

# 多層ニューラルネットワークを用いた マルチタスク学習による顧客購買行動予測

蓮本 恭輔<sup>1,a)</sup> 後藤 正幸<sup>2,b)</sup>

受付日 2021年7月16日, 採録日 2022年3月9日

**概要:** 顧客の購買行動を理解し、個々の顧客が将来どの程度自社の製品やサービスを利用してくれるのか、どれだけの利益をもたらしてくれるかを知ることは、企業にとって重要な課題である。近年、デジタルビジネスが拡大し、個々の顧客購買行動の情報が容易に収集可能になったことで、それらの情報をもとにした顧客購買行動の予測は顧客関係管理においてますます注目を集めている。顧客購買行動予測の重要なテーマとして、顧客生涯価値の予測と離反予測という2つのタスクがあるが、生涯価値の予測には離反時期の情報が必要なため、これらは強い関連性を有する。またこのような購買行動には、背後にある顧客ごとの潜在的な購買パターンや嗜好が強く影響していると考えられ、適切な予測にはそれらを特徴量としてモデルに取り込む必要がある。そのような潜在的な特徴をデータ駆動で抽出し、関連性のある複数のタスクを効率的に学習する仕組みとして、ニューラルネットワークを用いたマルチタスク学習がある。ニューラルネットワークによるマルチタスク学習では目的の異なるタスクを1つの学習器で同時に学習することで、タスクに共通に必要な潜在的な特徴を抽出し、全体のタスク精度を高められることが知られており、画像認識や文書生成においてさかんに研究が行われている。そこで本研究では多層ニューラルネットワークによるマルチタスク学習で、顧客生涯価値と離反の予測を行う手法を提案する。マルチタスク学習では各タスクの出力のモデル化とそれぞれのタスクの損失の重み統合が予測精度に大きな影響を与えるが、離反予測を顧客生涯価値予測の部分タスクとしてモデル化し、学習進捗に応じて重みをコントロールすることで、顧客生涯価値の予測精度を改善する手法を示し、実データを用いた実験を通して有効性を確認する。

**キーワード:** 顧客生涯価値 (CLV), 離反予測, マルチタスク学習, プラットフォームビジネス, ニューラルネットワーク, RFM 分析

## Multi-task Learning Using Multi-layer Neural Networks to Predict Future Customer Behavior

KYOSUKE HASUMOTO<sup>1,a)</sup> MASAYUKI GOTO<sup>2,b)</sup>

Received: July 16, 2021, Accepted: March 9, 2022

**Abstract:** It is important for modern businesses to understand the purchasing behavior of their customers and to understand how often each customer will use their services or products in the future and how much profit those customers will bring to the businesses. With the rapid expansion of digital business in the recent years, it becomes easier to collect information on individual customer purchasing behavior, leading to the increased attention on leveraging such information to predict future customer behavior in the field of customer relationship management. Customer lifetime value and customer churn are the key concerns for the prediction of future customer behavior. Since the prediction of customer lifetime value encompasses the prediction of customer churn, those tasks have a strong relationship. Future purchasing behavior, which influences customer lifetime value and customer churn, is strongly related to latent purchasing patterns and preferences. Multi-task learning with neural networks is known to be effective for extracting such latent features in a data-driven manner and efficiently learning multiple related tasks. It is also well known that multi-task learning with neural networks can improve the overall task accuracy by extracting latent features commonly required for different tasks while learning different tasks with a single learner simultaneously. In this study, we propose a method for predicting customer lifetime value and customer churn by multi-task learning with multi-layer neural networks. In multi-task learning, the modeling of each task and the integration of the weights for task losses have a significant impact on the prediction accuracy. Therefore, we propose a method that models churn prediction as a partial task of CLV prediction while controlling the weights of each task according to the learning progress to improve the prediction accuracy. We demonstrate the effectiveness of the method through experiments using real life data.

**Keywords:** customer lifetime value (CLV), customer churn, multi-task learning, neural networks, RFM analysis

## 1. はじめに

機械学習の手法が広く社会に浸透するにつれ、ビジネス上の課題を機械学習で解決することへの社会的期待はますます大きくなっている。特に顧客の将来の購買行動を予測しようとする試みは、アカデミックリサーチだけでなく、ビジネスの現場でも非常に関心が高く、様々な研究が行われている。これらの研究には顧客の嗜好を分析し、短期的に商品推薦を行うための研究 [1], [2], 顧客の離反や顧客生涯価値 (CLV: Customer Lifetime Value) [3] など長期的な購買行動を予測し、より戦略的なビジネスの意思決定に役立てようとする研究 [4], [5], [6], [7] などがある。これらの多くは過去の購買履歴や行動履歴、顧客属性をもとに、顧客の将来の購買行動を予測しており、機械学習などの統計モデルが活用されている。このような顧客購買行動の予測の中でも、離反や CLV の予測は企業のマーケティング戦略策定のうえで非常に重要である。たとえば、企業は顧客の離反を予測し、マーケティング施策などで適切な介入をすることで、顧客の離反を食い止め、長期的な関係を築くことができる。また顧客ごとの CLV を理解することで、適切なターゲットを見極め、より収益性の高いマーケティング施策に注力することができる。これらの顧客の離反行動や CLV の予測は目的は異なるものの顧客行動の予測という点では同質である。そこで本研究では顧客の離反と CLV に着目し、これらのタスクをマルチタスク学習の手法を用いて、より高い精度で実行する手法を提案する。

顧客の離反予測や CLV の予測は様々な業界を対象に様々な手法で研究されている。離反予測は、顧客の離反が観測でき、また離反による損失が大きい通信業界 [7], [8] や金融 [9] などの契約型ビジネスにおいて、活発に研究されてきた。これらの研究では離反自体は明確に観測できるため、単純な二値分類問題として機械学習の手法でモデル化されることが多い。一方で小売業などの非契約型のビジネスにおいては、離反は潜在的なイベントとして発生するため、確率モデルを用いたアプローチをとる研究が多い。確率的な手法については Schmittlein らによる Pareto/NBD モデル [10] が広く知られおり、その拡張手法も多数提案 [11], [12], [13] されている。また契約型ビジネスはほとんどの場合、その契約期間に CLV が比例するため、CLV 予測の研究 [6], [7], [14], [15] の中心は非契約型のビジネスとなっている。

本研究の対象事例は非契約型のビジネスであるプラット

フォームビジネス [16], [17] である。近年、アップルやアマゾンに代表されるプラットフォームビジネスは目覚ましいスピードで成長し、市場を席卷している。プラットフォームビジネスは多数の消費者と生産者を結び付ける機能を有しており、消費者の購買行動は単一の店舗を対象とした場合よりも複雑である。そのため、データの背後に潜む複雑な関係性をモデル化するには仮説に基づく統計的なアプローチよりも機械学習を用いたデータ駆動型のアプローチが有効である。そこで近年、多くの適用事例で有効性が示されている多層ニューラルネットワークを用いて、データ駆動型のアプローチにより、有効な特徴を抽出することを考える。従来、対象となる顧客の離反予測と CLV 予測は個別に扱われており、これらのモデルを別々構築し、施策検討に活用することもできる。しかし、これらのタスクには関係性があるため、マルチタスク学習によって2つのタスクの関係性を考慮したモデル化を行うことにより、両者の予測精度が向上する可能性がある。そこで、ニューラルネットワークの多層構造を活用し、マルチタスク学習モデルを構築することを考える。

