

時間単位 Clumpiness 指標を用いた解析手法の提案

——生存時間分析と LSTM-RNN を用いた利用頻度・離脱時間の予測への RFMC 分析の活用——

新 美 潤一郎

1. はじめに

マーケティング分野においては顧客関係管理 (Customer Relationship Management, CRM) の中で、企業が保有する顧客のロイヤルティを行動的ロイヤルティや態度的ロイヤルティとして把握することが試行されてきた (Jacoby and Chestnut, 1978)。その中でも、顧客の一定期間内の行動を Recency, Frequency, Monetary の3指標を用いて集計する RFM 分析 (Bult and Wansbeek, 1995) は、膨大に蓄積された行動データを簡潔に要約できることから、商品の購買やサービスからの離脱といった顧客の行動予測、さらには顧客生涯価値 (Customer Lifetime Value, CLV) の算出などに広く活用されてきた (Blattberg et al., 2008)。それに加えて、近年では新たに提案された Clumpiness を含む RFMC 分析 (Zhang et al., 2014) が関心を集めており、未だ学術的・実務的双方の側面から検討されるべき課題が残されているもののその活用方法が模索されている。

また、近年では機械学習を用いた行動予測も広く実施されているが、機械学習、特に深層学習手法はその精度の高低を特徴選択に大きく依存しており (Bengio, 2013)、いかに行動ログデータから予測精度を高める特徴量を抽出するかは機械学習の文脈においても重要な課題であ

る。さらに深層学習手法の中でも本研究で扱う Recurrent Neural Network (RNN; Rumelhart et al., 1986) は、行動ログ等に関してより原データに近い形式としての sequential data (連続データ) をそのまま解析に特徴量として投入することのできる手法であるが、RFMをはじめとした代表値を用いた要約を行わずに原データに近い豊富な情報を用いることで予測精度が高まるのか、あるいは次元が増加することから適切な変数を用いることで要約すべきなのか、その計算量と精度の高さの関係は特にマーケティング領域においては未だ不明瞭である⁽¹⁾。本研究では非契約型サービスからの離脱を扱うが、同様に離脱を扱った Spanoudes and Nguyen (2017) では、深層学習がパターン認識手法として出力 (関心のある変数) に影響する変動を適切に認識するという点で離脱予測や特徴選択にも適した手法であることが示されている。

そこで本研究では変数として主に RFMC の各指標を用いつつ、複数の手法による実データ解析を通じて、特に Clumpiness 指標の活用の可能性について統計モデリング・機械学習の双方から検討する。第2節で RFM 分析や Clumpiness 指標を扱った先行研究をレビューし、現状で Clumpiness 指標が抱える問題について整理する。第3節では本研究の提案指標を用いた統計モデリングによる行動予測を実施

(1) マーケティング領域でも大規模データの活用は様々試行されてきたが、その一方で解析にあたって条件抽出を繰り返すことが必要になるため、少なくとも深層学習等の応用領域の中で比較するとサンプルサイズは小さい。

し、その影響について確認する。第4節では機械学習手法を用いた離脱予測を行い、最後に第5節で2つの解析からの考察と今後の課題を述べる。

2. 先行研究

ここでは本研究に関連した分析手法について整理する。

2.1 RFM, RFMC 分析

まずはマーケティング研究として、従来から用いられているRFM分析やClumpinessを用いた分析手法について整理する。

2.1.1 RFM 分析の活用

そもそもRFM分析を基礎として顧客の行動やロイヤルティの高さを推計する手法には、RFM指標から顧客のCLVを予測するBuy Till You Die (BTYD; Schmittlein et al., 1987; Fader et al., 2005a) モデルや、3指標のウェイトを求め、その加重和をCLVとするWeighted RFM (WRFM; Shen and Chuang, 2009)、さらに本研究で扱うRFMC分析のように既存の3指標に別の指標を加えるExtended RFMモデルとして、品目数 (Count Item) を加えたRFMCI (Khajvand and Tarokh, 2011)、利用期間の長さ (Length) を考慮したRFML (Hosseini et al., 2010) などそのバリエーションは多数存在する。これらの先行研究からも分かるように、Extended RFMではいずれの場合でもRFMの3指標のみでは捉えることのできない要因の考慮を目的として提案されている。本研究ではあくまでも単一のゲームアプリを扱い、なおかつ消費者ごとにゲームの採用から30日間のデータを用いるため、品目数や利

用期間は考慮しない。

2.1.2 Clumpiness 指標とその拡張

本研究で扱うClumpinessは期間中の関心のあるイベント (購買, 利用, 訪問, 閲覧等) のまとまった発生 (e.g., 動画配信サービスにおけるbinge-watching⁽²⁾; Zhang, 2013) を捉える指標である。先述のようにExtended RFMでは3指標では捉えきれない要因を考慮してその識別性を高めることを目的に追加的な指標を導入するが、RFMCの場合にはその目的はイベント発生の不均一性を捉えることにある。例として、図1のように3名の顧客A, B, Cについてその一定期間内の行動をR, Fの2指標で集計した場合に、行動的ロイヤルティの観点で見ると顧客Aが他に比べてF指標で劣ることは明白だが、一方で顧客B, Cについては2指標では区別のつかない差異、より具体的には利用行動の発生タイミングが不均一性をもっているかどうかの差異が存在しており、なおかつこの問題はM指標を加えても発生しうる。このような問題への対処を目的として、行動ログ上にイベント発生の不均一性が存在しているかどうかを統計的に検定したのがC指標である。

このClumpiness指標を改善し行動予測等に応用した研究はいくつかあるが、たとえば新美・星野 (2020) では実務的な課題について整理した上で識別性を高める形で拡張したC比率を提案するとともに、ゲーム利用等における熱中の度合い、あるいは実店舗購買における競合他社の利用状況を表す代理変数としてClumpinessを活用できる可能性を示唆している。C比率は利用回数などの条件の異なる顧客間のclumpyさ (不均一性の高さ) の度合いを比較できるよう拡張されており、購買回数や離脱時期の予測にあたって高い説明力を持っている。

