

博士論文

時間の影響を考慮した
マーケティングの需要予測アプローチに関する研究

指導教員 松尾 豊 教授

東京大学大学院 工学系研究科
技術経営戦略学専攻

保住 純

概要

企業とは、主に営利を目的とした、資本を集めて事業を展開する経済主体であり、企業は自社の商品やサービスの価値、メッセージなどをより多くの顧客に届けるために、マーケティング活動を実施している。マーケティング活動には商品開発、価格設定、宣伝広告など様々なものが挙げられるが、その一つに市場の予測がある。近年多くの企業で行われている市場予測手法として、売上データや個人の行動履歴データを収集・分析して需要を予測し、それを次のビジネスアクションにフィードバックするという、最適化アプローチがある。

需要の予測には、商品や店舗の全体売上予測のような市場レベルの予測アプローチと、顧客個人が何を好んだり購入したりするかを予測する個人レベルの予測アプローチがある。市場レベルのアプローチによって正確な需要が予測できれば、計画的な生産を行えるようになり、期待される収益を残さず得ることができ、さらに過剰在庫を減らすなど、コストも削減することで増収につなげることができる。また、個人レベルの予測ができれば、潜在顧客をもれなく捉えてその商品を推薦することで購入に結びつけることによって、商品をより多く販売することができるようになり、最終的には全体的な売上を増加させることができる。そして、市場レベルの予測と個人レベルの予測は、それぞれサプライチェーンとデマンドチェーンの最適化に貢献するので、それらをどちらも行うことで全体最適がなされた商品販売につなげることができる。

ただし、これらのアプローチを実行する際に多くの場合に捨象される要素として、時間の影響がある。言い換えるならば、市場レベルと個人レベル、どちらの視点からも予測を行う際に、その予測対象の特徴が時間が経過しても不変であることを前提としているケースが多い。ユーザの趣味嗜好や意思決定モデルが、その予測を行うためにデータが収集された時点の状態のまま予測対象の時期においても維持されているか、その変化の影響がわずかであることから

予測モデルによってこれらについても学習することが暗に期待されている。

しかしながら、これらの時間の影響を意図的に反映させたモデルもこれまでに数多く提案されている。その明示的にモデルされるものとして、曜日や季節といった周期的な変化や、個人のライフスタイルの変遷などによる非周期的な変化がある。他にも、各消費財の消費周期など、個人ごとに異なるため一律に定義できないそれぞれ異なる時間的変化も存在する。さらに、これらの時間が予測対象に与えている影響は単一ではなく、複数の要因が重なって反映されており、データとしてはノイズも加えられた状況ですべて重なった状態でしか観測することができない。そのため、ここで挙げられるようなすべての時間の影響をあらかじめ把握した上で予測モデルを組み立てることは困難であると言える。

本研究では、市場レベル及び個人レベルの需要予測アプローチにおいて時間の要素を入力データから学習して取り入れるアーキテクチャを導入し、予測精度を向上させるアプローチを複数提案する。予測対象とするデータにおける時間の影響そのものを直接モデリングすることなく、その影響を学習できるモデルを用いることによって、既存の予測アプローチよりより高い精度を出すことができるということを示していく。

はじめに、市場レベルの需要予測のアプローチとして、複数の周期を自動的に取り入れられる売上に関する時系列の予測手法を提案する。これまでに商品の売上予測のために様々な時系列予測のアプローチが提案されているが、それに生じている周期を把握した上でモデリングを行い、その上でパラメータを調整しないと精度が出せないという問題を抱えている。この研究では、深層学習に畳込みネットワークを活用して、店舗や商材ごとの複数の周期を同時に学習し、予測結果に織り込むモデルを検討する。検証実験の結果、深層学習を用いた他の時系列予測手法と比べて、高い精度を記録することを示す。

次に、個人レベルの需要予測のアプローチとして、ユーザが複数の趣味嗜好を持ち、かつそれが時間とともに成熟するモデリングを行うことで推薦する手法を提案する。従来の商品推薦モデルはユーザの趣味嗜好が過去の行動履歴によって推定され、それが時間に対して不変であるということを前提としていた。しかし、商材によっては、その対象への成熟度が変化することで趣味嗜好が変わっていくことが仮定できる場合も考えられる。そこで、この研究ではユーザが対象商材に対し複数種類の趣味嗜好のカテゴリを持ち、それらが時間によって成熟してい

くということを仮定した上で、商品を推薦するモデルを検討する。対象として成熟度が変化するコンテンツに着目し、本提案手法によってより高い精度でより多くの商品を推薦の対象とできるようになることを示す。

最後に、先ほどの需要予測モデルの発展型として、ユーザの行動時刻を付与条件としてコンテンツを推薦する手法を提案する。近年、より少ないデータセットに対しより頑健性を持った推薦モデルとして、変分オートエンコーダを用いたものが注目されている。しかし、このモデルはユーザの行動順序のみに着目し、行動した時刻やその時間間隔を考慮に入れられていない。本研究では、このモデルに対してもユーザの行動時間を適切に条件として入力し、推薦精度を高めることができることを示す。時間情報を条件として付与する複数のアプローチを検証することによって、行動時間情報をより適切に取り扱うモデルを提案する。

以上の複数の研究を通じて時間の要素を取り入れることによる予測モデルの精度の向上を示し、これらの研究の共通点および相違点を整理する。それを踏まえて、市場レベル及び個人レベルの需要予測アプローチにおいて時間の要素を考慮したモデルを用いることが、そうでないアプローチに比べどのような利点があると考えられるか、精度以外の面からも考察する。そして、本研究による予測アプローチや、それ以外の市場や需要の予測アプローチ自体の限界を考察する。さらに、本研究から得られる知見が今後どのように発展してマーケティングプロセスの改善に貢献していくかを検討し、それに加えて、異なるドメインの予測タスクに対しても本研究がどのように貢献できるかを議論する。最後に、本研究の全体を総括する。

目次

第 1 章	序論	8
1.1	背景	8
1.2	本研究の目的	17
1.3	本論文の構成	17
第 2 章	関連研究	20
2.1	マーケティングにおける最適化の歴史	20
2.2	マーケティングの目的関数	21
2.3	時間を考慮した市場レベルの需要予測の研究	23
2.4	時間を考慮した個人レベルの需要予測の研究	26
第 3 章	複数周期に着目した畳み込みネットワークによる小売時系列データの予測	31
3.1	はじめに	32
3.2	提案手法	33
3.3	実験	37
3.4	まとめ	45
第 4 章	ユーザの行動履歴データを用いたコンテンツ興味推移カテゴリの可視化手法	46
4.1	はじめに	47
4.2	提案手法	50
4.3	評価実験	56
4.4	まとめ	62
第 5 章	行動時刻を考慮した条件付き変分オートエンコーダによる推薦システム	64
5.1	はじめに	65

5.2	前提知識	66
5.3	提案手法	69
5.4	実験	72
5.5	まとめ	77
第 6 章	考察	78
6.1	各論の考察	78
6.2	本研究全体の考察	80
6.3	本研究結果の制約と限界	82
6.4	本研究の他ドメインへの応用可能性	90
第 7 章	結論	93
	参考文献	97
	発表文献	102
付録 A	提案手法による可視化マップ	103

図目次

1.1	バリューチェーンの模式図	13
1.2	本論文の構成	18
3.1	CPNet の概念図	33
3.2	Cross CNN(CrCNN) の畳み込み例 (入力長 30、周期 7)	35
3.3	Multi-Window CNN の概念図	36
3.4	Rossmann データの一部概形	39
3.5	Walmart データの一部概形	39
3.6	Favorita データの一部概形	40
3.7	人工データの一部概形	40
3.8	線形スキップコネクションを用いたモデルにおける RMSE の変化 (Walmart、ラグ期間 8)	45
4.1	海外ビール市場の潜在的カテゴリ構造とその変化イメージ (Yang ら [Yang 14] より作成)	47
4.2	本研究の提案手法の模式図	48
4.3	各ユーザの興味レベル進展イメージ	52
4.4	興味行動時系列からのデータ収集の例	53
4.5	$K = 3$ 時に図 4.4 の興味推移シーケンスから作成されるグラフ例	55
4.6	提案手法による可視化結果の例 (マップ C)	61
5.1	Time-LSTM 3 の模式図	70
5.2	提案手法の模式図	71
5.3	MovieLens データセット内での Δt の分布	74
5.4	Last.fm データセット内での Δt の分布	74

表目次

3.1	テスト結果のまとめ（最も良い結果のものを太字で表記）	43
4.1	検証実験用データセットの概要	55
4.2	Yang らの手法による映画興味推移可視化の指標	58
4.3	提案手法による映画興味推移可視化の指標	59
4.4	Yang らの手法によるマンガ興味推移可視化の指標	60
4.5	提案手法によるマンガ興味推移可視化の指標	60
4.6	各可視化手法による映画・マンガの興味推移予測実験精度	61
5.1	データセットの概要	72
5.2	MovieLens 実験結果	75
5.3	Last.fm 実験結果	75
5.4	TSVAE 時刻情報の再構成の有無による比較実験 (Last.fm)	76
6.1	尺度水準とその性質	82
1	提案手法による MovieLens ユーザ興味推移可視化の結果 (1)	103
2	提案手法による MovieLens ユーザ興味推移可視化の結果 (2)	104
3	提案手法による MovieLens ユーザ興味推移可視化の結果 (3)	105

第 1 章

序論

本章では本論文の背景、目的、および以降の章構成について述べる。

1.1 背景

本節では、マーケティング、およびその中から本研究での対象とする部分と、その抱える課題について述べる。企業とは、主に営利を目的とした、資本を集めて事業を展開する経済主体である [松島 09]。近年では、持続可能な開発目標 (SDGs; Sustainable Development Goals) を経営の中に取り入れる企業が増えつつある。SDGs に掲げられたゴールやターゲットに企業として取り組むことは、単に社会貢献になるだけではなく、企業の持続可能を高めつつ新たなビジネスチャンスにもつながることが示唆されている。実際に、SDGs を経営に積極的に取り入れる企業は、そうでない企業に比べてパフォーマンスが高いことを指摘する研究もある*¹。また、SDGs は顧客やステークホルダーとのコミュニケーションのきっかけとしても利用されている。企業が SDGs への取り組みをアピールすることは、企業のブランドイメージを高め、市場からの評価を得ることにもつながるからである。そして、企業は自社の商品やサービスの価値、メッセージなどをより多くの顧客に届けるために、マーケティング活動を実施している。

マーケティング活動やその目的については、これまでに様々な定義がなされている。たとえば、Peter Ferdinand Drucker は「マーケティングの目的は、販売を不要にすることである」[Drucker 73]、Theodore Levitt は「マーケティングとは顧客の創造である」[Levitt 86]、Philip Kotler は「マーケティングとはニーズに応じて利益を上げること」[Kotler 16] と述べている。これらの定義はマーケティング活動全体を総括し、より端的に述べようとしているものである

*¹ SDGs 経営ガイド <https://www.meti.go.jp/press/2019/05/20190531003/20190531003-1.pdf>

が、Kotlerは同時に「マーケティングとは、個人や集団が商品や価値を創造し交換することで、必要とする物や欲しい物を入手するための社会的、経営的なプロセスである」という詳細な定義もしている。

本研究ではKotlerの后者の定義のように、マーケティングを一連のプロセスであるとみなし、マーケティングの定義にはアメリカ・マーケティング協会(AMA)の定義を採用する。AMAはマーケティングを以下のように定義している*2。

Definition of Marketing

Marketing is the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large. (Approved 2017)

(筆者和訳) マーケティングとは、顧客やクライアント、パートナー、社会全体にとって価値のあるものを創造し、伝達し、提供し、交換するための活動や制度、プロセスのことである。

以上に示されるように、マーケティングは幅広い活動を指した用語である。具体的には、マーケティング活動には商品開発、価格設定、宣伝広告など様々な活動が挙げられる。そして、それらはマーケティング・プロセスとして研究の対象となっている。また、これら個別の戦略を組み合わせる相乗効果を図ることはマーケティング・ミックスと呼ばれる。本章では、本研究が対象とするマーケティング・プロセスの中でもマーケティング・リサーチと、それを踏まえたプロセス最適化について、さらに、最適化についても時間の影響という観点から述べていく。

1.1.1 マーケティング・リサーチと最適化

マーケティング活動には上記の通り様々な種類のものがあるが、それらの活動効果を高めていくために、マーケティング・リサーチを行い、その分析内容を踏まえて種々の活動が行われる。マーケティング・リサーチもマーケティング・プロセスの一つとして捉えられ、プロセス全体の初期段階で実行されることが多い。アメリカ・マーケティング協会は、マーケティング・リサーチを以下のように定義している。

*2 <https://www.ama.org/the-definition-of-marketing-what-is-marketing/>

Definition of Marketing Research

Marketing research is the function that links the consumer, customer, and public to the marketer through information used to identify and define marketing opportunities and problems; generate, refine, and evaluate marketing actions; monitor marketing performance; and improve understanding of marketing as a process. Marketing research specifies the information required to address these issues, designs the method for collecting information, manages and implements the data collection process, analyzes the results, and communicates the findings and their implications. (Approved 2017)

(筆者和訳) マーケティング・リサーチとは、マーケティングの機会や問題点を特定・定義し、マーケティング施策を生成・改善・評価し、マーケティング・パフォーマンスを観察し、プロセスとしてのマーケティングの理解を深めるために使用される情報を通じて、消費者や顧客、一般の人々とメーカーとを結びつける機能である。マーケティング・リサーチは、これらの問題に対処するために必要な情報を特定し、情報収集のための方法を設計し、データの収集プロセスを管理し、実施し、結果を分析し、調査結果とその意味合いを伝達する。

マーケティング・リサーチの一つに、市場予測がある。そして、近年多くの企業で行われている市場予測の活用手法として、売上データや個人の行動履歴データを収集・分析して需要を予測し、それを次のビジネスアクションにフィードバックするという、最適化アプローチがある。これより、このアプローチの大まかな流れを述べる。

データの収集

まず、改善したいマーケティング・プロセスに関する情報を収集する。たとえば、商品の販売によって得られる利益を改善したい場合には、商品の売上やその在庫の状況を、顧客ごとの収益を改善したい場合には、購買履歴を収集することが考えられる。