マルチタスク学習 [18], [19] は複数の異なるタスクを同時に実行することによって、共通する特徴を効率的に学習するとともに、汎化性能を高める手法である。近年、ディープラーニングの研究が進むにつれ、画像認識 [20], [21] や文章生成 [22] の分野など多数の関連するタスクが存在する領域で、ニューラルネットワークによるマルチタスク学習の研究がさかんに行われている。ビジネス領域における適用事例は比較的少ないものの、購買データを入力とした顧客購買行動の予測など、一定の関連があるビジネスタスクの予測や識別には効果的であると考えられる。しかしマルチタスク学習においては複数のタスクの損失を統合する際の重みの設定が推定精度の向上に大きな影響を与えることが知られており [20], 単純にマルチタスク学習を適用するだけでなく、モデルに適した重みの統合が重要となる。そこで本研究では顧客の離反と CLV の関係性を考慮し、相対的に容易なタスクである離反予測の結果が、CLV 予測に活用されるようモデルを構築し、重みの学習のスケジュールする手法を提案する。またプラットフォームビジネスにおける実データを用いた実験により有効性を示す。

本研究の貢献は以下のとおりである。1) 多層ニューラルネットワークを用いたマルチタスク学習により、プラットフォームビジネスにおける顧客の離反と CLV の予測を行い、実データで有効性を確認した。2) 購買行動予測のタスクの関係性に着目し、マルチタスク学習を効果的に活用するモデルを提案し、その有効性を示した。

## 2. 準備

### 2.1 関連研究

顧客の購買行動を予測する研究には様々なものがあるが、

<sup>1</sup> 早稲田大学データサイエンス研究所  
Institute of Data Science, Waseda University, Shinjuku,  
Tokyo 169-8555, Japan

<sup>2</sup> 早稲田大学理工学術院  
Faculty of Science and Engineering, Waseda University,  
Shinjuku, Tokyo 169-8555, Japan

a) khasumoto@fuji.waseda.jp

b) masagoto@waseda.jp

ここでは本研究の対象領域となる顧客の離反予測と CLV 予測の従来研究について述べる。

### 2.1.1 顧客離反の予測

顧客の離反とは企業からサービスや商品の購入を行っていた顧客がなんらかの理由でその企業との取引をやめてしまうことである。離反した顧客からの収益は新規顧客を獲得することでカバーすることもできるが、新規顧客の獲得コストは既存顧客の維持コストの5倍かかるといわれており [23]、多くの企業にとって顧客の離反を防ぎ、既存顧客から継続的に収益を上げることは重要な課題となっている。そのため、顧客の離反予測は様々な業界で活発に研究されており、特に契約型顧客関係といわれ、顧客の離反が契約解除のよって明確に観測可能な通信業界や銀行では様々な事例がある [7], [8], [9]。一方で小売りやイーコマースなどの非契約型の顧客関係では、離反が明示的に観測できないため、ベイズ理論に基づく確率モデルで潜在的な離反を予測するか [10], [11]、離反を「一定期間の購入がない」、「購入額の相対的な減少」などで定義し、二値分類問題として機械学習の手法を用いて予測するモデル [4], [24], [25] がある。Ascarza [26] によれば、契約型の顧客関係では離反が明確なため、離反要因の調査が研究の中心であり、非契約型の顧客関係では離反自体が観測できないため、どの顧客が離反したのかを予測する研究に重きが置かれている。

### 2.1.2 顧客生涯価値の予測

CLV はある顧客が企業との取引を通じてその生涯で企業にもたらす利益を表す。顧客が最初の取引から離反するまでの期間に企業にもたらす収益を現在価値に割引することで計算される [3]。CLV の計算期間を  $T$ 、期間  $t$  に顧客が支払った価格を  $p_t$ 、コストを  $c_t$ 、生存確率を  $r_t$ 、割引率または資本コストを  $i$  とすると、CLV は式 (1) で計算される。ここで生存確率とは顧客が企業から商品を継続して購入する確率であり、離反確率を  $a_t$  とすると、 $a_t = 1 - r_t$  となる。そのため本質的に生存予測と離反予測は同義である。

$$CLV = \sum_{t=0}^T \frac{(p_t - c_t)r_t}{(1+i)^t} \quad (1)$$

CLV を予測するには顧客がいつ離反するのか、どの商品をどれくらい購入し、その取引からの利益はどれくらいなのかを知る必要があり、取引ごとの利益は企業のビジネスモデルに依存するため、顧客の離反予測よりも複雑なモデルが必要となる。一方で顧客の離反情報が CLV に影響するため、顧客の離反予測を適切に行うことができれば、CLV の予測精度も高まると考えられる。

CLV 研究においては確率モデルを用いた Pareto/NBD モデル [10] が広く知られている。Pareto/NBD では購入数をポアソン分布、継続購入期間を指数分布、顧客の異質性をガンマ分布のパラメータで表現することで将来の購買をモデル化している。これらの手法は非契約型の顧客関係

において、潜在的な離反を考慮したうえで将来の購買予測を予測できるという点で優れているが、CLV の予測に有用であると考えられる顧客の購買パターンや取引の利益構造などは考慮されていない。それらを潜在的なパラメータを確率的なモデルに組み込むこともできるが、モデルが非常に複雑となり現実的ではない。購買データの背後にある潜在的特徴を考慮して CLV 予測を行う手法としてニューラルネットワークを活用する方法があり、時系列データに Convolutional Neural Network (CNN) を適用する手法 [27] やマルチタスク学習を利用する方法 [28] などが提案されている。

### 2.1.3 マルチタスク学習

マルチタスク学習では関連するタスクを同時に学習することで、タスクに共通する表現を学習し、個別に学習する場合よりも各タスクの精度を高めることができる。またタスクを同時に実行するため、効率的に処理をこなすことができ、計算機資源が限られている場合などに有用である。さらにマルチタスク学習の効果として、Eavesdropping (盗み聞き) [18], [19] が知られている。これはタスクの難易度に差がある場合、難易度の高いタスクは難易度の低いタスクから情報を得て、より容易に学習が進められるというものである。そのため、本研究で扱う CLV 予測と離反予測はタスクの関連性およびタスクの難易度の観点でマルチタスク学習に適した問題設定であると考えられる。

近年、ディープラーニングにより画像や言語の処理能力が飛躍的に高まったこともあり、画像認識や文字列認識・生成においてニューラルネットワークを活用したマルチタスク学習の研究が活発に行われている。CLV 予測における研究事例は比較的少ないが、損失に対数正規分布を用いた Wang らの手法 [28] や航空旅客を対象に RNN を用いた提案 [29] がされている。Wang らの手法では誤差に対数正規分布を用いることで外れ値に頑健になると同時に、離反顧客と継続顧客を異なる分布と考え、指示関数を用いて損失を統合している。