(2) シリーズものとして構成されているドラマ等の番組を複数話から全話まとめて視聴すること。

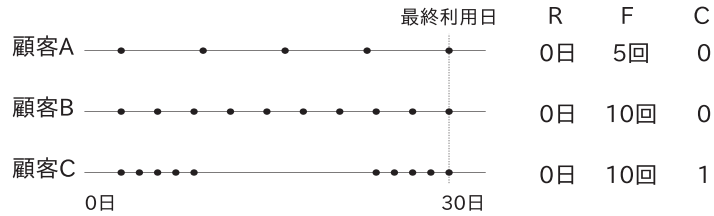


図1 RF指標で区別できない消費者の異質性

ることが示されている。一方で同様にC指標の改善と実データへの適用を行なった Nakayama and Araki (2020) では、C指標とF指標の交互作用項を投入することで予測誤差の減少につながるなどC指標について一定の有用性を評価してはいるものの、Zhang et al. (2014) で提示されたような clumpy な購買行動を行う顧客の将来購買が増加するような傾向については確認できなかったことを指摘している。いずれの既存研究においても今後も継続的にその性質を明らかにする必要性が述べられている。

2.1.3 異なる期間・単位時間での指標の集計

RFM分析で既存研究の多くが1ヶ月を想定した30日間を集計期間として用いていることについて、Fader et al. (2005b) では集計期間の長さや区切る時期によってRFMの各指標の値が変化してしまう解析の恣意性の問題が指摘されており、これはClumpinessについても同様である（新美・星野, 2020）。しかしながらClumpinessの計算にあたっては、それら集計期間による値の変化以上に重要な点として、集計の単位時間の設定の方法による値の変化の問題が考えられる。新美・星野（2020）では、スマートフォンの位置情報ゲームアプリであるPokémon GOのRFMCに基づく利用行動と離脱の関係について分析するにあたり潜在的な利用機会を日次単位（daily）で集計したC指標を解析に用いているが、これは時間単位

（hourly）での集計がアプリの利用行動以外に通勤・通学や睡眠などの生活要因による習慣性の情報を含意してしまうことによるものであるとしている。しかしながら、言い換えるとこの指摘は本来単位時間の設定とは無関係に利用行動のclumpyさの度合いを捉えるべきであるClumpinessが、実際には単位時間の設定次第で異なる情報を含有してしまうことを意味する。星野（2009）はウェブサイトの閲覧ログや購買履歴データといった行動ログデータは一般的に行動の結果のみを収録したデータであり、その行動がどういった嗜好や属性に起因して発生しているかを判別することはログデータのみからでは困難であることを指摘しているが、これは本研究で扱うようなスマートフォンの操作履歴に関しても同様である。つまり、Clumpinessの計算方法における単位時間の設定については、行動の結果でしかないログデータから消費者の異質性を捉えるという点でより詳細な検証が必要である。そこで本研究では、日次単位、時間単位それぞれで算出したClumpiness指標を解析に投入した場合の効果について複数の手法を用いて検証する。

本論文では以後、Zhang et al. (2014) がClumpiness指標を購買と訪問それぞれについて集計した指標をpurchase-Cとvisit-Cとしたことに倣い、単位時間を日次単位/時間単位で集計したC指標をそれぞれdaily-C/hourly-C（表内では略称としてdC指標/hC指標）、さらに新美・星野（2020）のC比率を同様に異な

る単位時間で集計したものをそれぞれ daily-Cratio/hourly-Cratio (同様に dCr 指標 /hCr 指標) と表記する。

ここで本研究で用いる hourly-C, hourly-Cratio を定義しておく、その算出方法としては daily-C, daily-Cratio が 30 日間の利用状況を日次の潜在機会 $N_d=30$ に離散化して集計するのと同様に、30 日間に対して時間単位の潜在機会 $N_h=24N_d$ 時間の潜在的な利用機会を用いて算出する。本研究で扱うスマートフォンゲームの利用状況のように、一般的な商品購買などとは異なり 1 日に複数回の利用が行われることもおかしくないような市場では、時間単位での集計を用いることでより豊富な情報を Clumpiness 指標として保持できる可能性がある。

2.1.4 C 比率の定義

本研究では既存研究と同様に C 比率の効果についても検証するが、そもそも C 比率を提案した新美・星野 (2020) の問題意識は、C 指標の算出過程 (Zhang et al., 2014, Appx.B) で用いられる H 指標と C 指標について、H 指標ではサンプル全体 Ω について、イベントの発生機会 N 、実際のイベント発生回数 n の一致した消費者同士 ($N_i=N_j$ かつ $n_i=n_j$, ただし $i \neq j$ かつ $i, j \in \Omega$) でしかその偏りの大きさを比較できないという問題があるが、一方でそれらの条件間で比較可能な C 指標は離散化されていることから大きな情報損失を伴うという点から出発しており、そこで C 比率を、C 指標の算出に用いる境界値 H_0 を用いて、

$$Cratio_i = \begin{cases} \frac{H_i}{H_0} & (H_i > H_0) \\ 1 & (otherwise) \end{cases} \quad (1)$$

と定義している (ここで $H_0=0$ となるのは全ての購買機会で購入を行う $n_i=N_i$ の場合であ

る)。本来の C 指標が H 指標を用いた N, n の条件下での不均一性の検定の結果だとすれば、C 比率は異なる N, n にわたる H 指標の標準化と考えることができる。

2.2 LSTM-RNN

次に本研究で用いる RNN や LSTM の構造や、それらを用いたマーケティング分析の先行研究を整理する。

2.2.1 RNN の構造

RNN は深層学習の一手法として、最も一般的な深層ニューラルネットワーク (Deep Neural Network, DNN) たる順伝播型ネットワーク (Feed-Forward Neural Network, FFNN) の機能に加え、帰還路 (recurrent connections) を用いた再帰的な入力が行われることが特徴である。RNN に入力する特徴量はタイムステップ $t=\{1, \dots, \tau\}$ について $x=(x_1, \dots, x_\tau)$ が逐次的に入力される。RNN のタスクには入力した連続データから単一の値を予測値として回帰 / 分類する Sequence-to-One (S2O) 型と、出力も連続データとして入力に対して随時回帰 / 分類の出力を行う Sequence-to-Sequence (S2S) 型が存在する。