これらのデータは商品販売の段階で収集しておくことが望ましく、その収集プロセスはシステム化されているケースが多い。たとえば、小売販売業の売上データは、POS(Point of sale)システムによって自動的に集計される。それに加えて、商品のカテゴリや色、サイズといったエ

ンティティの情報を商品納入の段階で記録しておけば、以降のプロセスでそれをもとにカテゴリライズし、より細やかな分析ができるようになる。他にも、顧客ごとに ID を発行し、その購買情報を顧客と紐付けて記録しておくことで、個人別で購買状況を収集することができるようになる。さらに、その ID ごとにもエンティティ情報を付与できるようにしておけば、より自分にあった商品が推薦されたり、クーポンが発行されたりするといった顧客が自らのエンティティ情報を提供することによる直接的なインセンティブを設けることによって、性別や年代、職業といった顧客情報（デモグラフィック情報）を収集することができ、商品と同様の、より細やかな分析につなげることができる。

他にも、マーケティング・プロセスに関する直接的な情報収集方法として、顧客の需要を直接把握するアプローチが挙げられる。その方法には、アンケートや電話によって消費者に直接需要を聴く質問法や、調査対象者となった顧客の行動を記録しておく観察法、仮想店舗や実験室内に顧客を招き、実際に購買行動でデータを収集する実験法、の三種類が挙げられる。ただし、これらは顧客に直接アプローチして行う手法であり、大量の顧客を対象としてデータを収集するには不向きである。

近年では、インターネットやスマートフォンが普及したため、その利用ログデータやインターネット上での公開データを用いてマーケティング・プロセスを改善するためのデータを収集することができるようになってきている。さらに、売上や顧客の属性を直接示しているデータでなくても、データマイニングによってこれらの相関を把握し、分析に活用する事例がある。たとえば、Abel らはブログの更新状況から本の売上ランキングを予測し [Abel 10]、Mestyán らは Wikipedia のページ編集状況とページ閲覧状況を用いて、映画の興行収入予測が可能であることを示している [Mestyán 13]。そして、それらのデータがより直接的に収集されたデータと比べてどれほど市場を予測できるか、その比較分析も行われている。たとえば、Goel らは Web 上の検索エンジンでの検索回数の特徴量を用いた市場予測は実データを用いた予測に精度は劣るものの、十分その代替として利用できることを示した [Goel 10]。Choi らは、検索エンジンでの検索状況は将来の予測には精度が上がりにくいことから使いにくい特徴量ではあるものの、実際の産業統計の公開よりも先に現在の市場動向を表示できる点で優れていることを示している [Choi 12]。

マーケティングプロセスの最適化

次に、上記の段階によって収集されたデータを用いて、市場動向や顧客動向の予測を行う。市場動向の予測や顧客行動の予測は、それぞれ時系列予測や商品推薦といった異なるタスクとして捉えられ研究されることが多いが、本研究では、市場動向の予測と顧客動向の予測を、マーケティング・プロセスにおける需要の予測タスクである同一の括りで捉え、需要の予測を、商品や店舗の全体売上予測のような市場レベルの予測アプローチと、顧客個人が何を好んだり購入したりするかを予測する個人レベルの予測アプローチに大別する。そして、市場全体の動向予測を市場（マクロ）レベルの予測、市場内の顧客個々人に関する予測を個人（ミクロ）レベルの予測と定義する。

市場レベルと個人レベルのどちらの予測も、より正確に予測ができるよう試みることになる。なぜなら、個人レベルでのアプローチによって正確な需要が予測できれば、計画的な生産を行えるようになり、期待される利益を十分に得ることができるからである。さらに、過剰在庫を減らすなど、顧客に販売されない商品を生産することによって発生するコストも削減することで増益につなげることができる。市場レベルの予測も、より正確にできるようにすることで、潜在顧客をもれなく捉えてその商品を推薦することで購入に結びつけることによって、商品をより多く販売することができるようになり、最終的には全体的な売上を増加させることができる。

そして、市場レベルの予測と個人レベルの予測は、それぞれマーケティングのバリューチェーンにおける上流であるサプライチェーンと下流であるデマンドチェーンの最適化に貢献する [Kastov 18]。在庫に過不足が発生しないように適切な量の商品を生産できるようサプライチェーンの最適化を行い、それらを想定する顧客に売り切れるように個人レベルの予測によるデマンドチェーンの最適化を行う。そして、実際の商品の販売状況を市場レベルの予測にフィードバックすることによって、その精度を高めていく。このサイクルを回すことによってバリューチェーン、ひいてはマーケティング・プロセス全体の最適化につながり、それらの相乗効果によってより効率的な商品販売へとつなげることができる。それは図 1.1 のように表される。

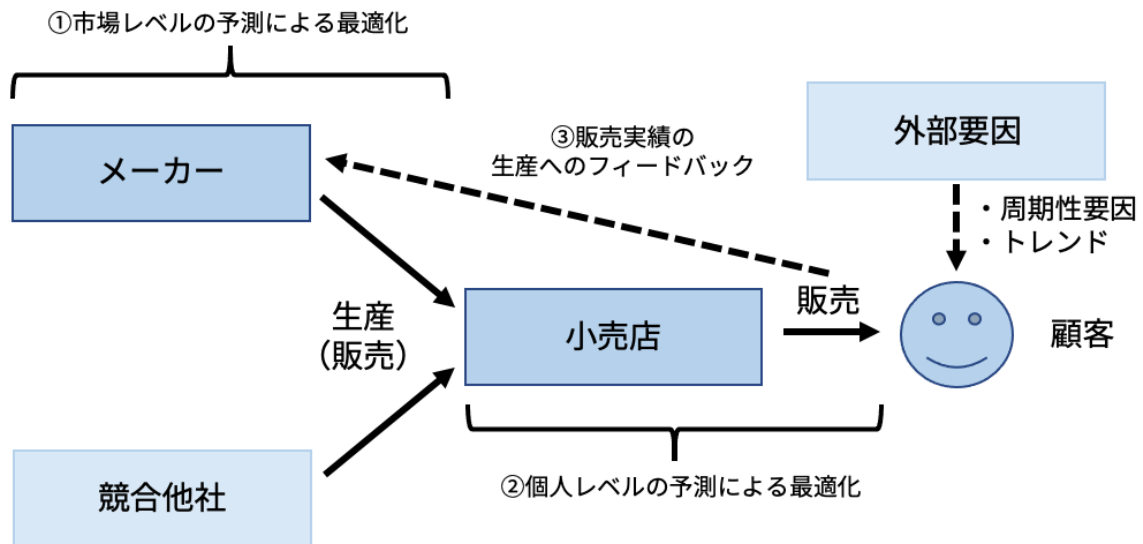


図 1.1 バリューチェーンの模式図

マーケティング施策の実行

ただし、需要の予測という最適化問題を解いただけでは売上の上昇には直結しないので、その内容を別のマーケティング・プロセス（施策）に反映させる必要がある。その際は、予測結果から数理的に最適だと判断される施策をとるのではなく、実際のビジネスの性質を捉えた意思決定と実行がなされる。たとえば、生鮮食品や生花では、市場レベルの予測によって需要や売上が予測できたとしても実際の売上とは乖離が生じるリスクが必ずあり、他の商材に比べて在庫を長期間保持することができず早期に破棄することになりやすいという理由から、その他の商材に比べて在庫が残るリスクを避けた納品計画を立てることになる。個人レベルの予測では、EC サイトなど顧客に向けて直接商品が推薦できる場合、予測結果をそのまま用いて商品の推薦やパーソナライズを行うことができる。店頭における小売販売のように、そのような直接的な推薦を行うことができない場合には、宣伝広告などを用いて顧客とコミュニケーションを取る施策などを検討することになる。その際には、そのペルソナが想像しやすいなど、この施策が実行しやすいクラスタを見出して定義した上で、ターゲティングを行う。

以上に挙げたように、市場レベルの予測は生産計画に影響を与えるものなので、実際に商品が納入されるまでのタイムラグを考慮して、それよりも早く予測をする必要がある。また、個人レベルの予測をターゲティングによって施策に反映させる場合には、ターゲットとするク

ラストがどのようなものであるか、施策の実行者が解釈可能である必要がある。このように、マーケティング・プロセスの最適化のために需要予測問題を解く際には、次の施策に結び付けられるような出力をする必要がある。ただし、近年では、無形物の販売を EC サイトやウェブサービスで完結させる場合などに、それらをすべてシステム化した上でマーケティング・プロセス全体を機能させる方法も用いられている。

このように、マーケティング・プロセスにおける最適化は以上の段階を追って行われる。

1.1.2 未来情報の観測不可能性

以上に示したようなマーケティング・プロセスにおける最適化の試みは、これまでに数多く行われてきた。しかし、コストや精度、採用できる難易度によって最適な手法は異なるため、そのための決定的な手法は確立されていない。ゆえに、どのようなプロセスを用いて最適化を行うかは、ビジネス上の意思決定の中で選択される。

需要予測には、商品のエンティティや顧客の属性情報、天気など外部要因など、様々な特徴量を予測時に考慮することができる。このような予測に用いる特徴量を引き出す概念の一つに、時間がある。時間、およびそれを規定する時刻の概念は、マーケティングの予測を行う上で非常に重要な手がかりとなる情報源であると考えられる。たとえば、小売店での購入は、多くの顧客が休みである土日や休日に増えるということが容易に想像される。このような売上増減の周期は様々な長さのものが考えられ、短いものでは通勤退勤時の売上増加といった一日の中で、長いものでは季節性の商品などの年単位での周期性を想定することができる。他にも、人気の上昇につれた全体的な売上増加のような、周期性のないトレンドを考慮することもできる。そして、実際の商品の売上には、それらの影響が重なり合って反映されていると考えられる。また、個人レベルの需要という観点では、顧客ごとの時間経過による購買傾向の変化を想定することができる。たとえば、各顧客が年月を重ねて成長するにつれてライフステージが変化し、購入する商品の特徴が変化していくことが考えられる。

ここまで述べた時間概念の影響は多くの人間や商品に共通する特徴であるが、人や商品によって変化の具合が異なるケースも考えられる。たとえば調味料やトイレットペーパーといった消費財は、顧客ごとに特有の周期で消費、購入されると考えられる。他にも、好きな嗜好品

やコンテンツに対するマニアックさ（熟練度）のように、年月が経つにつれてそれぞれ顧客によって異なるペースで趣味嗜好が変化することによって、購買傾向が変化するという非周期的な変化も考えられる [Abel 10]。

以上のように、時間の影響を取り入れることは予測に効果的であると考えられるため、それらを意図的に予測に反映させるモデルがこれまでに数多く提案されている。しかし、実際は、データ入手のコストや難易度といった観点から、よりシンプルな手法が選択されることが多い。実際に予測に用いることができる特徴量は限られるため、より正確な予測を行う際には考慮しなければならない様々な要素を捨象していくことになる。その際に、時間の概念もたびたび捨象の対象となる。

時間の概念を捨象する際には、それ自体を無視してしまうか、強い仮定を置いてしまうことが多い。具体的な一つの例として、市場レベルと個人レベル、どちらの視点からも予測を行う際に、その予測対象のもつ性質が、時間が経過しても変化しないことを前提としているケースが挙げられる。ユーザの趣味嗜好や意思決定モデルが、その予測を行うためにデータが収集された時点の状態のまま予測対象の時期においても維持されているか、その変化の影響がわずかであると仮定し、その影響の学習自体を前提としていない予測モデルによってこれらについても同時に学習することが暗に期待されている。

しかし、需要を予測する際に時間の影響を強く受けているため、捨象しない場合に比べて予測精度は低下してしまう。そのいくつか要因について、以下にまとめる。

需要発生要因の観測不可能性

データのみに基づく需要の予測では、データを収集した（サンプリングした）時点の状況しか観測できない。市場動向では過去の挙動、顧客行動については過去の行動に基づいて予測を行うが、完全にデータ依存の予測による意思決定では、過去の情報に基づいた決定しか行うことができなくなる。

たとえば、右肩上がりの売上といった、過去の状況を未来の状況に反映させるモデルを考えることも可能ではある。ただし、その際には現在の状況が未来にも継続するという外挿が行われることを前提とした予測モデルを採用しなければならない。実際は、自主的なセールの実施

やボーナスシーズンによる市場レベルでの売上の増加や、子どもの進学、結婚といったライフステージの変化による購買の増加など、ヒューリスティックに予測ができる売上の変化は存在するため、それを織り込んで予測に反映させることができる。また、そのような情報の織り込みを分析者によって行える手法も提案されている [Taylor 18]。しかし、それは多くの顧客に共通し、かつ観測や推定が可能な特徴を捉えているに過ぎず、実際に考慮しなければならない未来に与える時間の影響を分析者が把握しているとは限らない。

さらに、データの収集方法によっては、予測に必要な情報自体をデータから獲得することができない可能性がある。たとえば、サンプリングの周期によっては、予測に必要な情報全てを捉えきることができない可能性がある。たとえば、もし週に1度しかデータの収集が行えない場合、週単位の周期性があることを学習して予測をするということができないといった例が挙げられる。

観測の分解不可能性

売上や顧客の購買といったデータは、時間や時刻の影響すべてが重なり合った、結果としての状況でのみ観察できる。実際には、様々な要因が重なり合っており、人間の分解能にも限界があるため、影響の詳細までを分解しきって把握することができない。それを実現するために、フーリエ変換のような機械的な分解により強制的に分解した形式で表現する手法も考えられるが、施策の実行者にその結果の解釈ができるとは限らないため、意思決定や次の行動に結びつけることができない可能性があるといった問題を抱えている。

顧客の購買心理のモデリングにも用いられるような解釈可能な顧客モデリングのフレームワークは、多数のデータの観察によって人間の帰納により導かれたものである。たとえば、顧客が商品の購入に至るまでにどのような心理変容を経由するかを示した AIDMA モデルは、マーケティング・リサーチでの顧客観察を基にした、帰納的にモデリングされたフレームワークである。そのため、実際の顧客心理の遷移が正確に反映されているとは言い切れない。それはフレームワークの設計者が観測できたデータによるバイアスを多く受けているものであり、実際の需要予測には不適切である可能性も考えられる。事実、AIDMA には現在では反論も出され、AISAS、SIPS といった異なるフレームワークも提案されている。