マルチタスク学習では異なるタスクを同時に実行するため、それぞれの損失をどのように重みづけし、統合していくかが最終的な精度向上に重要となる [20]。損失の統合には様々な提案がなされており、学習進度に応じて重みを動的に変化させる手法 [21] や不確実性に基づいて損失を定義する手法 [20]、学習進度に応じて学習データや重みをスケジューリングして変えていく方法 [22] などがある。

## 2.2 対象事例

本研究の対象事例はプラットフォームビジネス [16], [17] である。プラットフォームビジネスには楽天市場のようなマーケットプレイス型ビジネス、フェイスブックのようなソーシャルプラットフォーム、アップルのようなデバイスやコンテンツを含む統合的なプラットフォームなど様々な

形態がある。本研究では多数の買い手と売り手が存在し、商品やサービスの購入がなされるプラットフォームを想定し、そのビジネスにおける顧客の離反や CLV の予測を行う。プラットフォームビジネスでは、買い手となる消費者だけでなく、売り手となる店舗も複数存在するため、顧客の購買行動は単一のビジネスよりも多様である。

顧客がプラットフォームのサービスを利用する際、事前にアカウントを作成し、サービスにログインしたうえで、そのサービスを利用している店舗から商品などの購入を行う。その際、売り上げの一部が店舗からプラットフォーム事業者へ手数料として支払われる。そのため、プラットフォーム事業者の収益は消費者の利用金額に比例するが、手数料率が事業者ごとにことなるため、実際の CLV はどのような事業者を顧客が利用しているかによっても異なる。また顧客の購買周期も利用店舗によって異なるため、潜在的な顧客の嗜好や購買パターンの理解が予測精度の向上には重要であると考えられる。

なお顧客は明示的にアカウントを閉じることができ、事業者は離反を確認することもできるが、多くの顧客は単純に利用を止めるだけで、明示的にアカウントを閉じることが少ない。そのため、実務上離反は一定期間の利用がないことで判断しており、ビジネスモデルとしては非契約型といえる。

### 3. 提案手法

#### 3.1 概要

本研究ではニューラルネットワークによるマルチタスク学習を用いた CLV と顧客離反の同時予測モデルを提案する。式 (1) に示されたとおり、CLV の算出には顧客の離反情報が必要であり、顧客の離反は CLV に比べると、比較的容易に推定可能である。そのため、CLV の予測に離反予測の結果を活用することで、CLV の予測精度を高めることが可能と考えられる。そこで本提案では離反予測の結果を CLV に利用し、同時にマルチタスク学習により推定精度を改善することを考える。離反予測は生存・離反の二値分類であるため、本モデルでは生存確率  $p$  を出力の 1 つとし、CLV 予測に利用する。ここで CLV は顧客ごとの生存確率  $p$  と購買量を表す観測されない未知の潜在変数  $z$  によって算出されると仮定する。この潜在変数  $z$  をもう 1 つのタスクの出力とし、CLV を間接的に予測する。さらにタスクの難易度を考慮して、エポック数に応じて重みを変化させることで Eavesdropping 効果が得られるように学習進捗を調整する。図 1 にモデルの概要を示す。このように CLV を直接予測せずに潜在的な変数を予測させたことには 2 つの理由がある。1 つは相対的に容易なタスクである離反予測を優先的に学習させるため、その結果を直接利用の方が効果的であると考えたことである。2 つ目は、ここで得られる潜在変数が顧客の購入量や顧客からの利益の指標とな

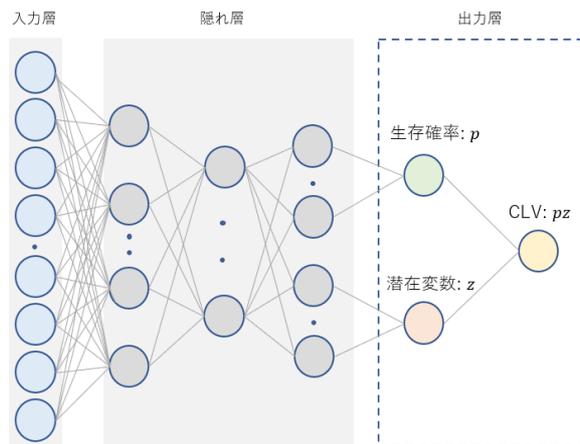


図 1 提案モデル概要

Fig. 1 Proposed model overview.

り、実務上有益な情報になると考えたためである。なお、モデルからの出力は生存確率になるが、生存予測と離反予測は本質的には同義のため、以降離反予測として扱う。

#### 3.2 定式化

一般に、マルチタスク学習の損失は、式 (2) のように複数のタスクに一定の重みをかけて合計したもので与えられる。ここで  $\lambda_t$  はタスクごとの重み、 $L_t$  はタスクごとの損失、 $t$  はタスク ID、 $T$  はタスク数である。

$$loss = \sum_{t=1}^T \lambda_t L_t \quad (2)$$

本研究では 2 つのタスクを同時に実行するため、離反予測と CLV 予測の誤差を交差エントロピーと平均二乗誤差を用い、式 (3) のとおり損失を定義した。  $N$  は学習データ数、 $p_i$  は生存確率、 $y_i$  は CLV 値である。文字上のハットは予測値であることを表し、CLV の予測値は生存確率  $\hat{p}_i$  と潜在変数  $\hat{z}_i$  の乗算で与えられる。

$$loss = \lambda_{alive} \left( \sum_{i=1}^N \{p_i \log \hat{p}_i + (1 - p_i) \log(1 - \hat{p}_i)\} \right) + \lambda_{clv} \sum_{i=1}^N (y_i - \hat{p}_i \hat{z}_i)^2 \quad (3)$$

ここで比較のために、CLV・離反予測のマルチタスク学習モデルである Wang ら [28] の手法における損失関数を以下に示す。  $\mu_i$  と  $\sigma_i$  はモデルの出力で、 $\mathbf{1}$  は生存・離反を示す指示関数である。

$$loss = \sum_{i=1}^N \{p_i \log \hat{p}_i + (1 - p_i) \log(1 - \hat{p}_i)\} + \sum_{i=1}^N \mathbf{1}_{\{y_i > 0\}} \left( \log \left( y_i \hat{\sigma}_i \sqrt{2\pi} \right) + \frac{(\log y_i - \hat{\mu}_i)^2}{2\hat{\sigma}_i^2} \right) \quad (4)$$

Wang らの手法は CLV 予測・離反予測の損失をそれぞれ

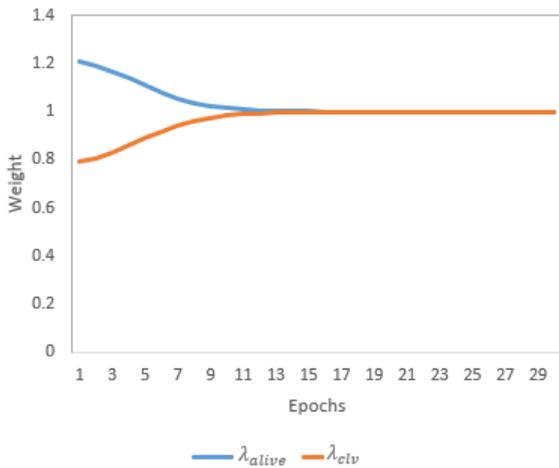


図 2 エポック数ごとの重み変化

Fig. 2 Change of weights over epochs.