帰還路を持つ t 期の隠れ層 h_t の状態は、 t 期の入力 x_t 、バイアス b_s 、ウェイト W_s 、活性化関数 f_s を用いて (ただし $s=\{i, r, o\}$ はそれぞれ入力、再起、出力を表す。以下の式(2)からもわかるように単純な構造の RNN では入力と再起に加重和を取り同一の活性化関数を適用する)、

$$h_t = \begin{cases} f_i(b_i + W_i x_t) & (t=1) \\ f_r(b_r + W_r x_t + W_r h_{t-1}) & (otherwise) \end{cases} \quad (2)$$

と表現でき⁽³⁾、つまり S2O 回帰での出力 o は活性化関数 f_o を用いて

$$o = f_o(b_o + W_o h_\tau) \quad (3)$$

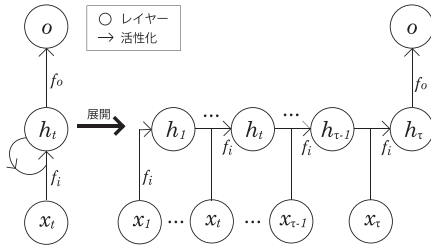


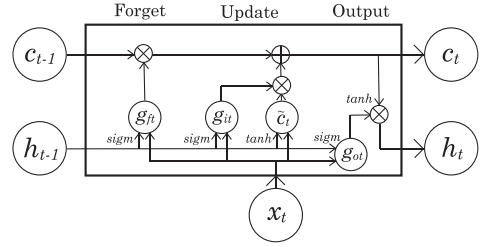
図2 S2O型RNNの構造と展開

と表すことができる⁽⁴⁾。

式(2)の隠れ層の表現からもわかるように、帰還路を持つS2O型RNNの構造は図2⁽⁵⁾のように展開することができる。これにより、RNNを時系列的な入力をラグとして扱う一つのFFNNとみなして、勾配計算（通時的誤差逆伝播法：Backpropagation Through Time, BPTT; Williams and Zipser, 1995; Werbos, 1988）を行うことが可能である⁽⁶⁾。つまり、FFNNと同様にパラメータの最適化には確率的勾配降下法（stochastic gradient descent, SGD）等の一般的な最適化手法を適用可能となる。

2.2.2 LSTMの構造

ここまででsequential dataを扱うことが可能なRNNの構造について述べた。しかしながら、その活用の中でRNNは長期の時系列的な依存関係を把握することが困難であることが指摘されており（Bengio et al., 1994）、これは一つには深層学習においてよく知られている勾配

図3 t 期 ($t > 1$) のLSTMセルの構造

消失問題（vanishing gradient problem; Hochreiter, 1998）に起因している。逆伝播による勾配計算の中では非線形関数、特にシグモイド関数による度重なる活性化により勾配の爆発や消失が発生することが知られているが、RNNでは図2の展開で示した通り長期的に時系列を遡ることで隠れ層を増やすと同様にネットワーク構造が深化するため、必ずしも隠れ層の多くない学習器でも長期のsequential dataを用いることで勾配消失が発生しやすい（岡谷, 2015）。

こういった問題を解決し、RNNで長期的な依存関係を捉えるための手法としてLong Short-Term Memory (LSTM; Hochreiter and Schmidhuber, 1997) 層が提案されている。LSTMでは内部に単位時間ごとに状態を変化させる「セル」と呼ばれるニューラルネットワーク的な構造を持ち、時間 t について隠れ層 h_t 以外にセル状態（cell state） c_t が存在していることが特徴である。

ここで t 期のLSTMセルの構造を図3⁽⁷⁾に

(3) 深層学習の隠れ層に適用する活性化関数には双曲線正接関数 $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ 等を用いることが一般的であることから、本研究でも $\tanh$ を採用している。

(4) 出力に適用する活性化関数には分類タスクとして softmax 関数等が用いられることが一般的である。本研究では回帰タスクとして離脱日数を予測するにあたり、非負の正規化線形関数 $\text{ReLU}(x) = \max(0, x)$ が出力の形態として合致していると判断した。

(5) Goodfellow et al. (2016) を参考に著者が作成。

(6) BPTT 以外に Realtime Recurrent Learning (RTRL; Williams and Zipser, 1989) を用いた計算方法も提案されている。

示すと、LSTMでは単純な再帰的構造とは異なり、忘却 (forget), 更新 (update), 出力 (output) という3つのプロセスとそのための「ゲート」が設定されており、そのゲートにより情報をどの程度保持するかが決定される。具体的には忘却ゲート、入力ゲート、出力ゲート、セル候補により制御されており、まず忘却ゲートで $t-1$ 期のセル状態から必要な情報を取捨選択、入力ゲートでどの情報を更新するかを決定、セル候補を構築してセル状態を更新、最後に更新されたセル状態に基づいて出力 h_t を得る。

t 期のそれぞれのゲート g_{st} の状態はウェイト W_s 、再帰ウェイト R_s 、バイアス b_s (ただし $s=\{f, i, o, c\}$ はそれぞれ忘却、入力、出力、セル候補を表す) とシグモイド関数 sigm を用いて

$$g_{st} = \text{sigm}(W_s x_t + R_s h_{t-1} + b_s) \quad (4)$$

同様にセル候補は双曲線正接関数 $\tanh$ を用いて、

$$\tilde{c}_t = \tanh(W_c x_t + R_c h_{t-1} + b_c) \quad (5)$$

これらから t 期のセル状態と隠れ層は、

$$c_t = g_{ft} \odot c_{t-1} + g_{it} \odot \tilde{c}_t \quad (6)$$

$$h_t = g_{ot} \odot \tanh(c_t) \quad (7)$$

と表すことができる。ここで $\odot$ は要素ごとの積を表す。この式からもわかるように、LSTMのゲートは基本的には入力に対する加重和をシグモイド関数で活性化する構造であり、それらを用いて過去の情報をどの程度保持するかが決