1.2 本研究の目的

以上の時間の捨象による予測の問題点を踏まえ、本研究では、市場レベル及び個人レベルの需要予測アプローチにおいて時間の要素を入力データから学習して取り入れるアーキテクチャを導入し、予測精度を向上させるアプローチを複数提案する。予測対象データにおける時間の影響そのものを直接モデリングすることなく、その影響を学習できるようなモデルを用いることによって、既存の時間の影響を捨象する予測アプローチよりも高い精度を出すことができるということを示していく。

本論文では3つの研究を紹介する。これらの研究では、時間の経過による顧客や市場の状態変化に対し、明示的な仮定を置いたモデルを設計した上で学習を進める統計モデリングや機械学習を用いたアプローチと、そのような強い仮定を置かずにモデルの学習によって変化の仕方を学習することを試みる深層学習を用いたアプローチの、2種類の予測アプローチを使用している。第1の研究では、市場レベルの需要予測を取り扱う。売上の時系列を予測するに際し、強い仮定を置かない深層学習を用いたモデルを複数設計して実験を行うが、それに加えて想定される変化に対し強い仮定がなされた線形モデルを併用することによって予測精度をさらに向上させることができるということを示す。第2、第3の研究では、個人レベルの需要予測を行う。第2の研究では、顧客の対象ドメインに対する興味の種類およびその推移パターンに対し明示的な仮定を置いたモデルを設計し、その上で学習を進める機械学習を用いたアプローチによる予測を行う。一方、第3の研究では、このような顧客の興味の変化のパターンをあらかじめ明示的に仮定することなく、深層学習を用いたモデルによってその変化の仕方を把握して予測を行うというアプローチを用いる。

1.3 本論文の構成

本論文の構成は図 1.2 のようにまとめられる。

2章では、本研究と関連する研究について述べる。マーケティングにおける最適化の歴史から、市場レベルや個人レベルの各種需要予測手法の概略、本研究に特に関連性の深い手法についての説明を行う。

3章では、市場レベルの需要予測のアプローチとして、複数の周期を自動的に取り入れられ

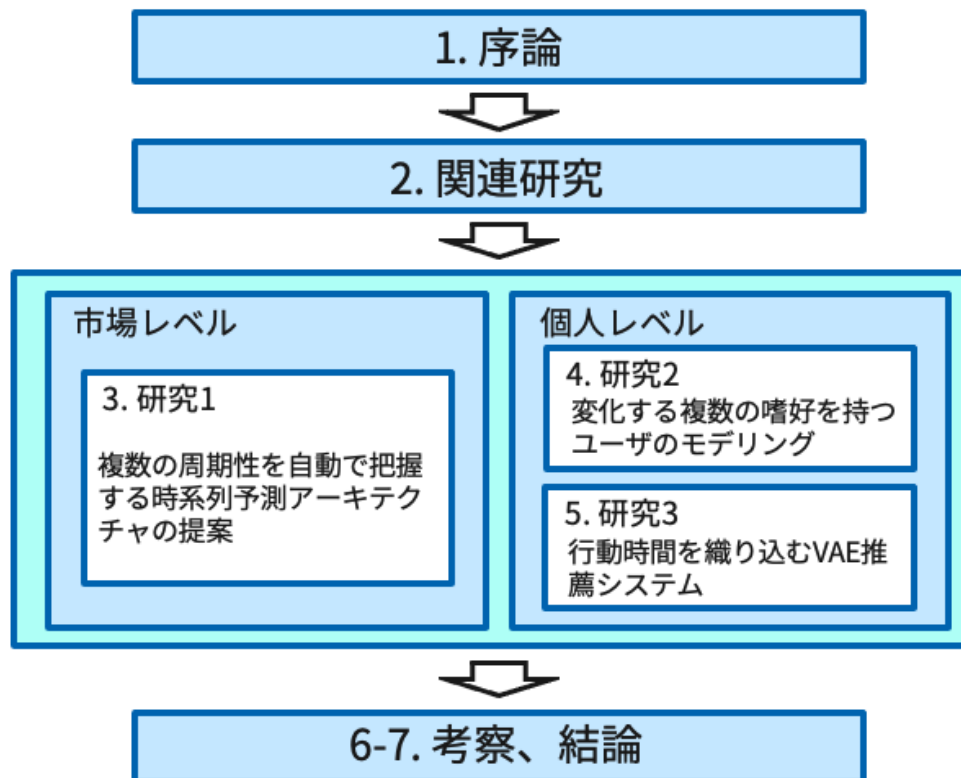


図 1.2 本論文の構成

る売上時系列予測手法を提案する。これまでに商品の売上予測のために様々な時系列予測のアプローチが提案されているが、それに生じている周期を把握した上でモデリングを行い、その上でパラメータを調整しないと精度が出せないという問題を抱えている。この研究では、深層学習に畳込みネットワークを活用して、店舗や商材ごとの複数の周期を同時に学習し、予測結果に織り込むアプローチを検討する。検証実験の結果、深層学習を用いた他の時系列予測手法と比べて、高い精度を記録することを示す。

4章では、個人レベル需要予測のアプローチとして、ユーザが複数の趣味嗜好を持ち、かつそれが時間とともに成熟するモデリングを行うことで推薦するモデルを提案する。従来の商品推薦モデルはユーザの趣味嗜好が過去の行動履歴によって推定され、それが時間に対して不変であるということを前提としていた。しかし、商材によっては、その対象への成熟度が変化することで趣味嗜好が変わっていくことが仮定できる場合も考えられる。そこで、この研究ではユーザが対象商材に対し複数種類の趣味嗜好のカテゴリを持ち、それらが時間によって成熟していくということを仮定した上で、商品を推薦するアプローチを検討する。対象として成熟度が変化するコンテンツに着目し、本提案手法によって、より高い精度でより多くの商品を推薦の対象とできるようになることを示す。

5章では、先ほどの需要予測モデルの発展型として、ユーザの行動時刻を付与条件としてコンテンツ推薦するモデルを提案する。近年、より少ないデータセットに対しより頑健性を持った推薦モデルとして、変分オートエンコーダを用いたものが注目されている。しかし、このモデルはユーザの行動順序のみに着目し、行動した時刻やその時間間隔を考慮に入れられていない。本研究では、このモデルに対してもユーザの行動時間を適切に条件として入力し、推薦精度を高めることができることを示す。時間情報を条件として付与する複数のアプローチを検証することによって、行動時間情報をより適切に取り扱う手法を提案する。

6章では、以上の複数の研究を通じて時間の要素を取り入れることによる予測モデルの精度の向上を示し、これらの研究の共通点および相違点を整理する。それを踏まえて、市場レベル及び個人レベルの需要予測アプローチにおいて時間の要素を考慮したモデルを用いることが、そうでないアプローチに比べどのような利点があると考えられるか、精度以外の面からも考察する。そして、本研究による予測アプローチや、それ以外の市場や需要の予測アプローチ自体の限界を考察する。さらに、本研究から得られる知見が今後どのように発展してマーケティングプロセスの改善に貢献していくかを検討し、それに加えて、異なるドメインの予測タスクに対しても本研究がどのように貢献できるかを議論する。

最後に、7章では、本研究の全体を総括する。

第2章

関連研究

本章では、マーケティングにおける最適化アプローチの歴史を紹介した上で、需要の予測に関する研究を、市場全体を需要を予測する手法と、顧客個人の需要を予測する手法と、に分けて紹介する。

2.1 マーケティングにおける最適化の歴史

もともと、マーケティングは学術研究の場ではなく、商業の現場から発展していった。Peter Ferdinand Drucker は [Drucker 73] の中で、マーケティングの元祖は日本の百貨店、三越伊勢丹の前身である越後屋にあると述べている。越後屋はすべての顧客に同じ値段で商品を販売する正札販売や、掛けではなく現金による取引など、当時としては革新的な施策を行っていた。他にも、屋敷売による顧客へのコンサルティングや、大福帳を用いた顧客管理といった、現代のマーケティング活動に通じるものも多く行われていた。ただ、この試みはあくまでも企業の経営努力の一つであったため、それが直接広がっていくことはなかった。

マーケティングという概念は、物々交換を行っていたり、生産者が商品を直接販売したりする時代には、考えられていなかった。マーケティング的な考え方が広まり始めたのは、第二次世界大戦後のアメリカであると言われている。イギリスより起こった産業革命により商品が工場で大量生産されるようになっており、特にアメリカで鉄道や通信網などが発展することで物流網が発達したことで、これまでより広いアメリカ全土を市場として捉えられるようになった。これにより、アメリカは大量生産・大量消費の時代に突入することになり、その需要の喚起や販売の最適化が課題として考えられるようになった。

各企業の努力で行われていたマーケティング活動は、マーケティング・サイエンスとして、学術研究の対象にもなっていた。そのさきがけは、Arch Wilkinson Shaw[Shaw 12] であると言われている。それは、Frederick Winslow Taylor が提唱した科学的管理法を、ビジネス、マーケティング課題に応用した試みであった。その後、P&G の Butler Ralph Starr がこのような試みを”Marketing”と命名し、ウィスコンシン大学で教鞭をとり、その後、このような講義がアメリカの多くの大学で開講されるようになった。

マーケティングは主に経営学の分野で研究対象となっているが、経済学、経営学、社会学、統計学、計算機科学など、文理を問わず分野を跨いで研究対象となっている。たとえば、理系の分野ではオペレーションズ・リサーチの対象として取り扱われた。オペレーションズ・リサーチとは、主に第二次世界大戦より発達した、数学や統計アルゴリズムを活用することで、さまざまな問題に際して最適化を行う方法である。顧客や売上の最大化や効率的な出店計画など、マーケティング・プロセスが抱える課題やそれに結びついた小問題が最適化問題として定義された上で、それをいかに解くかということが考えられるようになった。

ここまで挙げた通り、マーケティング活動は多岐に及び、それぞれの研究も無数に存在するため、本研究では、その一つである需要の予測に関する課題のみに焦点を当てていく。

2.2 マーケティングの目的関数

前項で述べたように、マーケティングの目的は売上の最大化である。そして、そのために販売者が行えることは、マーケティング・プロセスにおける意思決定や施策である。以上の内容も踏まえて、関連研究や本研究の目的を整理するために、これらの概念を数式を用いて表現していく。

y を顧客一人あたりの購入量や店舗ごとの総売上といった売上と関連する指標、 x は y を決定する要因となる各特微量とする。そして、 R を市場動向など最終的に収益を決定する環境モデルとする。このとき、マーケティングが一般社会における事象であることから、収益を左右する要因から収益が決定論的には一意に定まらないことを前提とすると、得られる売上は、条

件付き確率の表現 p を用いて以下のように書ける。

$$Income = R(p(y|x)) \quad (2.1)$$

そして、マーケティングにおいて、販売者は売上最大化のために実行可能なあらゆる施策から構成されるマーケティング施策空間 S から適切な施策 s を選択するものとする。販売者は販売に対し、施策 s を選択・実行することによって x に働きかけることになる。このことから、 x は s の関数であると表現できる。

$$x = x(s) \quad (2.2)$$

以上の内容を踏まえると、マーケティング・サイエンスの目的は、以下の最適化問題を解き、最適な s を探索することにあると言える。

$$s_{opt} = \operatorname{argmax}_{s \in S} R(p(y|x(s))) \quad (2.3)$$

2.2.1 説明変数 x のモデリング

x をボトムアップに設計することは、マーケティング・サイエンスのモデリング・アプローチであると言える。ボトムアップなモデリングでは最適な s をより明確に発見できるため、実務利用を念頭に置いた際には効果的であるが、実際の x はかなり複雑であり、正確にモデリングを行うことは困難であるため、機械学習によって相関関係をより忠実に表すモデルを設計する。前項より $y = p(x|s)$ と表現できたので、 $\hat{y} = x(s)$ であると近似した表現ができる。そして、 s と y の関係性を、より正確なモデルを仮定し実験することで検証するのではなく、大量のデータを用いてその近似器を学習する。これがデータ・マイニングの考え方である。モデリングの目的は主に理解と予測に大別できるが、機械学習は後者に重点を置いたものであり、本研究も後者のみを対象とする。

また、 $\hat{y} = x(s)$ は、一般に、 $x(s)$ の情報量が多いほど、その表現力が高まると言える。 $x(s)$ を決める要素には天気などコントロールできない変数があり、その影響を受ける。また、他の販売者の施策による影響によって干渉されることも考えられる。そのような要素をまとめて θ と置くと、以下のように書ける。本研究では、 θ の中でも、特に時間に関する情報 t に着目し

た上で \hat{y} を表現することを試みる。ただし、最終的に R を定め売上を決定する要素は、販売者とは異なる、顧客の購買の意思決定であることに留意する必要がある。

$$x(s) = x(s|\theta) \quad (2.4)$$

2.3 時間を考慮した市場レベルの需要予測の研究

マーケティングの需要予測における市場レベルの予測として、直接的な市場全体の需要予測が挙げられる。そして、需要は時間の経過に伴い変化するため、時系列予測モデルを用いた予測が用いられることが多い。このような予測アプローチの一つに将来の売上を予測し、それを将来の需要とみなす方法がある。正確な需要が予測できれば、計画的な生産を行えるようになり、期待される収益を残さず得ることができ、さらに過剰在庫を減らすなど、コストも削減することで増収につなげることができる。ただし、モデルを学習させる期間中に売り切れが発生していた場合、その売上は真の需要を表現していないということになる。十分な在庫を用意できていれば、売上が需要を正確に反映できているため、より正確な需要予測につながることで好循環を生むことができる。