れ計算して統合するという点においては提案手法と同様であるが、CLV 予測の損失に対数正規分布の尤度を使用している点、さらにその損失を指示関数により生存時のみ考慮している点で異なる。これは Wang らの手法では CLV が連続的な 1 つの分布ではなく、離反顧客と生存顧客で異なる分布と仮定しているためである。

提案手法では容易なタスクである離反予測の学習を優先的に行うためにスケジュール関数を用い、学習初期は離反予測の学習の損失重みを大きくし、徐々に均一な学習となるようスケジュールする。スケジュール関数は西塾らの手法 [22] を参考にシグモイドを使用し、以下の式で定義した。

$$\lambda_{active} = \lambda_{active}^{ini} \left( \frac{2w_{active}}{w_{active} + w_{clv}} \right) \quad (5)$$

$$\lambda_{clv} = \lambda_{clv}^{ini} \left( \frac{2w_{clv}}{w_{active} + w_{clv}} \right) \quad (6)$$

$$\text{ただし, } w_{active}(n) = \left\{ \frac{1}{1 + \exp\left(\frac{\alpha-n}{2}\right)} \right\},$$

$$w_{clv}(n) = \left\{ \frac{1}{1 + \exp\left(\frac{\beta-n}{2}\right)} \right\}$$

ここで  $\lambda_{active}^{ini}$ ,  $\lambda_{clv}^{ini}$  はそれぞれのタスクの重みの初期値で、事前に各タスクの学習収束時の損失を考慮したうえで決定される。 $\alpha$ ,  $\beta$  はスケジューラ用の係数でエポック数  $n$  に応じた重みの変化を制御している。本研究では事前に  $\alpha$ ,  $\beta$  を 1~10 の範囲でテストし、最も精度の高かった  $\alpha = 4$ ,  $\beta = 5$  の組合せを採用した。この結果、重みの初期値にかけられる係数  $w_{active}$ ,  $w_{clv}$  はエポック数が小さいときは生存予測で大きく、CLV 予測で小さく、徐々に均一になるように調整されている。このときのエポック数ごとの重みの変化を図 2 に示す。

### 3.3 入力変数

購買行動の予測において、適切な入力変数を使用するこ

とは非常に重要となる。さらに実運用を考えた場合、単純に予測に有効なだけでなく、精度が高く入手が容易なデータである必要がある。そこで本研究では過去の購買履歴より抽出可能なデータのみを使用する。購買履歴は売上データに紐づくため、どのような企業でも高い精度のデータが長期間保持されていると考えられる。本研究では、購買履歴より RFM 指標、購入店舗カテゴリごとの購入数、月次の時系列購入数を抽出し、すべてを入力変数として使用する。RFM 指標は RFM 分析に用いられる指標で顧客の購買傾向を示す指標として、アカデミックリサーチから実務まで幅広く利用されている。特に CLV の予測や離反予測においては有効な説明変数として多くの研究で用いられている [14], [15], [24], [25]。RFM はそれぞれの指標の頭文字を表し、それぞれ最新購買日 (Recency), 総購買回数 (Frequency), 総購買金額 (Monetary Value) を表す。最新購買日は顧客が最後に商品やサービスを購入した日にちであり、通常予測開始日前の最終購買日から予測開始日までの期間をとって数値化される。総購買回数は観測期間全体の購買回数合計である。総購買金額も同様に観測期間全体の購買金額合計となるが、回数と金額に相関があるため、本研究では平均購買金額を用いる。また基本的な RFM 指標に加えて、総購買店舗数 (Stores), 利用期間 (Tenure), 購買間隔 (Inter-purchase time) の 3 指標を追加し、extended RFM (eRFM) 指標として定義し、入力変数に使用する。総購買店舗数は重複を除いた購入先の店舗数で、プラットフォームビジネスのような複数の店舗が購入先となるビジネスにおいては店舗をまたいだ購入の多様性がその後の購買行動にも影響を与えられられる。Blattberg ら [30] によれば、クロス購買やマルチチャネルの利用は CLV 向上に好ましい購買行動とされており、これを裏付けている。利用期間はアカウント作成日から最新購買日までの日数で、購買間隔は利用期間中の購入と購入の間の日数を平均したものであり、ともに離反に一定の影響がある [4], [31] と考えられ、採用した。整理すると eRFM 指標は以下のとおりとなる。

$$\text{Recency} = \text{予測開始日} - \text{最終購買日}$$

$$\text{Frequency} = \text{総購買数}$$

$$\text{Monetary Value} = \text{総購買金額} / \text{Frequency}$$

$$\text{Stores} = \text{購買先の総ユニーク店舗数}$$

$$\text{Tenure} = \text{最終購買日} - \text{アカウント作成日}$$

$$\text{Inter-purchase Time}$$

$$= (\text{最終購買日} - \text{初回購買日}) / (\text{Frequency} - 1)$$

上記の eRFM 指標に加えて、購買店舗カテゴリごとの購入数と時系列情報である購入月ごとの購入数もニューラルネットワークへの入力変数として使用する。店舗カテゴリはプラットフォームビジネス上で商品やサービスを販売

している事業者に割り当てられた取扱商品のカテゴリでファッションや化粧品、書籍などのカテゴリを表し、その購入数は顧客の潜在的な嗜好の情報を含んでいると考えられる。またプラットフォームビジネスにおける取引ごとの利益は商品カテゴリや事業者によって異なることから、CLVを予測する上で重要な指標になると考えられる。さらに時系列の購入数は顧客の購買傾向や商品の購入サイクルを示す情報になりえるため、同様に予測に有効な入力と考えられる。

#### 4. 実験と考察

提案手法の有効性を確認するため、実データを用いて実験を行った。比較手法は、ニューラルネットを用いてシングルタスクで個別にCLV予測、離反予測を実行したモデル、マルチタスクを重みなしで実行したモデル、Wangらによって提案された損失統合手法を使用したモデル [28] とした。Wangらの手法では指示関数を用いることで、生存顧客のみCLVの予測誤差を損失に統合している。さらにCLVの誤差関数に対数正規分布を用いることで、外れ値に対して頑健になるようにモデル化している。しかし対数正規分布を用いると損失が負値となり、精度が著しく低下するため、本実験では対数正規分布は用いず、平均二乗誤差を使用した。実験は10分割交差検証を10回、すなわち100回の結果の平均値で評価を行った。評価指標はCLV予測については平均二乗誤差(MSE)と決定係数(R<sup>2</sup>)、顧客離反予測についてはF値、Accuracy、Precision、Recall、AUCを用いた。また参考として、線形回帰、ロジスティック回帰、ランダムフォレストで予測を行った結果も掲載した。

##### 4.1 実験データ

実験用のデータはプラットフォーム事業者A社から提供された約6万件の顧客の購買データを使用した。データ作成の概略を図3に示す。このデータはある年( $t$ 年度とする)にアカウントを開設したユーザで、当該年度を含めて3年分の購買情報が存在する。最初の2年を観測期間、最後の1年を予測期間と定義し、観測期間の購買情報をもとに予測期間におけるCLVおよび離反を予測する。入力データには、顧客の全体的な購買傾向を示す情報としてeRFM指標を観測期間全体のデータより作成し、また $t+1$ 年度



図3 実験データセットの概要

Fig. 3 Overview of experimental datasets.