定されている。この構造を導入した LSTM-RNN により、単純な構造の RNN と比較しても入出力の長期的な関係を捉えることが可能となり (Goodfellow et al., 2016), 特に自然言語処理や機械翻訳などの分野で離れた単語間の関係を捉えられるといった理由から高い精度を見せている。

2.2.3 LSTM-RNN を用いた解析

LSTM-RNN を用いたマーケティング分野における研究として、まず S2O 型を用いたものには、複数の LSTM セルや確率的に出力を無効にする Dropout (Srivastava et al., 2014) を組み合わせた構造の RNN を用いて、EC サイトの Clickstream データからあるセッションが購買・ショッピングカートの放置・閲覧のみのどの段階で終了するかを予測した Toth et al. (2017)⁽⁸⁾ や、契約型の金融サービスにおいて月次の RFM 指標を連続データとして用いて離脱を予測する Mena et al. (2019) などがある。

一方で S2S 型の活用はほとんど見られないが、RNN をアンサンブル学習におけるメタモデルとして用いて、複数の機械学習手法をスタッキング (i.e., stacked generalization; Wolpert, 1992) することで CLV を予測した Bauer and Jannach (2021) などがある。Bauer らで特徴的な点として、S2S 回帰を他の予測手法とスタッキングするにあたり、RNN では通常の S2S 回帰で連続的に予測値を算出している一方で、他のメタモデルとして用いている Gradient Boosting (Friedman, 2001) では時系列データのラグを取り、入力する説明変

(7) Hochreiter and Schmidhuber (1997), MathWorks lstmLayer (<https://jp.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.lstmLayer.html>) を参考に著者が作成。

(8) RNN の中で複数の LSTM ユニットを直列的に接続して S2O 型の出力を行う場合に、いずれの LSTM ユニットへの入力でも sequential data をとるためには、出力に最も近いユニットのみを S2O 型、その他のユニットは S2S 型に設定することが必要である。

数のタイムステップをずらして投入することで疑似的に S2S 回帰のようにシーケンシャルな出力を行い、それらをまとめてスタッキングしている点がある。

3 解析 1：統計解析を用いた利用状況の予測

本研究では daily-C, hourly-C それぞれの与える影響について実データを用いて検証するにあたり、スマートフォンゲームを対象に利用状況や離脱に関する解析を実施するが、まずは解析 1 として一般化線形モデル (GLM) や生存時間分析を用いたモデリングを行う。

3.1 使用データ

本節で使用するのは株式会社インテージの i-SSP モバイル調査で収集されたログデータである。2016 年から 2017 年までの 2 年間にわたってスマートフォンのウェブサイト閲覧、アプリ使用、PC のウェブサイト閲覧等について人口統計に基づいて収集した大規模なパネルデータである。本研究では対象となるゲームアプリとして先行研究と同様に Pokémon GO を取り上げる。このタイトルを取り上げるのは、社会現象として多くの利用者が存在すること、そしてデータとして得られている期間中にサービスを開始したことの二点による。特に本研究のようにログデータを用いた生存時間解析を行うにあたっては、離脱を観測できないままデータ期末まで到達した場合の右側打ち切りに加え、データ期首の時点より前からサービスを利用している場合の左側打ち切りが存在するが、

サービス開始時点からの解析では左側打ち切りの考慮が不要となる。よってサービス開始日である 2016 年 7 月 22 日を期首として 2017 年 12 月 31 日までのデータを利用している。期首から 30 日以内に利用を開始し、なおかつ少なくとも 3 回以上利用しているユーザーを解析の対象としており、サンプルサイズは 553 人となった。なお、先行研究と同様に利用開始から 30 日間利用がなかったことをもって離脱と判断し、離脱からの復帰は考慮しない。

本研究ではスマートフォンの利用ログを用いている都合上、ゲームへの支出額としての M 指標は取得することができない。とはいえ、スマートフォンで提供されるゲームは Free-to-Play (F2P)⁽⁹⁾ の形態を取っていることが多く、利用者の大半が無償で利用していることから、そのロイヤルティや離脱時間の予測にあたっては M 指標よりも利用時間の方が説明力が高くなる可能性がある。実際 Yan and Chen (2011) ではこういったアプリケーションの利用状況の指標化にあたり M 指標の代わりに D 指標 (duration, 利用時間) を活用した RFD モデルでユーザーごとの好みの度合いを把握することが可能であることを示しており、本研究でも各解析には M 指標の代わりに D 指標を用いる⁽¹⁰⁾。ただし Yan and Chen (2011) では 3 指標の加重和をとって一つの指標としているが、本研究では 3 指標それぞれの影響を確認するため独立した指標として用いる。

解析に先立って、表 1 に解析 1 で用いる 30 日間のデータから集計した各指標の相関係数を示す。まず dC 指標では弱い正の相関関係に

(9) ゲーム内で提供される機能のうち基本的なものは無償で利用できるものの、発展的なものや短時間でゲームを進めるための特典等の利用は有償となるような利用形態のこと。

(10) 一部の文献では D 指標として利用 / 購買 / 訪問等の平均間隔を用いているが、本研究のように集計期間を全顧客について N 日間に揃えると、消費者 i の平均利用間隔 $D_i = \frac{N}{F_i}$ となり F 指標と同じ情報しか持たないことから本研究では用いていない。

表1 RFDCの各指標間の相関係数

	R 指標	F 指標	D 指標	dC 指標	hC 指標	dCr 指標	hCr 指標	離脱日数
R 指標	1.00	-0.36	-0.34	0.15	0.40	0.39	0.49	-0.33
F 指標	-0.36	1.00	0.76	0.26	-0.41	-0.24	-0.32	0.40
D 指標	-0.34	0.76	1.00	0.25	-0.44	-0.24	-0.33	0.34
dC 指標	0.15	0.26	0.25	1.00	0.03	0.26	0.14	0.01
hC 指標	0.40	-0.41	-0.44	0.03	1.00	0.52	0.72	-0.31
dCr 指標	0.39	-0.24	-0.24	0.26	0.52	1.00	0.72	-0.25
hCr 指標	0.49	-0.32	-0.33	0.14	0.72	0.72	1.00	-0.31
離脱日数	-0.33	0.40	0.34	0.01	-0.31	-0.25	-0.31	1.00