2.3.1 ボトムアップのアプローチ

将来の市場売上全体を予測するために、売上全体を構成する数々の要素を組み合わせることによって予測を行う手法が数多く提案されている。たとえば、従来の小売データ予測に関する研究では、時系列の周期性の知識を取り入れる試みとして、予測対象が **Decomposable Time Series Model**[Harvey 90] で表すことができると仮定されたものがある。そこでは、予測できる時系列を以下の式で書き表している。

$$y(t) = g(t) + s(t) + h(t) + \epsilon \quad (2.5)$$

この式における $g(t)$ はトレンド、 $s(t)$ は季節性、 $h(t)$ は休日、 ϵ はノイズである。これは、予測対象とする時系列の周期性を $s(t)$ や $h(t)$ としてモデリングして捉える手法であるが、実際のデータに含まれる周期性はより複雑であると考えられるので、できるだけ柔軟に複数の周期性を捉えられるようにしたいということが、本研究の動機の一つである。

他にも、市場を構成する顧客に関する情報を収集し、それを積み上げることで予測を行う試

みも行われている。顧客の属性ごとに、その顧客の済む地域の発展具合や政策を表す地理変数、性別や年齢、所得などを表すデモグラフィック変数、顧客のこれまでのライフスタイルや所属集団の価値観を表す心理的変数、商品に対する理解や使用頻度を表す行動変数といった変数を導入して、それらの変数を組み合わせることによって予測が行われる。

2.3.2 トップダウンのアプローチ

本項では、時系列を入力データとし、その間接的な表現を組み合わせるのではなく、直接それ自体を学習して予測しようとする手法を紹介する。

古典的な時系列予測モデル

トップダウンなアプローチとして代表的なものとして、時系列モデルを導入し、それにフィッティングさせるといったものがある。これまでに研究されてきた古典的な時系列モデルの代表的な例として、自己回帰モデルが挙げられる。参照期間を N とした $AR(N)$ モデルは、以下の式で示される。

$$x(t) = \sum_{k=t-(N-1)}^{t-1} a_k x_k + c + \epsilon(t) \quad (2.6)$$

ここでの a は係数、 c は定数項、 $\epsilon(t)$ は正規分布に従うノイズである。これはスカラー時系列の予測モデルであるが、ベクトル時系列 $\mathbf{x}(t)$ を予測する VAR モデルも存在する。また、これらのモデルは予測する時系列にマルコフ性を仮定している。これは言い換えれば、現在の値によって将来の値が定まることを意味する。ただし、実際は厳密にはマルコフ性が成立しない状況においてもマルコフ性が成立するように特徴量を設計することがある。たとえば、金融分野で用いられる移動平均線がその一例である。このようなモデルを実際に小売業に関連した時系列予測タスクに活用した研究として、[Arunaj 15] が挙げられる。

機械学習による時系列予測モデル

有名な機械学習を使用した予測モデルとして、XGBoost[Chen 16] がある。これは勾配ブースティングと決定木とを組み合わせた、時系列にとどまらない予測タスクに用いられる手法である。また、決定木の刈り取りを内部で自動で行うことによって、学習の高速化を実現している。その後、この研究を踏まえ、より大きなデータセットに対応して高速な学習を実現する LightGBM[Ke 17] という派生手法も提案された。この手法は決定木の枝分かれの基準の探索を適切に省くことによって計算量を抑えることで、より大規模なデータに対して適用しやすく

なる工夫がなされている。これらの手法はデータ予測コンペティションにおいてたびたび利用されているものであり、実際にコンペティション・サイト Kaggle^{*1}において好成績を出したモデルにも多く採用されている。ただし、これらのモデルを用いて高い精度を出すためには、入力する特徴量の入念な設計やモデルパラメータのチューニングが必要になり、その値に応じて予測精度が大幅に変化する。そのため、高い精度を出すモデルにするには十分な調整が必要である。このことから、これらの手法は、機械学習を用いた予測タスクに不慣れである人が短時間に高精度な予測を行うための手法としては適切であると言い難い。

また、主に画像認識や翻訳タスクに用いられ、既存の機械学習による手法と比べて高い精度をあげてきた深層学習モデルは、近年では時系列の予測においても用いられるようになった。深層学習を用いた単純な例として、再帰的ニューラルネットワーク (RNN; Recurrent Neural Network) や学習実行時に過去の情報を失わないよう新たにゲートを設定した Long-short Term Memory(LSTM) や Gated Recurrent Unit(GRU) の使用が挙げられる。ただし、RNN による予測は逐次的に入力を行い更新を行うため、入力に長期の時系列を入力した際には、機械学習による手法に比べて非常に多くの計算時間やメモリが必要になる。この問題を踏まえ、入力を時間方向の 1 次元ベクトル (テンソル) と捉え、それを学習して将来を予測する畳込みニューラルネットワーク (CNN; Convolutional Neural Network) による時系列予測も行われるようになりつつある。深層学習による時系列予測の例には、音声認識分野では Wavenet[Oord 17] があり、それを時系列データ予測に応用した研究として [Borovykh 17] が挙げられる。

深層学習による実世界データの時系列予測の研究として、LSTNet[Lai 18] と SOCNN[Bińkowski 17] が挙げられる。LSTNet は深層ネットワークによるモデルと AR モデルの学習を並行して実行し、最終的に両方を予測に用いる手法である。深層ネットワークのみを予測に用いると、周期性のある情報が失われる可能性があるが、AR モデル側に適切な参照期間を設計すればそれを保持することができ、最終的な予測性能が高まるということを示している。一方、SOCNN は時系列方向とベクトル内での畳込みという二方向の畳込みを行うベクトル時系列予測モデルである。RNN ではなく CNN を用いる理由としては、前述した計算時間の削減のほか、畳み込みを重ねることでノイズが除去できる効果への期待がある。他にも、入力が一部観測できずに入力が離散的になってしまう場合にも、そのサンプリン

^{*1} <https://www.kaggle.com/>

グ間隔を同時に入力することでうまく予測を行える工夫もなされている。その結果、時系列の形成に様々な要素が考えられる金融時系列データの予測においても、他の時系列予測モデルに比べて高い精度を上げている。本論文にて提案する手法は主に [Lai 18] や [Bińkowski 17] から着想を得ている。これらの研究は深層モデルと線形 AR モデルの併用を指向している。他にも、深層学習を時系列予測の補助的に用いる手法に、R2N2[Goel 17] がある。R2N2 は主たる予測を AR モデルを用いて行い、その予測値と実際の値との残渣を RNN で予測し、両方を用いて AR モデルによる予測結果を修正するモデルである。

小売業における時系列データに深層学習による予測を実施した研究として、日本のスーパーマーケットにおける売上予測を試みた [Kaneko 16] が挙げられる。しかし、これは予測対象を、その直前に比べてどの程度売り上げが増減するかを予測する分類タスクに置き換えて、それを解くものである。本研究は売上そのものを直接予測し、出力がスカラ一列である予測タスクを取り扱う。

2.4 時間を考慮した個人レベルの需要予測の研究

市場はそれぞれ異なる属性や性質を持つ人間の集合体であると言える。かつて商品の販売は 1 対 1 の対面販売であったため、マーケティング・プロセスとしても、その対象は常に個人単位で考えられていた。その後、テレビや新聞といったマスメディアを用いて幅広く販促宣伝活動を行うマス・マーケティングが主流になったが、顧客への訪問販売やセールスレターを郵送するといった、ダイレクト・マーケティングも引き続き行われていた。近年では、EC サイトやスマートフォンアプリを通じたデジタル・マーケティングが可能となり、顧客一人一人に合わせたマーケティング施策が可能になったことから、個別の消費者の特徴を捉えそれに個別に合わせた施策を取ることができるようになった。このような背景から、顧客一人一人を対象とする個人レベルの需要予測は、その有効性や応用範囲が増していると言える。

ユーザに対し直接調査などを行わずに個人の購買行動動向を把握する方法に、個人の行動ログデータを用いて把握する手法が研究があげられる。個人の消費動向が詳しく把握できるようになると、よりその個人に適した推薦が実行できるようになると考えられるため、ショッピングサイトの収益向上を目的とした推薦エンジンを改善する際のアプローチとして消費動向把握が研究されてきた。そして、その予測やそれに対するフィードバックに必要なものが、顧客

ごとの行動履歴データである。事前に顧客ごとにデータを収集して、それを活用して施策を決定して実行する。その後、それに対するフィードバックデータを収集し、その結果をさらに施策に反映させることを繰り返すことで、分析と施策をブラッシュアップしていく。このようなアプローチはデータベース・マーケティングとも呼ばれる。

このようなマーケティング・プロセスには他にも、顧客の商品やブランド、店舗への忠誠心を表すロイヤリティの最大化といったものもなされている。顧客がその商品やブランドに費やすと考えられる金額総額は生涯顧客価値 (Life Time Value;LTV) と呼ばれ、業種によっては目の前の短期的な売上の向上だけではなく、生涯顧客価値を引き上げるマーケティング施策を取り続けることによって長期的な視野で売上を上げるアプローチも用いられている。

ユーザモデリングによるアプローチ

先述した AIDMA や AISAS のような顧客ごとにモデルを仮定し、行動履歴データをそれに従わせて学習する推薦手法がある。そのような一つの手法に、ユーザに成熟度の概念を取り入れる手法があり、これまでも成熟度の高いユーザを発見する研究がなされてきた。たとえば、Bian らは、質問投稿サイトで良質な解答を投稿するユーザを少ないデータ量から特定する手法を提案している [Bian 09]。また、時間の経過やその Web サービスの利用によってユーザの成熟度が進展するとみなし、ユーザ個人の成熟を連続的に捉え続ける研究も存在する。たとえば、Lauschke らは、Twitter ユーザの利用期間によるつぶやきの変化を [Lauschke 12]、Antin らは、Wikipedia ページ編集者の編集行動の変化を追跡している [Antin 12]。そして、Mcauley らは、消費者が高評価をつけるアルコール類が時間経過で変化することを示した上で各ユーザにレベルという属性を付与し、それをを用いることでアルコールへの評価予測の精度を高めた [McAuley 13]。さらに、Yang らはこの手法を病状の特定などにも応用できるように修整し、Mcauley らの手法をより幅広い時系列データに用いることのできる手法へと改良した [Yang 14]。

付加情報を使用した推薦システムの研究

推薦手法の精度を改善するアプローチの一つに、各ユーザに属性を付与する方法がある。実際の例としては、Sarwar らはユーザのクラスタリングによって推薦の精度を高めてい

る [Sarwar 02]。逆に、Oconnor らによってユーザに向けて推薦する商品をクラスタリングをするという手法も編み出されている [O' Connor 99]。ここで用いられるクラスタリングはハードクラスタリングと呼ばれ、一つの要素を一つのクラスタに属させる手法である。しかし、実際にはユーザをその性質によって単一のグループのみに属させることは不自然であり、性質によってはそのユーザを複数のグループに属することがより適切である場合も考えられる。そのため、各ユーザをそれぞれ一つのクラスタに配属するのではなく、ユーザにその性質に合ったタグを付与する方法が Sen らによって研究された [Sen 09]。しかし、タグ付けの基準によっては、どのタグも付与されないユーザが発生する可能性があり、そのユーザへの推薦がうまくできなくなる問題も生じた。そのため Bin らの研究では、ユーザとアイテムの双方を最低 1 つ以上のグループに属させるクラスタリング手法である Multiclass Co-Clustering 手法を提案し、それによってより広いユーザに対する推薦制度を高めている [Bin 12]。

多くの場合、ユーザに推薦をする際にその推定のための情報源として用いるデータは、ユーザの商品の購買履歴や、高評価レビューのログといったユーザごとの行動履歴データである。その他にも推薦精度の向上に寄与すると考えられる情報を推薦時に活用することができれば、推薦システムの精度が向上できることが期待される。このことから、これまでに様々な付加情報が検討され、その有効性が検証されてきた。たとえば、これまでに商品の画像情報を加える [Kang 17]、レビュー文章の情報を加える [Zheng 17]、SNS 上のユーザのつながりの情報を加える [Zhao 17]、直前のアクションからの時間間隔を加える [Zhu 17] といったものが挙げられる。特に、近年では深層学習を用いることによって様々な情報を似た形式の内部表現として取り扱うことができるようになったため、これらをまとめて推薦に活用するマルチモーダル推薦 [Liu 19] も行われている。商品やユーザのログデータに関連する様々な形式のデータが、推薦の精度向上に利用できる可能性があると言える。

また、これらの情報を単純にログデータと同時に入力するのではなく、その入力方法を工夫することによって推薦精度を向上させる研究も存在する。たとえば、[Cheng 16] は追加する情報を並列に学習し、最後の層の直前にユーザのログデータと組み合わせて推薦に使用している。他にも、[Beutel 18] では、ログデータに対しモデルを通す前後で文脈情報を付与する手法がとられている。[Beutel 18] ではこれらの情報を入力する箇所を工夫することによって何層ものネットワークの追加に相当する精度の向上が見込まれることが報告されているため、これらの情報付与に関する研究は、推薦システムの研究においても重要性があると考えられる。

2.4.1 相関関係の分析を基としたアプローチ

似た顧客をグルーピングし、その一人の顧客に有効だった施策を同じグループの異なる顧客にも実行する手法がある。それは機械学習の文脈ではクラスタリング、マーケティング用語としてはセグメンテーションと呼ばれる。顧客の行動を直接予測するのではなく、似た行動をした顧客を参照し、それを推定する手法がある。商品や顧客の潜在空間を学習し、そのマッチングであると考えられる。深層学習では、特徴量の学習であると捉えられるため、様々な手法が研究されてきた。

伝統的な推薦システムの手法には、アイテムベースや協調フィルタリング、Matrix Factorizationなどが挙げられるが、これらの一部に深層学習を活用することで、より推薦精度を高めた手法が提案されている。たとえば、[He 17]は多層パーセプトロンを用いた協調フィルタリング、[Xue 17]はMatrix Factorizationの一部に深層学習を活用している。

それらのアプローチの中に、VAEをベースとした手法がある。[Liang 18]は協調フィルタリングを拡張する推薦手法である。VAEを用いた推薦システムは、ユーザの行動履歴データを入力し、それとVAEのエンコーダからの出力との差分を見ることで商品を推薦する。VAEベースの推薦システムは、従来の深層学習を用いた推薦システムに比べて、より小さいデータセットに対する精度が向上し、頑健性も高まることが報告されている[Liang 18]。