のデータから直近の詳細な購買パターンを示す情報として購入先の店舗カテゴリごとの購入数と月次の購入数を抽出した。さらに $t+2$ 年度の情報から目的変数となるCLVおよび離反の有無を算出した。それぞれの次元数はeRFM指標が6次元、店舗カテゴリが47次元、月次の購入データが12次元の計65次元となる。対象顧客データについては、データ提供元企業の非公開情報の制約から生存比率や店舗カテゴリの実数をそのまま記載することができない。そこで、元データセットから、離反顧客と生存顧客が同数となるようランダムサンプリングにより抽出した。これは、目的変数のカテゴリ間でデータ数に偏りがある場合に、多数側のカテゴリからランダムにサンプリングしてデータ数を揃えるアンダーサンプリングと同等の手続きであり、データの加工手続き上の問題はない。また購買傾向確認の観点から過去に6件以上の取引がある顧客のみを対象としている。図4に提案モデルへの入力データのサンプルを示す。なお、実験データは事前に正規化されている。

##### 4.1.1 本研究での離反とCLVの定義

本研究で取り扱うプラットフォームビジネスのユーザはアカウントを明示的に閉鎖することができるが、多くのユーザは単純に利用をやめるだけで、その離反は潜在的である。この事業者では、1年間利用がない顧客を離反、ないしは休眠顧客と定義しているため、本研究では離反を予測期間の1年間に購買行動が観測されなかった顧客とする。CLVは理論上その顧客が生涯にわたってもたらす利益となるが、現実的に取得できるデータの期間は限られているため、これも実務上の利用が多いと考えられる予測期間1年間のCLVとした。

##### 4.2 事前分析

本研究で使用した顧客データを離反、生存の状態で層別し、CLVとeRFM指標の平均値を算出したものを表1に示す。生存顧客はCLVの大きさ順に5つのグループに分割した。CLVについては、グループごとに大きな差があり、上位に行くにつれて指数関数的に大きくなっており、データの偏りがみられる。最新購買日は離反顧客と生存顧客の間での差が最も大きく、離反の予測には重要な指標であることが分かる。一方でCLVが大きいグループ間での差異は小さく、CLV予測への影響は相対的に小さいと考えられる。平均購買金額はCLVの異なるグループ間では差

顧客ID	eRFM指標			購入店舗カテゴリ			購入月				
	Recency	Frequency	...	ソフトウェア	自動車用品	...	ファッション	1か月目	2か月目	...	12か月目
A001	31	3		1				1			1
A002	6	1						1			
A003	278	4		1			2	1			
A004	105	3			1			2	1		
A005	34	1						1			
A006	22	4						4			
A007	44	7					4	2			
A008	159	15		3			3	3			

図4 入力データの一部

Fig. 4 A sample of input data.

表 1 実験データの事前分析

Table 1 Preliminary analysis on experimental data.

状態	CLV	最新購買日 (Recency)	総購買数 (Frequency)	平均購買金額 (Monetary)	総購買店舗数 (Stores)	利用期間 (Tenure)	購買間隔 (IP Time)	顧客数
離反	0.00000	0.45463	0.01378	0.02279	0.00612	0.52847	0.29125	28,800
	0.00405	0.20989	0.01444	0.01373	0.00694	0.65421	0.33652	5,760
	0.01430	0.16102	0.01688	0.01548	0.00730	0.67664	0.30960	5,760
生存	0.03157	0.12439	0.01883	0.01879	0.00714	0.69230	0.28109	5,760
	0.06657	0.11094	0.02221	0.02821	0.00918	0.70011	0.25979	5,760
	0.26979	0.10704	0.03059	0.05574	0.01636	0.70409	0.23302	5,760
全体	0.03863	0.29864	0.01718	0.02459	0.00775	0.60697	0.28763	57,600

表 2 CLV/離反予測 実験結果

Table 2 The result of prediction for CLV and churn.

\*本実験の Wang らの手法では二乗誤差を使用している  
\*\*複号の後の数値は信頼区間 (95%) を示す

タスクタイプ	手法	重みづけ	CLV予測		離反予測				
			MSE	R2	F-score	Accuracy	Precision	Recall	AUC
シングルタスク	ニューラルネットワーク	-	0.00742 ± 0.00011	0.26834 ± 0.00558	0.73237 ± 0.00341	0.76106 ± 0.00161	<b>0.83256 ± 0.00350</b>	0.65604 ± 0.00701	0.84313 ± 0.00095
シングルタスク	線形回帰	-	0.00819 ± 0.00011	0.19346 ± 0.00398	-	-	-	-	-
シングルタスク	ロジスティック回帰	-	-	-	0.76869 ± 0.00113	0.75392 ± 0.00119	0.72518 ± 0.00117	0.81780 ± 0.00155	0.83281 ± 0.00118
シングルタスク	ランダムフォレスト	-	0.00754 ± 0.00011	0.25731 ± 0.00466	<b>0.78594 ± 0.00109</b>	<b>0.78881 ± 0.00107</b>	0.79681 ± 0.00131	<b>0.77540 ± 0.00142</b>	<b>0.86755 ± 0.00086</b>
マルチタスク	ニューラルネットワーク	-	0.00742 ± 0.00011	0.26923 ± 0.00485	0.77317 ± 0.00178	0.77735 ± 0.00107	0.78845 ± 0.00347	0.75980 ± 0.00584	0.85652 ± 0.00092
マルチタスク	ニューラルネットワーク	Wangらの手法*	0.00748 ± 0.00011	0.26356 ± 0.00517	0.77157 ± 0.00181	0.77557 ± 0.00110	0.78626 ± 0.00392	0.75904 ± 0.00632	0.85484 ± 0.00099
マルチタスク	ニューラルネットワーク	提案手法	<b>0.00728 ± 0.00010</b>	<b>0.28304 ± 0.00425</b>	0.77119 ± 0.00170	0.77605 ± 0.00104	0.78875 ± 0.00335	0.75559 ± 0.00552	0.85443 ± 0.00095

があるものの、離反顧客と全体の平均とでは大きな差がなく、CLV の予測には有用であるが離反顧客と生存顧客の判別には影響が少ないと考えられる。また平均購買金額や総購買店舗数は CLV が下位のグループでは大きな差がないものの、CLV 上位のグループでは突出しており、CLV が大きい優良顧客の予測には有効であることが考えられる。全体的に CLV や離反状態と eRFM 指標との間には非線形な関係がみられ、詳細な購買情報をもとに非線形なモデルを構築することで、より高い精度での予測が可能になると考えられる。