あったF, D指標だが, dCr指標では符号が転じている点がある。これはH指標からdC指標への検定のプロセスの中で, 二値化による情報損失が発生している可能性を示唆している。

3.2 解析1-1: 7日間の利用状況を用いた30日間の利用状況の予測

まず解析1として, 7日間の利用状況をもって30日間の利用状況を予測するモデルを複数作成して比較する。

3.2.1 解析の目的と概要

この解析の問題意識として, 第一に従来の解析で一般的に行われているような30日間のデータを用いた利用状況の予測ではサービスの利用開始からごく短期間で離脱するトライアルユーザーを識別できないという点がある。第二にClumpinessとの兼ね合いとして, 既存研究では一般的に日次単位でC指標の計算を行なっているが, 集計の日数を短くした場合たとえば1週間一には, daily-Cの潜在的な購買機会 $N_d=7$ となってしまう, C比率を用いてもなお消費者の利用パターンの識別が不十分となることが考えられる。こういった場合に識別に

十分な購買機会としてhourly-Cが有用となる可能性がある。そこで本研究では, RFMC分析の中でも特にC指標について, その集計の単位時間を変更した場合に解析結果に与える影響について実データを用いて検証する(R, F, D指標については集計期間が変わらない限りは単位時間を変更しても値は変化しない)。

解析にあたっては7日間の利用状況から30日間の利用状況を予測するため, 一般化回帰分析(GLM)による一定期間内のイベント発生回数のモデリング手法としてマーケティング分野において盛んに利用されているポアソン回帰分析を用いる。

3.2.2 解析結果

表2に作成した各モデルのパラメータ推定値と情報量規準を示す。まず全体を俯瞰するとベースライン(Model 1)と比較してそれ以外のいずれの場合にも情報量規準は改善している。C指標とCr指標の効果を比較すると, daily, daily+hourlyの2つに関してはCratioを用いた場合に情報量規準が改善しているが, その一方でhourlyに関してはModel 6よりもModel 3の方が情報量規準が改善する結果と

表2 解析 1-1：パラメータ推定値の比較

モデル	Model 1 (Baseline)	Model 2 (dC)	Model 3 (hC)	Model 4 (dC, hC)	Model 5 (dCr)	Model 6 (hCr)	Model 7 (dCr, hCr)
定数項	1.75 ***	1.60 ***	1.67 ***	1.51 ***	1.50 ***	1.72 ***	1.54 ***
R 指標	-0.37 ***	-0.34 ***	-0.39 ***	-0.36 ***	-0.48 ***	-0.38 ***	-0.47 ***
F 指標	0.84 ***	0.83 ***	0.84 ***	0.83 ***	0.85 ***	0.84 ***	0.85 ***
D 指標	0.02 ***	0.01	0.02 ***	0.01 ***	0.03 ***	0.02 ***	0.03 ***
dC 指標	—	0.33 ***	—	0.32 ***	—	—	—
hC 指標	—	—	0.05 ***	0.05 ***	—	—	—
dCr 指標	—	—	—	—	0.22 ***	—	0.23 ***
hCr 指標	—	—	—	—	—	0.01 ***	-0.02 ***
AIC	51231.65	50791.44	51161.09	50720.86	50510.97	51218.33	50489.66
BIC	43987.63	43551.75	43921.40	43485.50	43271.28	43978.64	43254.30

Note. 太字はその情報量基準を用いた場合に最も当てはまりの良いモデルを表す。

* $p < .10$. ** $p < .05$. *** $p < .01$.

なっている。この点については、C 指標、Cr 指標それぞれのグループ (Model 2-4, 5-7) での daily と hourly の効果を比較した場合に、いずれのグループにおいても確認できる傾向として hourly-C は単体 (Model 3, 6) で用いると、有意水準こそ 1% で満たしてはいるものの、情報量規準の改善幅としてはそもそも小さいことがわかる。それでも各グループにおいて daily に加えて hourly を投入した場合 (Model 4, 7) で最も改善していることから、daily での集計に加えて hourly を補助的な指標として投入することにより予測力が向上することが示唆される。ただし、dCr、hCr をともに投入した Model 7 では hCr のパラメータ推定値が逆転していることには留意が必要である。

3.3 解析 1-2：30 日間の RFMC から離脱時間を予測

次に 30 日間の利用状況から離脱時間を予測するモデルを作成する。

3.3.1 解析の目的と概要

解析には Cox 比例ハザードモデル (Cox, 1972) を用いる。この手法は購買時期や離脱時期といったイベント発生時期のモデリングを目的に、マーケティング分野では広く用いられている。たとえばロイヤルティプログラムが顧客の離脱時期や支出額に及ぼす影響を調べた Meyer-Waarden (2007) や、顧客の金融サービスからの離脱をモデリングした Van den Poel and Lariviere (2004) など多数存在する。

30 日間の RFD を投入したモデルを基準として、そこに Clumpiness の関連指標を投入することによるパラメータ推定値や情報量規準の変化を確認する。サンプルサイズは 553 人で、期間中に離脱したのは 438 人であった。

3.3.2 解析結果

表 3 にモデルごとのパラメータ推定値を示す。まず C 指標と Cr 指標を比較すると、いずれの場合 (Model 5, 6, 7) にも Cratio を用いた場合の方が情報量規準が改善していることがわかる。ほとんどの場合でベースライン (Model

表3 解析1-2：パラメータ推定値の比較

モデル	Model 1 (Baseline)	Model 2 (dC)	Model 3 (hC)	Model 4 (dC, hC)	Model 5 (dCr)	Model 6 (hCr)	Model 7 (dCr, hCr)
R 指標	0.36 ***	0.34 ***	0.30 ***	0.28 ***	0.24 ***	0.21 ***	0.19 ***
F 指標	-0.07	-0.10	-0.07	-0.10	-0.10	-0.08	-0.09 ***
D 指標	-0.04	-0.04	-0.04	-0.04	-0.06	-0.08	-0.08 ***
dC 指標		0.14		0.14 ***	—	—	—
hC 指標			0.33 ***	0.34 ***	—	—	—
dCr 指標					0.29 ***		0.17 *
hCr 指標						0.19 ***	0.13 **
AIC	4717.44	4717.89	4710.47	4710.82	4705.70	4704.10	4702.86

Note. 太字はその情報量基準を用いた場合に最も当てはまりの良いモデルを表す。

* $p < .10$. ** $p < .05$. *** $p < .01$.