他にも、単純なVAEを発展させることでさらに精度を向上させる手法が提案されている。たとえば、[Sohn 15]は、エンコーダとデコーダの入力に条件を付与することで、より再構成誤差を小さくすることができる。このように工夫がなされたVAEを用いた推薦システムの研究も進められている。たとえば、ユーザの行動の順序を考慮して予測に活用できるようにするために、系列情報を扱えるよう拡張された手法であるSVAEが提案されている[Sachdeva 19]。[Sachdeva 19]では行動の順序のみが考慮されており、各行動の日時は考慮されない。推薦時に用いる情報に行動の時間の情報を取り入れることができれば、商品の購入間隔や商品カテゴリに対するユーザの成熟度といった各行動間の時間間隔による情報が取り入れられるようになるため、さらなる推薦精度の向上が期待される。本研究の5章では、エンコーダに時間情報を考慮するTime-LSTM[Zhu 17]を導入し、精度の向上を検討する。さらに、条件付きVAEを

使用し、その条件に行動時刻から得られる情報を取り入れることによって、さらに精度を高める手法を提案する。

第3章

複数周期に着目した畳み込みネットワークによる小売時系列データの予測

小売業において、将来の商品の需要を正確に予測することは売上を増加させる上で重要であるため、これまでに計量経済学や統計学といった様々な分野で数々の時系列予測手法が研究されている。近年では、さらに予測性能を上げるために、深層学習を活用する試みもある。ただし、これまでの深層学習を用いた手法は小売データに特化して設計された手法ではないため、小売データにおける有用性は不明瞭である。ゆえに、その特徴を捉える機構をモデルに組み込むことができれば、これまでの手法を超える予測精度を達成できる蓋然性は高い。このことを踏まえ、本研究では小売データが持つ周期性に関するドメイン知識を取り入れた深層学習による予測モデルを提案する。この周期性に関するドメイン知識とは、曜日や月（季節）のような複数の商品に共通する売上の周期変動の他に、商品ごとに特有の異なる消費の周期があり、それぞれ商品の売上データごとにそれらの異なる複数の周期が反映されていると考えられることである。従来手法では売上の周期をあらかじめ把握した上で予測モデルを設計する必要があるが、本研究では、既存手法を時系列ごとに異なる複数の周期を捉えられるように拡張することで周期性に関するドメイン知識を考慮して予測を行う手法を提案する。そして、複数種類の実際の小売データや人工的に作成したデータを用いた予測実験を行い、その精度の向上をもって提案手法の有効性を確認する。本手法を活用することによって、購買の周期についての事前情報がない商品においても需要の予測ができるようになり、より収益を高められる発注や販売が可能になると考えられる。

3.1 はじめに

小売業において、将来の商品の需要を正確に予測することは売上を増加させる上で重要であり、企業が下す意思決定においても、需要予測は重要な位置を占めている [Tsoumakas 18]。企業は、将来の各商品需要を予測した上で商品の発注を行うが、その際に、需要を高く見積もりすぎると過剰在庫が発生し、廃棄などによる経済的損害が生じる要因となる。一方で、需要を低く見積もりすぎると、売り切れが発生してしまい、商品販売の機会損失が発生してしまう。そのため、売上をより正確に予測できるようにすることが、最適な発注や、ひいては収益の最大化につながる。

近年、このような需要予測において、属人性を排除することや、より正確な予測を行うことを目的として、計量経済学や統計学といった様々な分野で数々の時系列予測手法が研究されている。たとえば、主に計量経済学分野で使用される自己回帰 (AR; Autoregressive) モデルがある。他にも、機械学習による予測モデルとして XGBoost[Chen 16] や LightGBM[Ke 17] といった手法も提案され、ライブラリとして公開されている。近年では、さらに予測性を上げるために、深層学習を活用しようとする試みもある [Borovykh 17]。たとえば、Long- and Short- term Time-series Network (LSTNet)[Lai 18] は線形モデルに深層学習による予測モデルを組み合わせている。また、Significance-Offset Convolutional Neural Network (SOCNN)[Bińkowski 17] は線形モデルのパラメータや入力の補正項を深層モデルによって学習させて予測を行っている。

ただし、これらの深層学習を用いた手法は小売データに特化して設計された手法ではなく、異なるドメインの系列データに向けて提案されたものであり、小売データにおける有用性は不明瞭である。たとえば、LSTNet[Lai 18] は交通量や電力といったドメインでの検証を行っている。一方で、SOCNN[Bińkowski 17] は金融というドメインに特化してモデルを構築し、検証を行っている。[Bińkowski 17] のように、小売データの特性を踏まえ、それを捉える機構をモデルに組み込むことができれば、これまでの手法を超えるような予測精度を達成できる蓋然性は高い。

以上の内容を踏まえ、本研究では図 3.1 で示されるような小売データが持つ周期性に関するドメイン知識を取り入れた深層学習による予測モデル (Consume Prediction Net; CPNet) を提案する。この周期性に関するドメイン知識とは、それぞれ商品の売上データごとに異なる複数の周期が含まれているということである。たとえば、消費財の売上データの時系列には、曜日

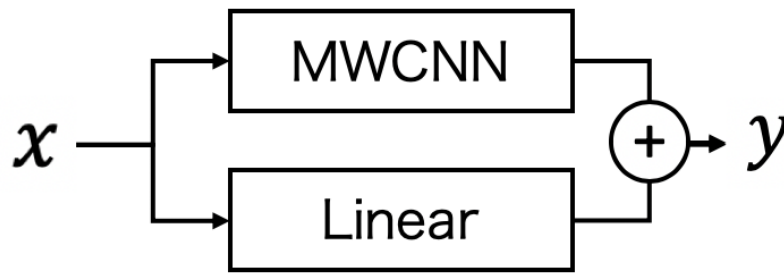


図 3.1 CPNet の概念図

や月（季節）のような複数の商品に共通する売上の周期変動の他に、商品ごとに特有の異なる消費の周期があり、それら複数の周期が反映されていると考えられる。予測モデルにこのような知識を取り入れることができれば、予測精度を向上させる可能性は高い。

本研究では、既存手法である LSTNet を基に、時系列ごとに異なる複数の周期を捉えられるように拡張することで周期性に関するドメイン知識を考慮して予測を行う手法を提案する。そして、複数種類の実際の小売データや人工的に作成したデータを用いた予測実験を行い、その精度の向上をもって提案手法の有効性を確認する。

3.2 提案手法

3.2.1 問題設定

本論文で提案する手法は、ダウンロードコンテンツのような在庫の概念を想定しない商材ではなく、実際に在庫を抱えるような小売業分野での使用を念頭に置いた手法である。また、学習時にはなかった新商品や新店舗の売上予測を行いたい際にも学習したモデルを適用することができるようにするため、本研究では固定された次元数の多次元時系列 $\mathbf{x}(t)$ ではなく、1次元時系列 $x(t)$ を入力とし、その予測を行う ($t \in \mathbf{N}$)。一般に、複数次元のベクトル時系列を扱う予測モデルのほうが入力される情報量が多くなり、特に深層学習を使用したモデルではその次元間の相互関係を学習することで予測精度が高まることが考えられる [Bińkowski 17]。一方で、このような手法は入力する時系列の次元を固定する必要があるためコールドスタート問題に対応できなくなり、たとえば、学習時にはなかった新商品や新店舗の売上予測を行いたい際に、学習したモデルを適用することができなくなる。これは小売業における実用性を失ってしまうことにつながるため、今回は1次元時系列の予測モデルを取り扱う。同様の理由で、曜日や天気、プロモーションの有無といった小売業時系列における追加の条件は入力しない。また、予

測の粒度は日や週単位とし、予測対象とする時系列には一ヶ月から最長で一年程度の周期が含まれていることを想定している。

そして、本手法は予測までのラグのある時系列予測タスクを行う、すなわち $x_{k+L}(L > 0)$ を $x_{[1:k-1]}$ より予測する。これは、予測結果を参考にした発注を行い、その商品が納品されるまでに期間を要することを前提とするからである。また、直後 x_{k+1} の予測を繰り返すことで逐次的に将来を予想するものではなく、ラグ L 後の値を直接予測する。そのほうが逐次的に予測する手法と比べて誤差が累積されず、最終的な予測精度が高くなるからである [Chevillon 07]。商品ごとのリードタイムや企業の実態に応じて適切なラグは異なるため、本研究では具体的な L の値についての詳細な検討は行わないが、その代わりに、複数種類のラグの場合による実験を行い、異なるラグに対する予測モデルの汎用性を確認する。

3.2.2 周期性に関するドメイン知識

今回対象とする時系列の特徴の一つに、複数の周期が存在しているという点が挙げられる。人間の行動は一週間や一ヶ月、一年といった複数の異なる周期の影響を強く受けており、小売データなど人間の行動が反映されたデータにも、その周期が反映されている。その長さは時系列データのサンプリング周期によって異なるが、たとえば日ごとにサンプリングされた日時データの場合、週 (7 日) や月 (約 30 日) といった周期が反映されていると考えられる。また、このような周期は、季節のような多くの商品に共通しているものもあるが、商品ごとに特定の周期で消費されて再購入されるものもあり、その消費周期は商材によって異なる。人の手によって将来の売上予測を行うときは、これら複数の時間幅で時系列を捉え、その周期性などの傾向を捉えた上で予測が行われる [Kastov 18]。それぞれのデータに反映された周期をすべて捉えて明示的にモデリングするには実務上限界があるため、その周期性を捉えることができるモデルを設計し、それを用いて予測させるということが有効であると考えられる [Taylor 18]。

従来の小売データ予測に関する研究では、時系列の周期性の知識を取り入れる試みとして、予測対象が Decomposable Time Series Model [Harvey 90] で表すことができると仮定されていた。そこでは、予測できる時系列を以下の形式で書き表している。

$$y(t) = g(t) + s(t) + h(t) + \epsilon \quad (3.1)$$

この式における $g(t)$ はトレンド、 $s(t)$ は季節性、 $h(t)$ は休日、 ϵ はノイズである。これは、予測対象とする時系列の周期性を $s(t)$ や $h(t)$ としてモデリングして捉える手法であるが、実際のデータに含まれる周期性はより複雑であると考えられるので、できるだけ柔軟に複数の周期性

---	---	---	---	---	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

図 3.2 Cross CNN(CrCNN) の畳み込み例（入力長 30、周期 7）

を捉えられるようにしたいということが、本手法を提案する動機の一つである。

次に、上記のような複数周期性に関する仮定を深層モデルで捉える方法として、カレンダー状の畳み込み、および複数ウィンドウによる畳み込みの二種類を検討する。

3.2.3 Cross CNN(CrCNN)

はじめに、予測対象の時系列が単調なトレンドと単純な周期を持つことを想定して設計した Cross CNN(CrCNN) を提案する。この手法は、折り返し周期を事前に設定した上で、直近から折り返し周期分の感覚を開けたサンプリングを行った畳み込みと、最後から折り返し周期分のみ注目した畳み込みの二種類の畳み込みを行い、それらの結果を結合して予測結果を出力する。この処理を直感的に示すと、カレンダーのように入力データの折返しを行い、その縦方向と横方向に畳み込み処理をかける手法である。この処理の実例として、入力長が 30 で折り返し周期 7 に設定した場合の畳み込みの実例を図 3.2 に表す。数式で表すと、折返し周期を T として、以下の式 3.2 と表現できる。

$$c = (1 - u)(W_h \mathbf{x}_{[-T]} + b_h) + u(W_v \mathbf{x}_{[:T]} + b_v) \quad (3.2)$$

ただし、入力を \mathbf{x} 、重みを W 、バイアス項を b 、活性化関数を σ 、縦横の畳み込みの重みの比率を u とする。日次にサンプリングしたデータに対しては週に相当する $T = 7$ で、週次でサンプリングしたデータに対しては月に相当する $T = 4$ で折返しを行うことを想定する。

この手法により、単調なトレンドについては縦方向の畳み込みで、周期については横方向の

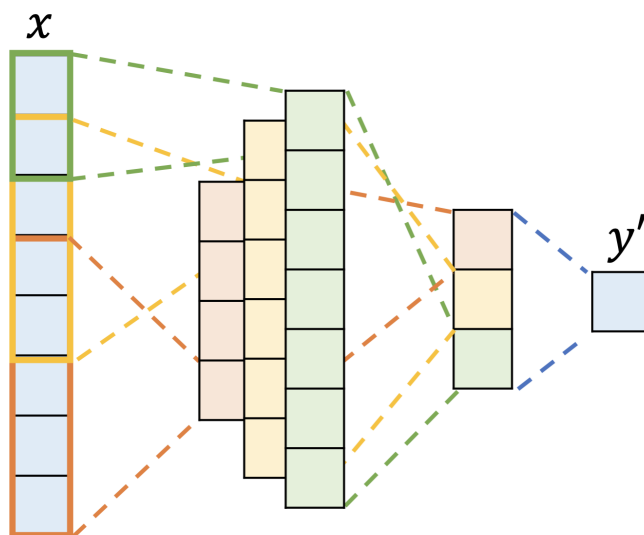


図 3.3 Multi-Window CNN の概念図

畳み込みによって学習できることが期待される。この二方向の畳み込みにより、週や月といった明確な周期に関する知識を予測結果に取り入れられると考えられるが、それ以外の商品ごとに異なる周期を捉えることができないという欠点がある。このことを踏まえ、次に、異なるサイズのウィンドウを用いて畳み込みを行う手法を提案する。

3.2.4 Multi-Window CNN(MWCNN)

先述したように、小売データには複数の周期があることを前提とし、それらをなるべく多く捉えて予測に反映させることが予測精度の向上につながると考えられる。これを踏まえて、この人間が行うような着眼を予測モデルに組み込み直接実装する Multi-Window CNN(MWCNN) を提案する。具体的には、以下の演算を行う。この概念は図 3.3 のように描ける。

$$c = \sum_{k=1}^N w_k \cdot \sigma(W_k * \mathbf{x} + b_k) \quad (3.3)$$

ただし、 W_k のウィンドウサイズは $2^n (n \in N)$ とし、そのウィンドウサイズごとの重みを w で示す。本論文の実験では $k=1 \sim 5$ に設定する。これは、小売データに想定される日足で記録されたデータに対して 1 ヶ月に相当する 30 日分をカバーする 32 をフィルタサイズに設定すべきと考えられるからである。