### 4.3 実装

ニューラルネットワークの構築には PyTorch を使用した。入力層のノード数は 65、共有の隠れ層として、64 ノードと 32 ノードの層があり、その後タスク別の隠れ層として、16 ノードの層を介して最終的に 1 次元の出力で、生存確率および潜在変数  $z$  の推定値を出力する。epoch 数は 30、バッチ数は 256 とし、学習の最適化は Adam を学習率 0.01 で使用した。マルチタスク学習の重みの初期パラメータは事前実験の結果より学習収束時に損失が均等になるように、 $w_{alive}^{ini} = 1$ 、 $w_{clv}^{ini} = 0.05$  とした。活性化関数は最初の 2 層が Relu 関数、残りの 2 層は恒等関数を使用した。ただし離反予測の最終層はシグモイド関数を使用している。

### 4.4 実験結果

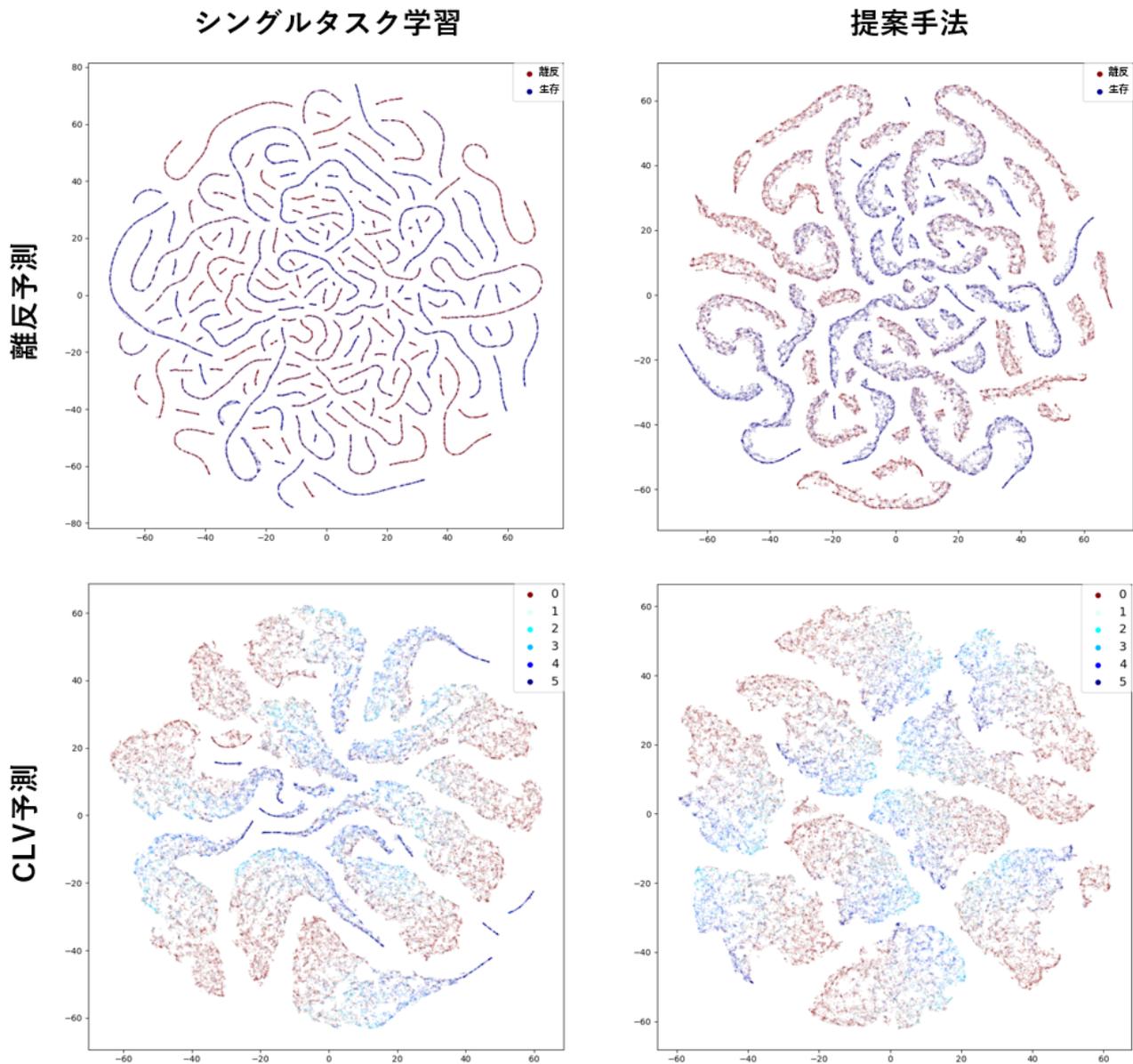
表 2 に CLV 予測、離反予測の結果を示す。タスクタイプがシングルタスクの場合は手法に限らず、CLV 予測、離

表 3 CLV グループごとの生存確率  $\hat{p}_i$  と潜在変数  $\hat{z}_i$  の平均値

Table 3 The average of the alive probability  $\hat{p}_i$  and the latent variable  $\hat{z}_i$  for each CLV group.

状態	CLV	$\hat{z}_i$	$\hat{p}_i$	顧客数
離反	0.00000	0.05195	0.30769	28,800
	0.00405	0.04708	0.56505	5,760
	0.01430	0.05173	0.63763	5,760
生存	0.03157	0.05787	0.70097	5,760
	0.06657	0.07902	0.73420	5,760
	0.26979	0.13920	0.75688	5,760
全体	0.03863	0.06347	0.49332	57,600

反予測を個別に学習・予測しており、マルチタスクは 2 つを同時に学習・予測している。全体的にニューラルネットワークは CLV 予測で高い精度を示しているが、その中でも提案手法は最も高い精度を示しており、提案手法の有効性が確認できた。一方で、マルチタスク学習であっても、シングルタスク学習より必ずしも高い精度となっている訳ではなく、損失統合のための重み調整の重要性が再確認された。離反予測については相対的に予測タスクの難易度が低いため、手法間の差異は比較的小さかったが、ランダムフォレストによる予測が最も良い値を示した。ニューラルネットワークを用いたシングルタスクの離反予測では、Recall が相対的に低く、Precision が高く出ており、予測に偏りがみられるが、マルチタスク学習ではどちらも高い数値を示しており、バランス良く学習していることが分かる。またマルチタスク間では、信頼区間も考慮すると離反予測の精度に大きな差異はみられなかった。離反予測に関しては、タスク自体が相対的に容易なため、予測結果の差異が小さいと考えられる。表 3 は CLV グループごとの生存確率  $\hat{p}_i$  と



\*CLV 予測の凡例は CLV の大きさ順に 5~0 となっており、離反顧客は 0 である

図 5 t-SNE による中間層分析結果

Fig. 5 An analysis of latent representation by t-SNE.