1) よりは改善する結果となっているが、Model 2 に関してのみ基準となる Model 1 よりも情報量規準は漸増している。この点については、dCr を用いた場合 (Model 5, 7) では有意水準 1% を満たして離脱を早める方向に影響していることや、それらのモデルで情報量規準が改善していることから、clumpy な行動を日次レベルで捉えること自体は有用である一方で、Cratio としてその程度を捉えることが必要であることを示している。最終的に daily-Cratio, hourly-Cratio を同時に投入した場合に情報量規準は最も改善する結果となった。

ここで留意すべき点として、解析 1-1 とは異なり dCr を用いた場合 (Model 2, 5) の改善幅が hCr を用いた場合 (Model 3, 6) よりも小さくなっていることがわかる (それでも dCr 単体で dC+hC の場合よりもモデルは適合している)。また、この解析では hourly-Cratio の効果はいずれも clumpy な行動が離脱を早める方向に影響している。

4 実データ解析 2：機械学習と連続データでの活用

次に解析 2 として RNN を用いた顧客のサービスからの離脱有無の予測を行う。

4.1 解析の目的と概要

本解析で RNN を用いることの問題意識は、行動ログデータの要約の度合いと予測精度の関係を調べることにある。そもそもマーケティング領域に限らず統計解析全般において行動ログデータ (モバイルアプリの使用時間等) を用いる場合、原データを図 4 右端の月次単位のパネルデータまで要約した上で解析に投入することが一般的であるが、近年では sequential data の投入がより容易に実現できるようになったことにより、様々な形式のデータを活用することが可能となってきている。データの要約に関しても、その程度によって原データから月次単位の集計データまでの表現は図 4 に示すように異なっている。本研究で用いる sequential data は、RNN で扱いやすい構造である時間単位 (対象×日付×時間)、日次単位 (対象×日付) の集計データを指す。

原データ			集計データ（時間単位）					集計データ（日次単位）				集計データ（月次単位）				
ID	datetime	D	ID	date	hour	F	D	ID	date	F	D	ID	year	month	F	D
1001	2017/09/01 09:00:00	20	1001	2017/09/01	9	2	20	1001	2017/09/01	7	208	1001	2017	9	240	6451
1001	2017/09/01 09:12:35	82	1001	2017/09/01	10	1	82	1001	2017/09/02	8	321	1002	2017	9	186	4917
1001	2017/09/01 10:34:12	56	1001	2017/09/01	12	1	38	1001	2017/09/03	5	120	1003	2017	9	311	8020
1001	2017/09/01 12:07:28	38	1001	2017/09/01	13	3	68	1001	2017/09/05	12	580	1004	2017	9	247	5660

図4 データの要約の程度による表現の差

代表値を用いたデータの要約は、それにより情報が毀損しているとも考えられるが、一方でデータの次元の増加は次元の呪い (Curse of dimensionality; Stone, 1982) と称されるように、それにより必要となるサンプルサイズを指数関数的に増加させることや、必要な学習時間や計算資源の増加にも繋がる。したがって現段階ではその要約の程度とデータの豊富さはトレードオフの関係下にあるが、ここで本研究で用いている Clumpiness を活用することで、日次データのみでも十分な予測精度を担保できる可能性がある。

そこで、本解析では同じ原データを元に、要約の程度をいくつかに分けたデータを作成し、それぞれを入力として用いた場合の予測精度の変化を確認する。類似した解析を行なっている Mena et al. (2019) では LSTM-RNN を用いた場合に正則化ロジスティック回帰よりも高い精度で予測が可能であることが示されていることから、本研究でもベースラインとして L2 正則化ロジスティック回帰分析を採用する。

手法の比較にあたっては、学習データ、テストデータそれぞれでの Accuracy（正答率）、Precision（適合率）、Recall（再現率）、F 値を用いるが、評価にはそのうち Accuracy と F 値を用いる。Precision と Recall についてはその定義上どちらか一方を高めることは容易に可能であり⁽¹¹⁾、相反する関係にあるそれら 2 指

標の調和平均たる F 値で評価することが必要となるためである。また、データ全体からランダムに 70% を学習用、15% を検証用、15% をテスト用として分割し、学習済み推定器をテストデータに適用した場合の Accuracy, F 値の大小を算出、それらを汎化誤差として評価する。

4.2 解析結果

表4に解析結果を示す。まず手法内での比較を行うと、RNN では hourly を基準にした場合に Cratio まで段階的に汎化誤差が改善する結果となった。一方のロジスティック回帰では学習時で改善幅は微増、テスト時にはほぼ変わらないという結果になった。手法間では、Cratio を用いたロジスティック回帰では汎化誤差が部分的に RNN のそれを超える場合はあるものの、全体として汎化誤差が最も改善したのは RNN (Cratio) の場合であった。また、データの粒度を細かくして情報量が増加したと考えられる RNN (hourly) では、学習に要する時間も長くなった上にその精度も大きく低下した。

ここで全体的な傾向として、学習時の評価値が高いことが決して汎化誤差を縮小することにつながっていない点がある。学習時に他の手法の方が優れていた予測精度も、テスト時には C 比率を用いた場合に最も精度が向上する結果となった。

また結果には LSTM セルのサイズが 5,

(11) 例として、 $y=\{0, 1\}$ の分類問題において全ての観測値に対して 1 と予測した場合に Recall の値は上限値 1 となるが、一方で Precision は極端に低下する。