この手法は先述した CrCNN に比べ、予測に活用する周期を明示的に定義してモデルに反映させたものではないが、商品ごとに異なる周期も捉えて予測に反映させられると考えられる。

3.2.5 線形スキップコネクション

ここまでに周期性を学習できるような CNN 構造を提案したが、[Lai 18] や [Bińkowski 17] でも述べられているように、予測対象に周期性や線形のトレンドが存在することが明らかである場合には、線形 AR モデルの式を直接使えることが予測精度を担保する上で望ましいと考えられる。それを踏まえ、前述の CrCNN や MWCNN を用いた予測と同時に AR モデルによる予測を並行して行う。AR モデルも採用することによって、予測する時系列に強い周期性やトレンドがある場合に、より早期にその特徴を捉えられると考えられるようになると思われる。このような深い深層学習モデルと浅い AR モデルを組み合わせる考え方は、特に [Lai 18] で明示的に導入されているが、異なるタスクにおいても、たとえば推薦アルゴリズムである Wide and Deep Learning[Cheng 16] にも取り入れられている。

[Lai 18] では線形 AR モデルと深層モデル両方の出力を単純に合算したものを最終出力とし、それが実際の値と近くなるように学習を進めるが、本手法では AR モデルを学習させると同時に、深層モデルと線形 AR モデルの重みを合算する際の比率 $u \in [0, 1]$ も学習させ、それらを合算する。線形スキップコネクションの重み及びバイアスを W_l, b_l と置くと、式 3.3 の c に引き続き、出力 y を以下のように書ける。

$$y = (1 - u)(W_l \mathbf{x} + b_l) + u \cdot c \quad (3.4)$$

CPNet は以下に説明する Multi-Window CNN と線形スキップコネクションによって構成され、その概形は図 3.1 で表される。

3.3 実験

本論文における手法が、実際の小売業の時系列予測に対して有効であるかどうかを、複数の小売業のデータに対して予測実験を行うことで検証した。また、ここではラグ期間を複数設定し、ラグの長さの変化による予測結果の変化も検証した。

各モデルの学習は、訓練データを用いて行う。ハイパーパラメータが設定できるモデルについては、異なるハイパーパラメータの組み合わせごとに学習を行い、バリデーションデータによる予測結果を比較することで、最適なパラメータを選択する。最後に、モデルの最終的な予測精度を、テストデータを用いた予測精度によって検証する。

3.3.1 データセット

本研究では小売業に関する公開されている実データとして Rossmann と Walmart と Favorita、そして人工的に作成したデータを用いて実験を行った。また、Rossmann と Walmart、および Favorita は、実店舗において顧客が商品を購入したログデータを用いた実測の売上データであり、系列の一部に欠損が含まれていたため、その前後の値を用いて線形補完を行っている。

Rossmann

ヨーロッパに展開するドラッグストアである Rossmann による、店舗ごとの日次の売上総額を記録したデータセット^{*1}の訓練用データセット (train.csv) における、Sales カラム時系列データである。店舗数は 1155 である。この売上データは、日曜日が定休日であるため七日ごとに売上が必ず 0 になるという特徴があるため、事前に日曜日に相当する分の売上データは取り除き、予測の対象としないこととする。例としてそのデータの概形を図 3.4 に表す。訓練データ期間は 2013/1/1~2013/12/31、バリデーションデータ期間は 2014/1/1~2014/12/31、テストデータ期間は 2015/1/4~2015/7/31 である。説明変数となるモデルへの入力長は 90 (三ヶ月相当) とし、目的変数としてラグ期間を挟んだ後の売上を予測する。予測までのラグは 30 (一ヶ月相当) および 60 (二ヶ月相当) に設定する。

Walmart

アメリカに展開するスーパーマーケットである Walmart の、店舗ごとの週次の売上総額を記録したデータセット^{*2}の訓練用データセット (train.csv) における、Weekly_Sales カラム時系列データである。その一部の概形を図 3.5 に表す。このデータセットには店舗 (Store)45 店分のデータセットが含まれるが、それがさらに 81 種類の部門 (Dept) コードが別に付与されているので、店舗ごとではなく各店舗の部門ごとに、週次での売上予測を行う。訓練データ期間は 2010/2/5~2011/9/30、バリデーションデータ期間は 2011/10/07~2012/3/30、テストデータ期間は 2012/4/6~2012/10/26 である。説明変数となるモデルへの入力長は 24 (六ヶ月相当) とし、目的変数としてラグ期間を挟んだ後の売上を予測する。予測までのラグは 8 (二ヶ月相当) および 12 (三ヶ月相当) に設定する。

^{*1} <https://www.kaggle.com/c/rossmann-store-sales>

^{*2} <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting>

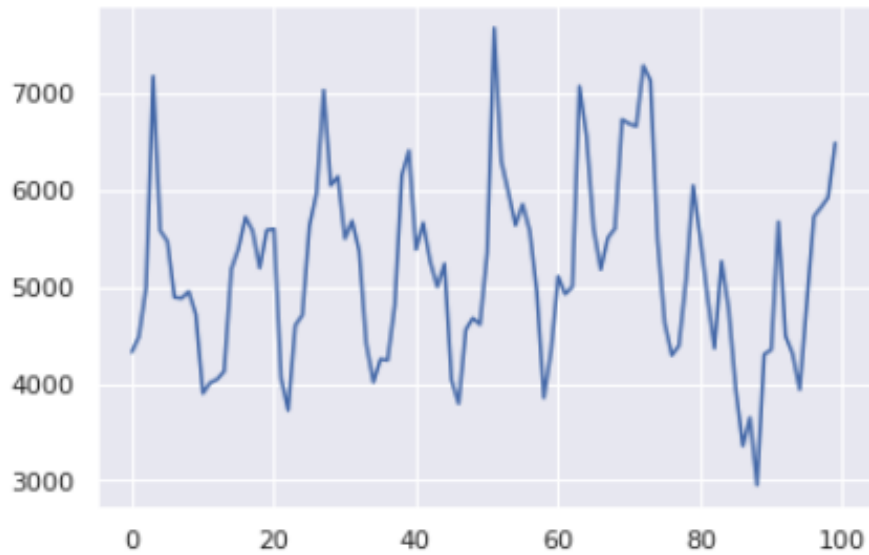


図 3.4 Rossmann データの一部概形

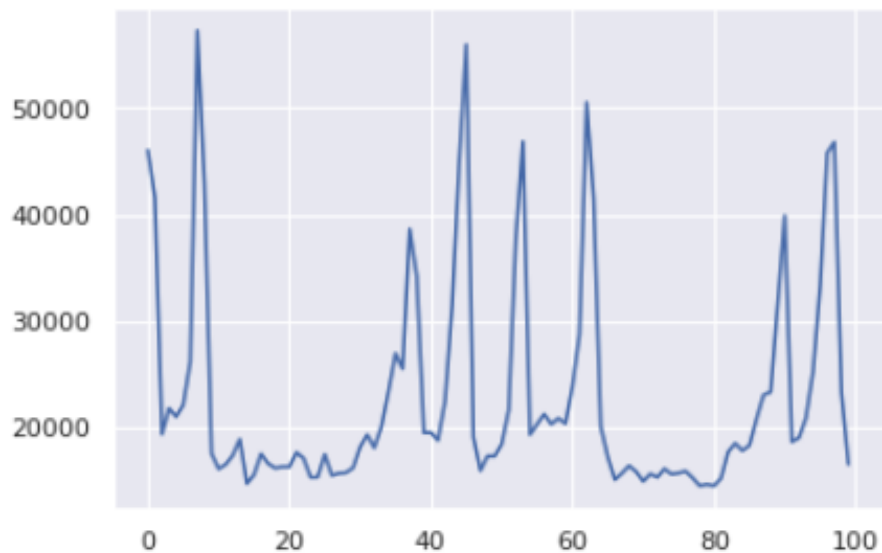


図 3.5 Walmart データの一部概形

Favorita

エクアドルに展開する大手食料品小売企業である Corporacin Favorita による、店舗ごとの日時の商品別販売数を記録したデータセット^{*3}の訓練用データセット (train.csv) における、unit_sales カラム時系列データである。対象となる店舗別販売商品数のレコードは全部で

^{*3} <https://www.kaggle.com/c/favorita-grocery-sales-forecasting>

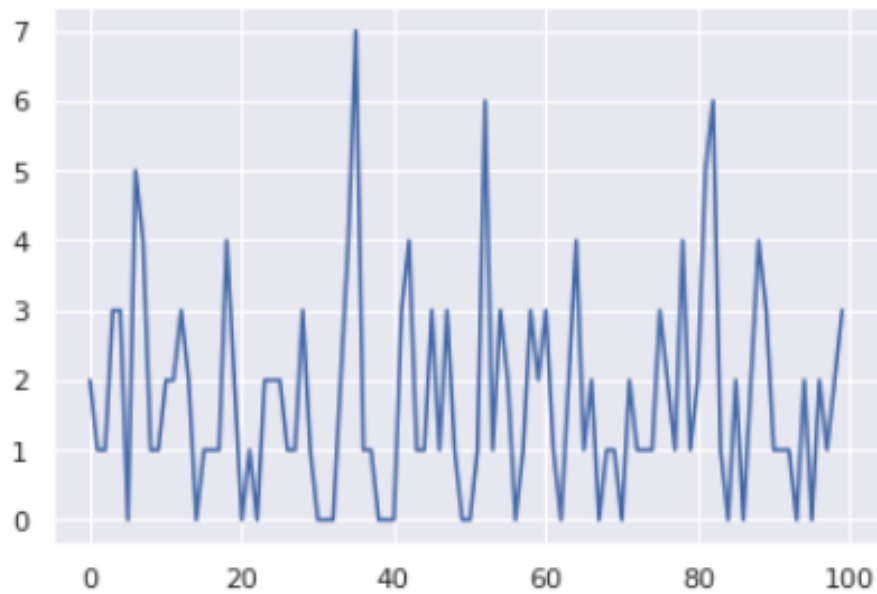


図 3.6 Favorita データの一部概形

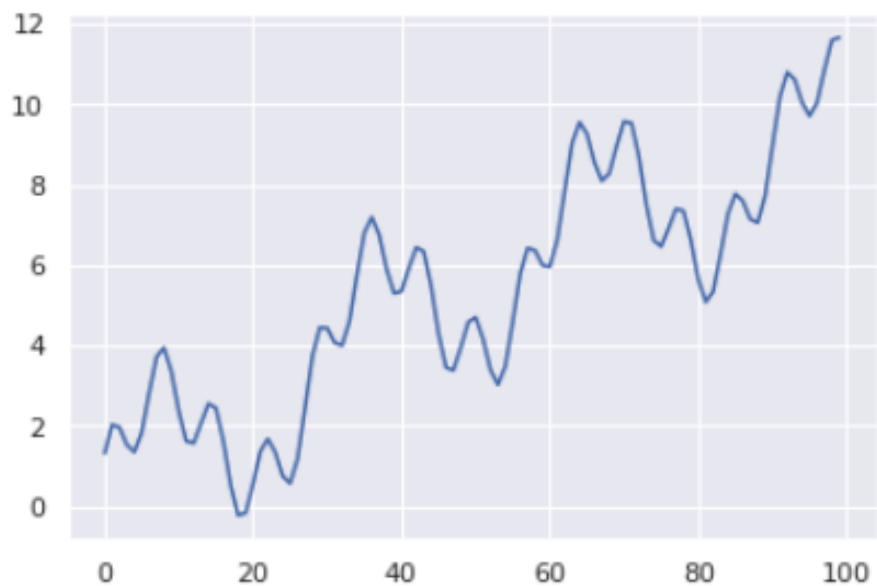


図 3.7 人工データの一部概形

105635 種類あるが、それらを全て予測に用いるとメモリ不足が発生するモデルがあるため、そのうちの 20% に相当する 21127 商品のレコードを無作為に選出し、それらを使用する。例としてそのデータの概形を図 3.6 に表す。訓練データ期間は 2013/1/1~2013/12/31、バリデーションデータ期間は 2014/1/1~2014/6/30、テストデータ期間は 2014/7/1~2014/12/31 である。説明変数となるモデルへの入力長は 90 (三ヶ月相当) とし、目的変数としてラグ期間を挟ん

だ後の売上を予測する。予測までのラグは 30（一ヶ月相当）および 60（二ヶ月相当）に設定する。

人工データ

以上の公表されている実世界での時系列データの他に、以下の式で表される時系列を作成した ($t \in \mathbf{N}$)。

$$x(t) = \sin\left(\frac{2\pi}{7}t\right) + 2 \sin\left(\frac{2\pi}{30}t\right) + 4 \sin\left(\frac{2\pi}{365}t\right) + 0.0005t \quad (3.5)$$

これは、人間の活動によって生み出される時系列データが週次（7日）、月次（30日）、年次（365日）の周期と右肩上がりのトレンドが含まれる時系列であることを仮定して、この特徴が単純に反映された時系列データを直接正弦波を用いて実装したものである。その一部の概形を図 3.7 に表す。訓練データ期間は $t=0\sim 4379$ 、バリデーションデータ期間は $t=4380\sim 5839$ 、テストデータ期間は $t=5840\sim 7299$ をとる。説明変数となるモデルへの入力長は 90（三ヶ月相当）とし、目的変数として入力からラグ期間を挟んだ後の売上を予測する。予測までのラグは 30（一ヶ月相当）および 60（二ヶ月相当）に設定する。

3.3.2 実験設定

本研究では提案手法との比較対象として線形自己回帰モデル (AR)、Prophet、XGBoost、多層パーセプトロン (MLP)、CNN、RNN、LSTNet、SOCNN、DeepAR を設定し、それらとの比較実験を行った。また、MWCNN と線形スキップコネクションそれぞれの予測性能を評価するために、提案手法から線形スキップコネクションを省略したモデルである CrCNN 及び MWCNN のみによる予測実験も行った。