潜在変数  $\hat{z}_i$  の平均値を示したものである。潜在変数  $\hat{z}_i$  は CLV から生存確率の取り除いた値となるため、潜在的な購入ボリュームや顧客からもたらされる利益額を示す変数になるが、CLV に比例して大きくなっており、期待した値が抽出されていると考えられる。また生存確率  $\hat{p}_i$  も離反したグループと生存したグループで大きな差異がある一方で、生存グループ間では変化がなだらかになっており、離反予測が機能していると考えられる。

#### 4.5 t-SNE による中間層分析

ニューラルネットによって抽出された特徴は、離反予測や CLV 予測に重要な指標であるため、それらを活用した顧客のセグメンテーションや顧客購買行動の分析に活用す

ることができると考えられる。そこで提案手法で予測に重要な特徴が抽出されているか、シングルタスク学習で抽出された特徴と違いがあるのかを確認するため、t-SNE を用いて、最後の隠れ層の特徴量を 2 次元で可視化した。本分析では 10 分割交差検証を個別に 1 回実行し、そのテストデータを統合して使用した。t-SNE の Perplexity は 50 である。図 5 にシングルタスクおよび提案手法の可視化データを示す。CLV 予測のプロットについては事前分析で使用した CLV の大きさ順のグループを使用しており、大きい順に値が 5 から振られており、離反顧客は 0 である。まず離反予測のデータについては、どちらの手法においても離反データ、生存データがある程度分かれてプロットされており、判別に必要な内部表現が獲得できていることが分か

る。特に提案手法においては、入り組んではいるものの比較的生存顧客が画像中央に集まっており、相対的に識別しやすいものと考えられる。同様に CLV 予測においても、離反・生存でプロットが分かれており内部表現による識別ができていていると考えられる。一方で生存顧客内での CLV の大きさごとの判別は可視化された 2 次元空間のデータからは判別が困難であった。これは CLV 予測精度がそれほど高くないため、中間層では違いが明確に可視化できなかったと考えられる。ただし提案手法とシングルタスク学習を比べた場合、提案手法は相対的にプロットに広がりがあるため、非線形識別であれば局所的に存在する CLV が異なるユーザー群をよりとらえやすいのではないかと考えられる。

## 5. 考察

CLV の予測精度については、提案手法が最も精度を示した。これはマルチタスク学習における Eavesdropping 効果により、汎化性能が向上した結果と考えられる。一方で同じマルチタスク学習でも、重みの調整をしない場合はシングルタスク学習と大きな差異がなく、重みの適切な設定が重要であることが確認された。また Wang らの手法では予測精度がシングルタスク学習を下回った。この手法では離反顧客と生存顧客の CLV が異なる分布であると仮定し、指示関数を用いて生存顧客のみ予測誤差を学習している。しかし本研究の顧客データの CLV は表 1 から 0 付近にも多数存在しており、連続して分布していると考えられる。それらの CLV が小さく離反しそうであるが実際には生存している顧客の予測誤差が十分に考慮されず、結果的に全体の予測精度が低下した可能性がある。提案手法では生存確率をそのままモデルに組み込んだことで、実際の離反有無にかかわらず、全顧客データをバランス良く学習でき、精度が高くなったと考えられる。離反予測においては、タスクが比較的容易なため、手法間で大きな差異は出なかったが、シングルタスク学習よりはマルチタスク学習の精度が高くなっており、マルチタスク学習による改善がみられた。一方で単純な精度比較ではランダムフォレストの予測精度がニューラルネットを上回った。これには大きく 3 つの可能性が考えられる。1 つ目は離反予測に利用されるデータの特徴による影響である。表 2 から離反顧客と生存顧客では最新購買日 (Recency) に大きな差異があり、離反には最新購買日が強く影響していることを示している。このような結果に大きな影響を与える入力変数がある場合、複雑な交互作用などの潜在的な特徴を購買データから集約しながら利用するニューラルネットワークよりも、個々の入力データを直接活用するランダムフォレストの方が高い精度となる可能性がある。2 つ目はニューラルネットの過学習である。ニューラルネットは層を深くすることで複雑なタスクをモデル化できる一方で、その柔軟性から過学習に陥りやすい。離反予測は比較的容易なタスクのため、単純

な手法でも比較的モデル化しやすいが、CLV 予測という複雑なタスクを実現するうえで中間層を 3 層とした。マルチタスク学習における双方のタスクの汎化性向上は一定程度あったと考えられるが、離反予測においては過学習が影響し、十分に性能が発揮できなかったと考えられる。最後に訓練データの量の影響である。本研究では約 6 万件の顧客データを使用した。ニューラルネットワークの学習で非線形構造を抽出するには十分とはいえず、相対的にシンプルなモデルであるランダムフォレストの予測精度が良好となる領域に、この学習データ量とタスクの難易度が入っていたのではないかと考えられる。一方、ニューラルネットワークで抽出される潜在的な特徴は CLV や離反に関連する特徴として、顧客のセグメンテーションや優良顧客判別に活用できる可能性があり、単純な精度向上以外の用途でも幅広い活用が期待される。そのため、「モデルの自由度」、「必要となる学習データ数」、「モデルの説明可能性」などもふまえて、ビジネスのニーズに合わせて適切な手法を選択することが重要である。

## 6. まとめと今後の課題

本研究では、ニューラルネットワークによるマルチタスク学習により、顧客生涯価値と顧客離反を同時に予測し、予測精度を改善する手法を提案した。またプラットフォームビジネスにおける実データを用いた評価実験により、提案手法の有効性を確認した。さらに中間層表現を分析することで、ニューラルネットワークを通して、予測に重要な特徴が抽出されていることを確認した。本研究では顧客離反の予測タスクが顧客生涯価値の予測タスクに内包される点に着目し、マルチタスクニューラルネットワークのモデリングを行ったが、同様の問題設定は異なるビジネス上のタスクでも存在し、応用可能なものと考えられる。

今回のような重みのスケジューリングやネットワークの設計はドメインナレッジが必要となり、現段階では自動的に決定することができない。また利用できる入力変数はすべて使用したが、ニューラルネットワークの構造や学習パラメータとの相性によって、入力変数の選択が予測性能に影響を与える可能性もある。これらもふまえて、タスクの難易度を考慮して、スケジューリング関数の設計やパラメータ設定、モデリング全体の自動化は今後の課題である。また現段階の CLV の予測精度は十分であるとはいえない。必要とされる精度のレベルは目的や環境により異なるため、明示することは難しいが、筆者の個人的経験や事業者との議論からは、実運用に向けて利用者の協力を得るには、決定係数で 0.5 以上が求められることも多い。さらなる精度の向上と、顧客セグメントレベルでの予測など実社会での適用を想定した現実的な運用方法も検討が必要である。

