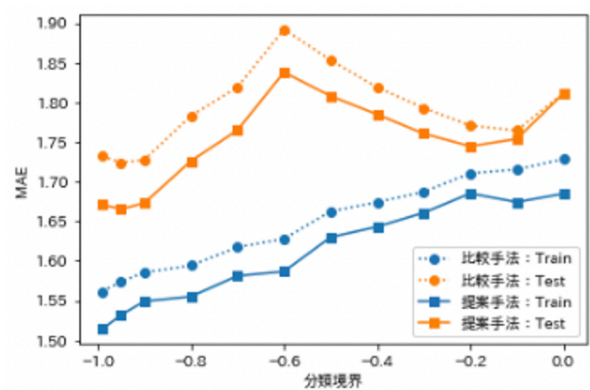


図 1: 分類境界の変化に対する MAE の推移

図1から学習データを絞った提案手法の方がテストデータに対するMAEが良好化していることがわかる。分類器で切り替えやすいとされたユーザのみを回帰モデルの学習データとした方が精度が向上していることが確認できる。

また、全ユーザへの予測と提案手法(SVMで絞り込んだユーザへの予測)の学習データとテストデータに対するMAEの比較結果を図2に示す。点線の比較手法は、クレジットユーザ全体を学習データとし、予測はテストデータの全クレジットユーザに対して行った結果を示している。また、実線の提案手法は、SVMでクレジットユーザと判断されたクレジットユーザのみを学習データとし、予測は分類境界の閾値以上のデータに絞り込んだ場合の結果である。

図2より、実線で表される提案手法におけるMAEが学習データ、テストデータ共に低くなっており、推薦精度が向上していることがわかる。分類モデルによりユーザを抽出することで精度が向上したといえる。

4.3.2. 回帰モデルのみを用いた場合との推薦精度の比較

表2で表されるように、各モデルにおけるテストデータに対するMAEを比較すると、比較手法よりも提案手法の

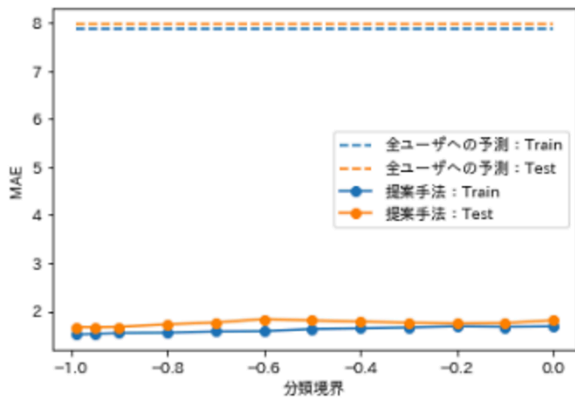


図 2: 全クレジットユーザを学習するモデルと提案手法の比較

方が MAE の値が低く、推薦精度が向上した。

表 2: 各モデルのテストデータに対する MAE

手法	MAE
線形重回帰分析	7.89
Random Forests 回帰	7.23
Tobit 回帰	7.33×10^4
提案手法	1.53

4.3.3. 提案手法を用いたポイントユーザへの切替期間の分析

本研究で構築したモデルを用いてポイントユーザへの切り替え期間の予測を行った。学習データのクレジットユーザのうち分類器でクレジットユーザと分類されたユーザ 3,760 件に対して、切り替え期間の予測モデルを学習を行った。対象データは、ポイントユーザのアンダーサンプリング後のテストデータとし、分類境界の閾値は図 1 で表される提案手法のテストデータに対する MAE が最も低くなる -0.95 とした。ポイントユーザの利用期間と切り替え期間の予測値における関係を図 3 に示す。

点線より下にいるユーザは、切り替え期間の予測値が実際の利用期間よりも短くなっており、本来クレジットユーザになっていると予測されているにも関わらず、切り替えを行っていない。このことから、OP カードの利用履歴データには表れない外乱的な要因によって予測不可能な要因で切替が行われていないユーザが一定数存在することが考えられる。

点線より上にいるユーザは、切り替え期間の予測値が実際の利用期間よりも長くなっており、今後切り替えのタイミングが来る可能性があるということが分かる。このように、ポイントユーザ毎の切り替え日数を考慮して DM などのマーケティング施策を打つタイミングを変えることで、切り替え促進を行うことができる。

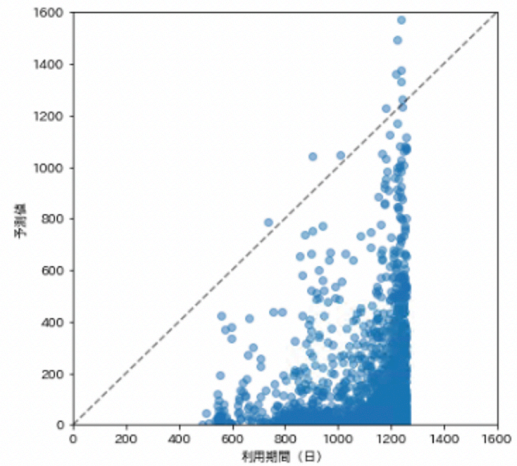


図 3: ポイントユーザの切り替え期間予測

5. 考察

提案手法により、分類モデルと回帰モデルを併用し、SVM を用いて切り替えを行いやすいユーザ群を絞ったのちに、線形回帰を行うことで切り替え期間の予測を行うことで予測精度が向上した。これにより、分類器を用いてクレジットカードへ切り替える可能性のあるユーザを抽出できていることを示した。その上で回帰モデルを用いることで、回帰モデルのみを用いた比較手法に比べて提案手法の優位性が示された。

6. 結論と今後の課題

本研究では、ポイントユーザがクレジットユーザへ切り替えるための難易度を評価する手法を提案することを目的とし、ポイントユーザがクレジットユーザに移行するまでの切り替え期間の予測モデルを構築した。実データ分析を通じ、従来の回帰モデルを用いた手法に比べて提案手法の推定精度が改善することを示した。

今後の課題として、切り替え期間の予測を行ったポイントユーザに対する詳細な分析を行い、今後のマーケティング施策の一助とすることが挙げられる。

謝辞

本研究では、小田急電鉄株式会社から提供を受けたユーザの利用履歴データを使用致しました。貴重なデータの提供に深く感謝致します。

参考文献

- [1] Michael Lee Bryant and Frederick D. Svm classifier applied to the mstar public data set. *Algorithms for Synthetic Aperture Radar Imagery*, Vol. 3721, pp. 355–360, 1999.
- [2] 清水良太郎, 坂元哲平, 山下遥, 後藤正幸. アンケートデータを考慮した ec サイトの購買履歴分析モデルの提案. *経営システム*, Vol. 27, No. 2, pp. 70–76, 2017.
- [3] 平野洋介, 楊添翔, 雲居玄道, 阿部永, 立花徹也, 後藤正幸. 顧客成長を促す施策立案のための特徴転移型クラスタリングモデル. *情報処理学会論文誌*, Vol. 62, No. 10, pp. 1704–1715, 2021.