表4 解析2: RMSE の比較

	投入変数	Train				Test			
		Accuracy	F-value	Precision	Recall	Accuracy	F-value	Precision	Recall
Random	—	0.453	—	—	—	0.335	—	—	—
RNN (hourly)	hR, hF	0.502	0.459	0.452	0.466	0.661	0.553	0.520	0.591
RNN(Baseline)	dR, dF	0.740	0.740	0.677	0.817	0.710	0.679	0.559	0.864
RNN (C)	dR, dF, dC	0.751	0.733	0.712	0.756	0.758	0.681	0.640	0.727
RNN (Cratio)	dR, dF, dCr	0.734	0.724	0.682	0.771	0.774	0.696	0.667	0.727
Logistic(base)	R, F	0.713	0.694	0.707	0.681	0.742	0.680	0.654	0.708
Logistic (C)	R, F, dC	0.713	0.694	0.707	0.681	0.742	0.680	0.654	0.708
Logistic(Cratio)	R, F, dCr	0.720	0.699	0.718	0.681	0.742	0.680	0.654	0.708

Note. 太字はその評価指標で最も精度が高いことを示している。結果はRNN (hourly) ではEpoch 10, それ以外のRNN ではEpoch 5で学習した場合であり, これらネットワークの構造や学習方法の妥当性については4.2節を参照。

Epoch 数が5の条件で学習を行った精度を示しているが, DNNの構造の決定にあたっては予めEpoch数や隠れ層のサイズを変更した条件で網羅的な解析を実施している。たとえばEpoch数を1万程度まで上げることでRNN (Baseline) のAccuracy (Train) が0.90程度まで改善する結果も得られたが, いずれの場合にもそれに対して汎化誤差が大きく増加しており, 網羅的な解析を行った中でも表に示すRNN (Cratio) で汎化誤差が最も改善した。

5 まとめ

本節では以上の解析結果をまとめてその知見を示した上で, 本研究の限界について述べる。本研究では非契約型サービスからの利用・離脱の分析として, 時間単位で集計したClumpiness指標 (hourly-C指標) を活用して, GLMや深層学習を用いた解析例を示した。まず全ての解析に通じる傾向として, 当初からdaily-Cとは異なる情報を持っていると推察されていたhourly-C指標を解析に追加的に用いることでモデルの当てはまりや予測の精度が向

上することが示された。ただし, パラメータ推定値を確認すると特にhourly-Cratioにおいて符号が逆転する場合があることから, 変数として持つ意味合いについては今後も検討していくことが必要である。

解析1の2つの解析結果では, 解析1-1での月間の利用頻度の増加と解析1-2での早期離脱は一見すると相反する結果とも捉えられるが, これは利用行動が短期的か長期的かの差として理解することができる。つまりこの解析結果は短期的な利用行動としての月間の利用頻度は増加するものの, 長期的に見た場合に離脱を早期化させる可能性を示している。2.3節で述べた通り, 本研究で用いているRFDモデルに限らずRFM分析全般に関して集計期間による解析の恣意性は以前から指摘されている問題であり, 本研究のように複数の手法を用いて多面的に解析することや, 数ヶ月にわたる長期的な利用行動を解析に用いることの重要性を示しているといえる。

また, 解析2でも示した通りRNNは原データに近いsequential dataを直接扱えることから, hourly単位の豊富なデータを用いた場合

に精度が高まることが予想されたが、実際には日次単位での集計データと比べて十分な精度が出ず、日次データを Clumpiness 指標を用いて要約した場合に予測精度がより改善するという結果が得られた。この点については、次元の呪いとして、計算量が飛躍的に増加した一方で計算に十分なサンプルサイズが得られなかったことによるものではないかと予想される。たとえば本研究で扱った Clumpiness 指標ではないものの、それに類似した要約手法である多様性変数を用いて行動ログを要約して深層学習で予測した場合にも、データの効率的な要約により予測精度が向上する類似した結果は確認されている（新美・星野，2017）。また、解析結果でも述べたように、無闇な Epoch 数の増加は in-sample での学習時の精度のみが上がり、汎化誤差はむしろ悪化するという点で学習データへの過適合に陥りやすいため、ネットワーク構造や学習方法の決定にあたっては事前に十分な検討が必要である。

最後に本研究の課題として、本研究では時間単位で集計した Clumpiness 指標を従来の日次集計の指標に加えることを検討したが、指標の高低による心理的ロイヤルティへの影響までは明らかにできていない。今後は調査を組み合わせるなどして clumpy な行動を行う消費者の心理的側面との関係についても検証する必要がある。

また本研究の LSTM-RNN ではあくまで一般的な予測精度の算出を目的として最適化には一般的な SGD を用いており、その他先行研究で用いられているような Dropout や複数の LSTM を重ねた構造なども採用していないことから、予測精度の向上のみを目的とする場合にはアーキテクチャのさらなる検討により改善できることが予想される。

謝 辞

本研究の実施にあたり株式会社インテージより i-SSP モバイル調査の貴重なデータをご提供いただきましたこと御礼申し上げます。また名古屋大学大学院経済学研究科の後輩諸君より有益なコメントをいただきました。本研究はデータ取得にあたり名古屋大学 未来エレクトロニクス創成加速 DII 協働大学院プログラムの助成を受けています。

参考文献

- Bauer, Josef and Dietmar Jannach (2021) “Improved Customer Lifetime Value Prediction With Sequence-To-Sequence Learning and Feature-Based Models,” *ACM Transactions on Knowledge Discovery from Data (TKDD)*, Vol. 15, No. 5, pp. 1–37.
- Bengio, Yoshua (2013) “Deep learning of representations: Looking forward,” in *International Conference on Statistical Language and Speech Processing*, pp. 1–37, Springer.
- Bengio, Yoshua, Patrice Simard, and Paolo Frasconi (1994) “Learning long-term dependencies with gradient descent is difficult,” *IEEE transactions on neural networks*, Vol. 5, No. 2, pp. 157–166.
- Blattberg, Robert C, Byung Do Kim, and A. Neslin Scott (2008) *Database Marketing: Analyzing and Managing Customers*: Springer.
- Bult, Jan Roelf and Tom Wansbeek (1995) “Optimal selection for direct mail,” *Marketing Science*, Vol. 14, No. 4, pp. 378–394.
- Cox, David R (1972) “Regression models and life-tables,” *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34, No. 2, pp. 187–202.
- Fader, Peter S, Bruce GS Hardie, and Ka Lok Lee (2005a) ““Counting your customers” the easy way: An alternative to the Pareto/NBD model,” *Marketing science*, Vol. 24, No. 2, pp. 275–284.
- Fader, Peter S, Bruce GS Hardie, and Ka Lok Lee (2005b) “RFM and CLV: Using iso-value curves for customer base analysis,” *Journal of*