各モデルの損失関数には二乗平均平方根誤差 (RMSE) を採用し、それが最小化されるよう学習を進める。いずれのモデルも最適化関数は Adam[Kingma 14]、活性化関数は LeakyRelu を採用し、勾配が最大 1.0 になるよう Gradient Clipping を課している。バッチサイズはいずれも 1028 とした上で、Rossmann は 2000、Walmart は 200000、Favorita は 100000、人工データは 20000 エポック学習させた。これらモデルは、AR モデルは $N=7$ 、深層モデルはいずれも入力層、隠れ層、出力層の三層で構築している。中間層のユニット数は、ハイパーパラメータ自動最適化フレームワークである Optuna^{*4}を用いて最適な値を探索した結果、いずれのモデ

^{*4} <https://optuna.org/>

ルも 120 前後が最適であったため、一律で 120 に設定した。ウィンドウサイズは $k = 3, 5, 7$ より最も予測精度が高くなるものを選択する。XGBoost および Prophet, DeepAR において設定可能な各種ハイパーパラメータについても、Optuna によりそれぞれ 50 回の試行により最適な値を探索し、その値を設定した。

3.3.3 評価指標

各モデルの性能の特徴を多面的に評価するために、評価指標は以下の三種類を採用した。正解の値を y 、予測値を \hat{y} とする。

二乗平均平方根誤差 (RMSE)

先述の通り、RMSE(Root Mean Squared Error) は学習時の目的関数でもあり、各予測手法は訓練データにおいてそれが最小となるように学習を進行させる。そして、テストデータにおいてこの値が小さいほど精度の高い予測であると言える。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2} \quad (3.6)$$

平均平方二乗誤差率 (RMSPE)

先述の RMSE は、売上のスパイクなど急激な変化における誤差をそうでない時点と比較して過剰に考慮してしまう傾向がある。このため、正解値と予測値との誤差を正解値で割ることでスケールした値の二乗誤差を見る RMSPE(Root Mean Square Percentage Error) も評価指標として用いる。RMSPE も、テストデータにおいてこの値が小さいほど精度の高い予測であると言える。

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2} \quad (3.7)$$

相関係数 (CORR)

テスト結果における予測結果と実際の値の相関係数を算出し、実際の値と予測結果の関係性を確認する。この値が 1 に近いほど、予測値と実際の値の大小関係の一致具合が高く、よりよい予測であると言える。

$$CORR = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.8)$$

表 3.1 テスト結果のまとめ (最も良い結果のものを太字で表記)

データセット		Rossmann		Walmart		Favorita		人工データ	
		ラグ期間		ラグ期間		ラグ期間		ラグ期間	
モデル	評価指標	60	90	8	12	60	90	30	60
AR	RMSE	2117.8	2414.4	8705.6	9284.5	0.9293	0.9731	0.133	0.142
	RMSPE	0.277	0.347	0.865	0.876	1.426	2.912	26.78	25.94
	CORR	-0.131	-0.147	0.052	-0.101	0.025	0.014	0.673	0.694
Prophet	RMSE	1861.5	1876.4	17.998	18.065	0.1412	0.1773	0.041	0.046
	RMSPE	0.223	0.281	0.166	0.264	1.122	1.561	4.422	5.782
	CORR	0.554	0.511	0.446	0.390	0.392	0.262	0.710	0.695
XGBoost	RMSE	1980.6	2012.4	25.881	26.229	0.1759	0.1561	0.066	0.052
	RMSPE	0.272	0.336	0.171	0.229	2.285	2.447	5.886	6.211
	CORR	0.456	0.441	0.420	0.363	0.282	0.231	0.588	0.624
DNN	RMSE	2144.4	2331.3	442.15	621.96	0.0075	0.0076	0.512	0.528
	RMSPE	0.272	0.282	1.480	1.210	1.167	1.771	9.238	9.927
	CORR	0.144	0.161	0.221	0.171	0.401	0.382	0.337	0.227
CNN	RMSE	2666.1	2641.3	521.97	518.85	0.7023	0.7043	0.372	0.536
	RMSPE	0.411	0.458	0.239	0.290	3.556	4.741	10.466	8.161
	CORR	0.216	0.149	0.126	0.121	0.365	0.298	0.789	0.689
RNN	RMSE	3038.8	3680.8	2219.8	5462.5	0.0556	0.0605	0.538	0.759
	RMSPE	0.691	0.882	0.977	0.953	4.628	6.925	8.877	8.796
	CORR	-0.053	-0.043	-0.009	-0.018	0.022	0.017	0.126	0.112
LSTNet	RMSE	2735.6	2868.0	18.526	17.758	0.0575	0.0578	0.209	0.250
	RMSPE	0.291	0.288	0.147	0.299	0.923	1.112	26.33	29.42
	CORR	0.267	0.354	0.226	0.243	0.342	0.388	0.512	0.551
SOCNN	RMSE	2036.7	2325.5	18.046	16.251	0.7196	0.7156	0.039	0.040
	RMSPE	0.276	0.286	0.387	0.408	1.183	0.997	26.18	26.64
	CORR	0.373	0.364	0.399	0.416	0.352	0.382	0.694	0.621
DeepAR	RMSE	2140.4	2405.5	20.09	21.05	0.0480	0.0492	0.033	0.039
	RMSPE	0.327	0.462	0.236	0.258	1.266	1.678	4.727	4.998
	CORR	0.511	0.488	0.486	0.456	0.254	0.250	0.661	0.622
CrCNN	RMSE	2549.7	2633.1	180.92	184.61	0.0168	0.0198	0.051	0.053
	RMSPE	0.489	0.494	0.372	0.381	0.781	0.664	25.33	26.72
	CORR	0.332	0.343	0.388	0.376	0.414	0.355	0.409	0.442
CrCNN+AR	RMSE	1809.8	2132.2	18.597	16.942	0.0187	0.0205	0.044	0.058
	RMSPE	0.258	0.269	0.189	0.212	0.536	0.628	5.445	6.778
	CORR	0.470	0.443	0.446	0.479	0.454	0.506	0.737	0.643
MWCNN	RMSE	2208.5	2368.2	17.986	17.755	0.0012	0.0017	0.038	0.039
	RMSPE	0.400	0.391	0.572	0.596	0.698	0.776	26.13	28.32
	CORR	0.496	0.449	0.395	0.381	0.389	0.360	0.179	0.085
CPNet (MWCNN+AR)	RMSE	1486.9	1512.9	17.672	16.519	0.0006	0.0007	0.031	0.036
	RMSPE	0.237	0.268	0.149	0.199	0.542	0.556	4.569	4.824
	CORR	0.578	0.532	0.478	0.481	0.516	0.487	0.795	0.696
MWCNN +Prophet	RMSE	1576.6	1980.8	18.915	20.27	0.0018	0.0021	0.036	0.055
	RMSPE	0.360	0.411	0.268	0.321	0.551	0.578	6.146	7.272
	CORR	0.517	0.466	0.414	0.429	0.463	0.428	0.563	0.497

3.3.4 実験結果

実験結果を表 5.2 に記す。一つを除き、提案手法である CPNet が高いスコアを示していることが確認された。ラグ期間を広げた場合には、全体として予測精度が低下する傾向があることも確認された。

CrCNN および MWCNN は通常の CNN と比べ、特に人工データの場合に大きく RMSE が改善されていることから、予測対象データに含まれる周期性が明確な場合に、より有効な手法であると考えられる。そして、線形スキップコネクションを用いた場合、CrCNN と MWCNN のどちらの場合においてもそれを使用しなかったモデルに比べて予測精度が向上していることから、その有効性が確認できる。一方で、RNN ではどの実験においても良い精度が出ずに他のモデルに比べて精度が悪くなっている。これより、RNN は予測対象までのスキップ期間が空いてしまう時系列予測タスクにおいては予測に不向きであることから、小売データの予測において CNN に比べて不向きであるとも考えられる。

また、CrCNN、CrCNN+AR、MWCNN 及び CPNet(MWCNN+AR) における Walmart (ラグ期間 8) データ予測時のエポックごとの RMSE の変化を図 3.8 に記す。線形スキップコネクション (AR) を併用するモデルは、そうでないモデルに比べて早急に収束することが分かる。

さらに、表 5.2 には、MWCNN のスキップコネクションとして線形モデルの代わりに Prophet を用いたモデル (MWCNN+Prophet) も作成し、それによる予測実験を行った結果を記している。結果からも分かるように、このモデルは MWCNN と比べて大きな改善が見られなかった。一般に、線形モデルに加えて Prophet 側のほうが容量が大きいことからこのモデルのほうが MWCNN よりも精度が高くなるのではないかと考えられるが、本来 Prophet で学習することが期待される通常の線形モデルでは捉えることができない複数の周期を、既に MWCNN 側で学習して予測に反映できていることが理由であると考えられる。ただし、Prophet は祝日やキャンペーンと言った非周期性の突発的なイベントに関する情報を予測に反映させることができる機構を備えており、非周期生の情報は情報を MWCNN は捉えることが出来ないことから、予測にこれらの情報を用いることができる際には、この Prophet を一部使用したモデルを検討する余地があると考えられる。

各予測モデルにおいて RMSE および RMSPE の傾向を比較すると、両方の大小関係は各モデルや各データに対して概ね一致していることが確認される。一方で、特に人工データを用いた実験結果に注目すると、線形スキップコネクションを加えたモデルは深層学習のみを用いたモデルに比べて RMSE の下がり具合と比較して RMSPE がより大幅に下がっていることも確認される。周期状やパルス状といった短時間に大幅な変化がある時系列を予測する際、RMSE が同程度で RMSPE が大幅に異なっていた場合、RMSPE が大きくなっている予測モデルはそうでないモデルに比べてその大幅な変化を予測できていないと考えられる。たとえば、周期性のある単純な時系列に対しその移動平均をとって予測するモデルを用いて予測をすると、その周期性の変化が平滑化されて RMSE は小さくなったとしても、周期性の変化が捉えられない

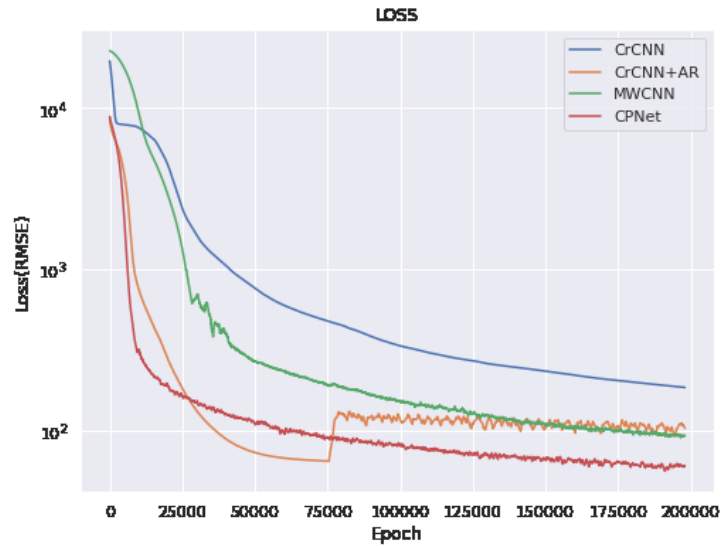


図 3.8 線形スキップコネクションを用いたモデルにおける RMSE の変化 (Walmart、ラグ期間 8)

ことで RMSPE が大きくなる傾向がある。このことから、線形スキップコネクションと深層モデルの併用したモデルのほうが大幅な変化をより正確に捉えており、予測対象とする時系列の大小の傾向だけではなく周期性そのものがより正確に捉えられていることが示唆される。

3.4 まとめ

本論文では、小売業での実務利用のための、深層学習と AR モデルを活用した時系列予測モデル (CPNet) を提案した。複数周期を捉える Multi-Window CNN と、AR モデルを並行させる線形スキップコネクションによる時系列予測モデルであり、実際にその精度が従来手法よりも高いことを実験により確認した。本論文の提案手法は、予測対象とする時系列の周期の把握をする必要がなく、予測までのラグ期間に依らずに一定の精度が期待できるため、特徴量の設計を行わずに深層学習を用いて直接将来予測する場合におけるベースライン手法になると考えられる。

第4章

ユーザの行動履歴データを用いたコンテンツ興味推移カテゴリの可視化手法

本研究では、Web サービスなどのユーザの行動が記録されたログデータを用いて、ユーザの興味対象がどう推移するかを表すカテゴリを可視化する手法を提案する。あるコンテンツに対し興味を示している証拠となる行動を興味行動と定義し、ログデータからユーザの興味行動の時系列を抽出、ユーザをソフトクラスタリング、そして各クラスタ内のユーザの興味推移状況をレベルで表すことで、クラスタごとに興味対象がどう推移するかを可視化する。従来手法による可視化では、各ユーザは単一カテゴリへの興味しか持たないとみなしてしまうために、本来分けられるカテゴリを重ねてしまったり、そもそも可視化されないカテゴリが発生する可能性があった。そのため、本提案手法では、ユーザが同時に複数のジャンルに興味を持つという前提を考慮し、新たにソフトクラスタリングを導入している。検証実験の結果、本手法によって消費者の興味推移動向が従来手法よりも詳細に可視化でき、商品の推薦にも活用できる可能性があることを示した。本手法によってボトムアップにクラスタが可視化されることによって、人間では考えつかなかったクラスタを把握できるようになり、より効率的なターゲティング施策の立案にも活用できると考えられる。

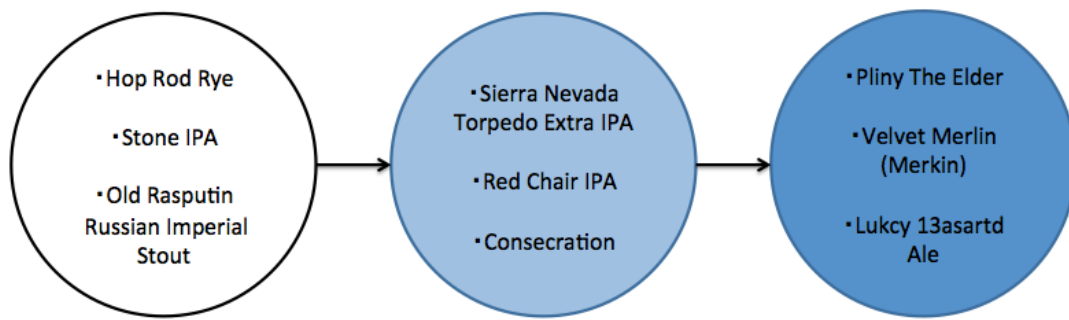


図 4.1 海外ビール市場の潜在的カテゴリ構造と其の変化イメージ (Yang ら [Yang 14] より作成)