謝辞 本研究にあたり、多くのご助言をいただいた湘南工科大学の三川健太先生、上智大学の山下遥先生、後藤研

研究室の方々，また貴重なデータを提供いただいたプラットフォーム事業者 A 社の皆様に深く感謝いたします。

参考文献

- [1] Koren, Y., Bell, R. and Volinsky, C.: Matrixfactorization techniques for recommender systems, *IEEE Computer Society*, Vol.42, pp.30–37 (2009).
- [2] Grbovic, M., Radosavljevic, V., Djuric, N., et al.: E-commerce in your inbox: Product recommendations at scale, *Proc. ACM SIGKDD*, pp.1809–1818 (2015).
- [3] Gupta, S. and Lehmann, R.D.: *Managing customers as investments*, Wharton School Publishing (2005). スカイライトコンサルティング (訳): 顧客投資マネジメント, 英治出版 (2005).
- [4] Buckinx, W. and Van den Poel, D.: Customer base analysis: Partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting, *Eur. J. Oper. Res.*, Vol.164, No.1, pp.252–268 (2005).
- [5] Lemmens, A. and Croux, C.: Bagging and boosting classification trees to predict churn, *J. Mark Res.*, Vol.43, No.2, pp.276–286 (2006).
- [6] Khajvand, M., Zolfanghar, K., Ashhoori, S., et al.: Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study, *Procedia Computer Science*, Vol.3, pp.57–63 (2011).
- [7] Chamberlain, B.P., Cardoso, A., Liu, C.H.B., et al.: Customer lifetime value prediction using embeddings, *Proc. ACM SIGKDD*, pp.1753–1762 (2017).
- [8] Huang, B., Kechadi, M.T. and Buckley B.: Customer churn prediction in telecommunications, *Expert Syst. Appl.*, Vol.39, No.1, pp.1414–1425 (2012).
- [9] Xie, Y., Li, X., Ngai, E. and Ying, W.: Customer churn prediction using improved balanced random forests, *Expert Syst. Appl.*, Vol.36, No.3, pp.5445–5449 (2009).
- [10] Schmittlein, C.D., Morrison, G.D. and Colombo, R.: Counting your customers: Who-are they and what will they do next?, *Management Science*, Vol.33, No.1, pp.1–24 (1987).
- [11] Abe, M.: “Counting Your Customers” One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model, *Marketing Science*, Vol.28, No.3, pp.541–553 (2009).
- [12] Fader, P.S., Hardie, B. and Lee, K.L.: “Counting Your Customers” the Easy Way: An Alternative to the Pareto/NBD Model, *Marketing Science*, Vol.24, No.2, pp.275–284 (2005).
- [13] Platzer, M. and Reutterer, T.: Ticking Away the Moments: Timing Regularity Helps to Better Predict Customer Activity, *Marketing Science*, Vol.35, No.5, pp.779–799 (2016).
- [14] 阿部 誠: RFM データを用いた顧客生涯価値の算出: 既存顧客の維持介入と新規顧客の獲得, *マーケティングジャーナル*, Vol.34, No.1, pp.73–90 (2014).
- [15] 蓮本恭輔, 雲居玄道, 後藤正幸: 非負値行列因子分解を用いたプラットフォームビジネスにおける顧客生涯価値予測, *情報処理学会論文誌*, Vol.60, No.7, pp.1283–1293 (2019).
- [16] 富士通総研・早稲田大学ビジネススクール根来研究室 (編著): プラットフォームビジネス最前線, pp.2–13, 翔泳社 (2013).
- [17] Van Alstyne, M.W., Parker, G.G. and Choudary, S.P.: Pipelines, platforms, and the new rules of strategy, *Harvard Bus. Rev.*, Vol.94, No.4, pp.54–62 (2016).
- [18] Caruana, R.: Multitask learning, *Machine Learning*, Vol.28, No.1, pp.41–75 (1997).
- [19] Ruder, S.: An overview of multi-task learning in deep neural networks, arXiv preprint arXiv:1706.05098 (2017).
- [20] Kendall, A., Yarin, G. and Roberto, C.: Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.7482–7491 (2018).
- [21] Liu, S., Johns, E. and Davison, A.J.: End-to-end multi-task learning with attention, *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.1871–1880 (2019).
- [22] 西埜 徹, 三沢翔太郎, 狩野竜示ほか: マルチタスク学習による一貫性のある求人記事分類及び職種フレーズと記事見出し文の生成, *言語処理学会第 25 回年次大会発表論文集*, pp.1277–1280 (2019)
- [23] Kotler, P., Keller, K.L. and Manceau, D.: *Marketing Management*, 15e edition, Pearson Education (2016).
- [24] Hopmann, J. and Thede, A.: Applicability of customer churn forecasts in a non-contractual setting, *Innovations in Classification, Data Science, and Information Systems*, pp.330–337, Springer (2005)
- [25] Jahromi, A.T., Stakhovych, S. and Ewing, M.: Managing B2B customer churn, retention and profitability, *Ind. Mark. Manag.*, Vol.43, No.7, pp.1258–1268 (2014).
- [26] Ascarza, E., Netzer, O. and Hardie, B.G.: Some customers would rather leave without saying goodbye, *Market. Sci.*, Vol.37, No.1, pp.54–77 (2018).
- [27] Chen, P.P., Guitart, A., del Río, A.F., et al.: Customer lifetime value in video games using deep learning and parametric models, *2018 IEEE International Conference on Big Data* (2018).
- [28] Wang, X., Tianqi, L. and Jingang, M.: A deep probabilistic model for customer lifetime value prediction, arXiv preprint arXiv:1912.07753 (2019).
- [29] Haofei, Y., Youfang, L., Zhihao, W., et al.: A Time-aware Multi-task Learning Model for Customer Value Prediction in Civil Aviation, *Pro. 2019 9th International Workshop on Computer Science and Engineering*, pp.588–598 (2019).
- [30] Blattberg, R.C., Malthouse, E.C. and Neslin, S.A.: Customer lifetime value: Empirical generalizations and some conceptual questions, *Journal of Interactive Marketing*, Vol.23, No.2, pp.157–168 (2009).
- [31] Ascarza, E., Neslin, S.A., Netzer, O., et al.: In Pursuit of Enhanced Customer Retention Management: Review, Key Issues, and Future Directions, *Cust. Need. Solut.*, Vol.5, pp.65–81 (2018).



蓮本 恭輔 (正会員)

1998年東北大学地球工学科卒業。2010年コーネル大学ジョンソン経営大学院修了(MBA)。2017年早稲田大学大学院創造理工学研究科博士後期課程。2020年早稲田大学データサイエンス研究所招聘研究員。株式会社

SAP ジャパン, アマゾンジャパン株式会社を経て, 現在, PayPal Pte. Ltd. にてマーケティング業務に従事。機械学習のマーケティングへの応用に興味を持つ。



後藤 正幸 (正会員)

1994年武蔵工業大学大学院修士課程修了。2000年早稲田大学大学院理工学博士課程修了。博士(工学)。1997年早稲田大学理工学部助手。2000年東京大学大学院工学研究科助手。2002年武蔵工業大学環境情報学部助教授。

2008年早稲田大学創造理工学部経営システム工学科准教授。2011年同大学教授。情報数理応用とデータサイエンスの研究に従事。著書に、『入門パターン認識と機械学習』コロナ社(2014), 『ビジネス統計～統計基礎とエクセル分析』オデッセイコミュニケーションズ(2015)等。IEEE, 電子情報通信学会, 人工知能学会, 日本経営工学会, 経営情報学会等各会員。