- marketing research*, Vol. 42, No. 4, pp. 415–430.
- Friedman, Jerome H (2001) “Greedy function approximation: a gradient boosting machine,” *Annals of statistics*, pp. 1189–1232.
- Goodfellow, I, Y Bengio, and A Courville (2016) “Sequence modeling: recurrent and recursive nets,” in *Deep Learning*: Cambridge: MIT press, Chap. 10, pp. 367–415.
- Hochreiter, Sepp (1998) “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 6, No. 02, pp. 107–116.
- Hochreiter, Sepp and Jürgen Schmidhuber (1997) “Long short-term memory,” *Neural computation*, Vol. 9, No. 8, pp. 1735–1780.
- Hosseini, Seyed Mohammad Seyed, Anahita Maleki, and Mohammad Reza Gholamian (2010) “Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty,” *Expert Systems with Applications*, Vol. 37, No. 7, pp. 5259–5264.
- Jacoby, Jacob and Robert W Chestnut (1978) *Brand loyalty: Measurement and management*: John Wiley & Sons Incorporated.
- Khajvand, Mahboubeh and Mohammad Jafar Tarokh (2011) “Estimating customer future value of different customer segments based on adapted RFM model in retail banking context,” *Procedia Computer Science*, Vol. 3, pp. 1327–1332.
- Mena, C Gary, Arno De Caigny, Kristof Coussement, Koen W De Bock, and Stefan Lessmann (2019) “Churn Prediction with Sequential Data and Deep Neural Networks. A Comparative Analysis,” *arXiv preprint arXiv: 1909.11114*.
- Meyer-Waarden, Lars (2007) “The effects of loyalty programs on customer lifetime duration and share of wallet,” *Journal of retailing*, Vol. 83, No. 2, pp. 223–236.
- Nakayama, Yuji and Nagateru Araki (2020) “An Exploratory Study on the Clumpiness Measure of Intertransaction Times: How Is It Useful for Customer Relationship Management?” in *Advanced Studies in Classification and Data Science*: Springer, pp. 329–339.
- Van den Poel, Dirk and Bart Larivière (2004) “Customer attrition analysis for financial services using proportional hazard models,” *European journal of operational research*, Vol. 157, No. 1, pp. 196–217.
- Rumelhart, David E, Geoffrey E Hinton, and Ronald J Williams (1986) “Learning representations by back-propagating errors,” *nature*, Vol. 323, No. 6088, pp. 533–536.
- Schmittlein, David C, Donald G Morrison, and Richard Colombo (1987) “Counting your customers: Who are they and what will they do next?” *Management science*, Vol. 33, No. 1, pp. 1–24.
- Shen, Chia-Cheng and Huan-Ming Chuang (2009) “A study on the applications of data mining techniques to enhance customer lifetime value,” *WSEAS Transactions on Information Science and Applications*, Vol. 6, No. 2, pp. 319–328.
- Spanoudes, Philip and Thomson Nguyen (2017) “Deep learning in customer churn prediction: unsupervised feature learning on abstract company independent feature vectors,” *arXiv preprint arXiv: 1703.03869*.
- Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov (2014) “Dropout: a simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, Vol. 15, No. 1, pp. 1929–1958.
- Stone, Charles J (1982) “Optimal global rates of convergence for nonparametric regression,” *The annals of statistics*, pp. 1040–1053.
- Toth, Arthur, Louis Tan, Giuseppe Di Fabbri, and Ankur Datta (2017) “Predicting shopping behavior with mixture of RNNs,” in *eCOM@SIGIR*.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin (2017) “Attention is all you need,” in *Advances in neural information*

- processing systems*, pp. 5998–6008.
- Werbos, Paul J (1988) “Generalization of backpropagation with application to a recurrent gas market model,” *Neural networks*, Vol. 1, No. 4, pp. 339–356.
- Williams, Ronald J and David Zipser (1989) “A learning algorithm for continually running fully recurrent neural networks,” *Neural computation*, Vol. 1, No. 2, pp. 270–280.
- Williams, Ronald J and David Zipser (1995) “Gradient-based learning algorithms for recurrent networks,” *Backpropagation: Theory, Architectures, and Applications*, pp. 433–486.
- Wolpert, David H (1992) “Stacked generalization,” *Neural networks*, Vol. 5, No. 2, pp. 241–259.
- Yan, Bo and Guanling Chen (2011) “AppJoy: personalized mobile application discovery,” in *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pp. 113–126.
- Zhang, Yao (2013) “New clumpiness measures and their application in customer evaluation,” *Journal of Applied Statistics*, Vol. 40, No. 11, pp. 2533–2548.
- Zhang, Yao, Eric T Bradlow, and Dylan S Small (2014) “Predicting customer value using clumpiness: From RFM to RFMC,” *Marketing Science*, Vol. 34, No. 2, pp. 195–208.
- 岡谷貴之 (2015) 『深層学習』, 講談社.
- 新美潤一郎・星野崇宏 (2017) 「顧客行動の多様性変数を利用した購買行動の予測—Deep Learning を応用した実店舗・Web・モバイルの多面的な分析—」, 『人工知能学会論文誌』, 第 32 卷, 第 2 号, B-G63 1–9 頁.
- 新美潤一郎・星野崇宏 (2020) 「RFMC 分析における Clumpiness 指標の拡張と自社顧客の行動予測への応用: Clumpiness を活用した離脱時期と競合利用の予測手法の提案」, 『行動計量学』, 第 47 卷, 第 1 号, 27–40 頁.
- 星野崇宏 (2009) 『調査観察データの統計科学: 因果推論・選択バイアス・データ融合』, 岩波書店.