4.1 はじめに

現代のマーケティングでは、市場内の消費者を複数個のグループに分割し、その中からターゲットとするグループを選択する手法が行われている。それらの手法は、セグメンテーション（市場細分化）とターゲティングと呼ばれる。セグメンテーションの切り口には、年齢や職業、居住地域などがあるが、それぞれのセグメントに属する個々人の趣味嗜好や消費行動がそれぞれ異質であることが重要である。良いセグメンテーションができているときには、そのセグメントごとに趣味嗜好の対象になる商品がそれぞれ異なるため、それら商品をカテゴリとして分類することができる。また、逆に、そのようなカテゴリを特定することができれば、そのカテゴリを嗜好する消費者を集めることで、マーケティングを行う上で有効なセグメントを作成することができると考えられる。そして、そのアプローチは、消費者を複数のセグメントに属させることを可能にする。

商品をカテゴリ構造に分類する方法については、消費者行動論の研究において複数の種類があることが知られている [新倉 05]。その一つに、各コンテンツを商品ごとの特性に応じて既定のカテゴリに割り振ることで構成する分類学的カテゴリがある。ただし、分類学的カテゴリは、特にコンテンツ産業における商品では、有効に機能しない場合がある。たとえば、伊藤によると、漫画にはスポーツものや SF といった典型的なジャンル（カテゴリ）は一応は存在するが、漫画家はそのカテゴリの境目に位置づけられる漫画を制作しようとするため、ここで述べられるようなジャンルは意図的に破壊されて、拡張、変化し続けるものであると述べられている [伊藤 05]。また、分類学的カテゴリとは対称的にボトムアップに形成されるカテゴリと

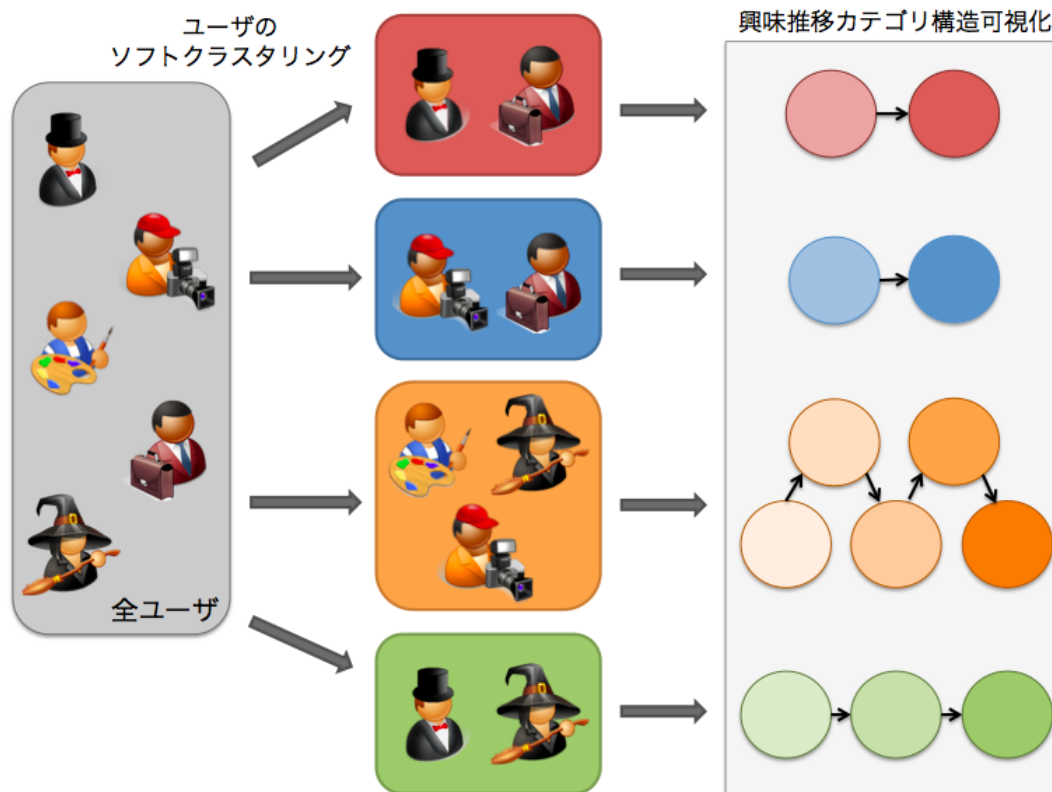


図 4.2 本研究の提案手法の模式図

して、消費者のコンテキスト*¹によって生まれるカテゴリがあると指摘されており、そのカテゴリはアドホック・カテゴリと呼ばれる [新倉 05]。たとえば、マンガの場合は、「泣けるマンガ」「人に薦めたいマンガ」などが挙げられる。アドホック・カテゴリは消費者の意識や行動によって生まれるカテゴリであり、その構造やカテゴリ内に含まれる商品も消費者によって異なるため、必ずしも生産者側から正確に認識することはできないが、コンテキストの近い消費者間ではその認識が共有される可能性が高くなるため、アドホック・カテゴリに着目すれば、そのカテゴリへの認識を共有している市場内セグメントを発見できる可能性がある。

以上を踏まえると、効果的なセグメンテーションを実現するカテゴリの発見は、市場における消費者の実際の趣味嗜好や消費行動から特定することが有効であると考えられる。ただし、消費者から直接そのカテゴリを特定しようとする、属人的に消費者を直接調査する必要があるため、コストがかかる。他にも、海外市場でそれを行う場合、社会制度や文化が日本と異なる場合があり、コンテキストを把握すること自体が困難な可能性がある。

さらに、セグメントを取り巻くコンテキストが日々変化するということや、各消費者がコン

*¹ 消費者の置かれている状況や、その背景のこと。

コンテンツの鑑賞経験を重ねることでカテゴリへの造詣が深まり嗜好する商品が変化していくなどの理由から、セグメント内のユーザの趣味嗜好は常に固定されているものではなく、年月が経つにつれ動的に変化するものである [McAuley 13]。もし、その変化の傾向を理解できれば、ユーザが将来に嗜好する商品が予測できるようになるため、より精度の高い商品の推薦ができるようになるなど、様々な応用も期待される。

過去に行われたユーザの実際の行動からボトムアップに対象物のカテゴリとそのカテゴリ内の趣味嗜好の変化を可視化しようとした試みとして、Yang らの研究 [Yang 14] が挙げられる。Yang らがビール情報サイトユーザのビールの嗜好の変化からカテゴリ構造を可視化した例を、図 4.1 に示す。Yang らは [Yang 14] において、ユーザを行動の変化のパターンによってクラスタリングし、さらに、ユーザにレベルを付与することで、各クラスタの各レベルのユーザにどのような酒類が好まれるかなどを集計した。その結果、ボトムアップに対象物のカテゴリ構造を可視化し、ユーザの趣味嗜好が年月が経つにつれて変化していくことを示した。

しかし、Yang らの手法は、可視化を行う上でユーザをそれぞれ単一のグループに属させるため、ユーザが複数のカテゴリへの趣味嗜好を持つことを想定していないものである。そのため、この手法をそのままコンテンツの興味推移を可視化するために用いると、多くのユーザが一つのカテゴリ内の商品のみに興味嗜好を持つわけではないため、本来分けて可視化することができるカテゴリ構造を重ねた状態で可視化してしまったり、多くのユーザが趣味嗜好を持つものの、ユーザごとにそれへの興味を示す行動の回数が少ないカテゴリ構造が可視化されなかったりする可能性がある。ユーザが複数のカテゴリへの趣味嗜好を持つことを考慮するためには、1 ユーザを複数のクラスタに当てはめられるソフトクラスタリングを導入することが考えられるが、Yang らの手法 [Yang 14] では、1 ユーザを 1 クラスタのみに当てはめるハードクラスタリングとユーザのレベル判定が同時に行われる。そのため、ソフトクラスタリングを導入するためには、Yang らの手法からレベルの判定のプロセスを分離させ、ソフトクラスタリングの後に行うよう修整する必要がある。ただし、ユーザのクラスタリングとレベル判定を分離できれば、各クラスタごとにレベル上限を個別に設定できるようになるため、それぞれに最適なレベル上限を設定できれば、より過不足の少ない可視化が実現できるようになると考えられる。

以上の内容を踏まえ、本研究では、Web サービス等のユーザの行動が記録されたログデータを用いて、ユーザの興味対象がどう推移するかを表すカテゴリをボトムアップに可視化する手法を提案する。あるコンテンツに対し興味を示している証拠となる行動を興味行動と定義し、ログデータからユーザの興味行動の時系列を抽出した後、図 4.2 に表されるユーザのソフトク

ラスタリング、そして各クラスタ内のユーザの興味推移状況をレベルで表すことで、クラスタごとに興味対象がどう推移するかを可視化する。

本提案手法の新規性と有用性は、以下の通りである。

- 従来手法は1ユーザを1クラスタのみに当てはめるハードクラスタリングを用いることでコンテンツのカテゴリ構造を可視化しているが、本手法はユーザが同時に複数のジャンルに興味を持つという前提を考慮して、ソフトクラスタリングを用いた設計となっている。このことによって、市場の状況をより正確に反映した可視化が実現できる。
- 本手法によって得られる可視化結果は、コンテンツ提供側から把握することが難しい消費者の実際の興味動向に基づくボトムアップなカテゴリを可視化したものである。それは市場の実態に則したものであるため、実際のマーケティング活動においても有益な示唆をもたらすものである。

4.2 提案手法

従来手法によってユーザのコンテンツ興味推移を可視化すると、各ユーザが単一カテゴリへの興味しか持たないとみなしてしまうことから、可視化の結果にコンテンツの重複や欠落が生じるという問題がある。本手法は、ユーザが同時に複数のジャンルに興味を持つという前提を考慮し、ソフトクラスタリングを導入、さらに、各クラスタごとの最大興味レベルを一律にせず、それぞれ最適な値を設定することによって、従来手法に比べて過不足が少なくコンテンツが表示できる可視化を実現する。

本手法の概要は、以下の流れで説明できる。

1. ユーザごとの興味集合ベクトルと興味推移シーケンスの抽出
2. Fuzzy C-means 法を用いた興味集合ベクトルのソフトクラスタリング
3. 各ユーザごとの興味推移グラフの作成
4. 各クラスタごとの所属する各ユーザの興味推移最適経路の探索
5. クラスタ内の全ユーザの興味レベルごとの上位コンテンツの集計
6. 表示コンテンツが最大となるような各クラスタごとの可視化結果の選択

本章では以降、提案手法の前提を述べた後に、手法の詳細を説明する。

4.2.1 前提

用語の定義

本手法では、各ユーザが興味を抱いているコンテンツ集合を興味ジャンル、各ユーザの興味ジャンルごとの造詣の深さ（成熟度）を興味レベル、各ユーザがあるコンテンツに対し、それに興味を持っていたことで行われた行動を興味行動と定義する。興味行動の例としては、コンテンツの購入や、コンテンツに対する好意的なレビューの投稿などが挙げられる。ただし、商品の売り切れのようなコンテンツに興味があっても興味行動が起こせないケースは今回は想定しない。

ユーザの性質の前提

本手法では、ユーザは興味レベルが変化するに伴い、強く興味を示すコンテンツが変化していくものとする。たとえば [McAuley 13] において、アルコール類のレビューサイトである RateBeer^{*2}において、飲酒の経験が深まるに連れて、よりアルコール度数の高い酒類を高く評価するようになることが確認されている。

さらに、各ユーザの興味について、[Yang 14] で用いられる以下の前提を置く。

- 各ユーザは、興味ジャンルを 1 つ以上持つ。
- 各ユーザの興味レベルは時間経過につれ 1 レベルずつ上昇するが、下がることはない。

各ユーザの持つ興味ジャンルごとの興味レベルはそれぞれ独立であるため、それぞれ値が異なる場合がある。また、各ユーザの興味が必ずしも最初の興味レベル（レベル 1）から始まるものではないことに注意する。なぜなら、ユーザがログデータを収集する Web サービス上で興味行動を示す以前に、既に興味レベルが高くなっている可能性があるからである。そして、興味レベルが進展するペースもユーザやその興味ジャンルによって異なり、場合によっては時間が経過しても全く興味レベルが増加しない場合もあり得るものとする。ある興味ジャンルにおける、実際のユーザごとの時間経過における興味レベルの変化の例を、図 4.3 に示す。

^{*2} RateBeer: Great beer made easy <http://www.ratebeer.com/>

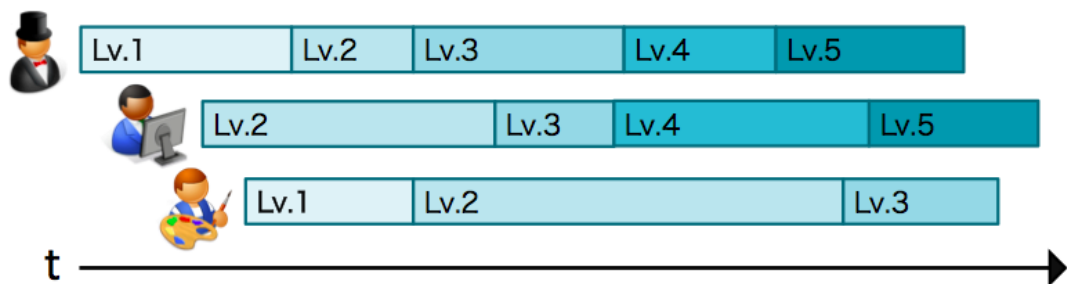


図 4.3 各ユーザの興味レベル進展イメージ

データセットの条件

本手法によってログデータから興味推移動向を可視化するには、そのログデータ内に以下の3つの要素が存在することが必要である。

- 複数のユーザ
- 各ユーザの複数のコンテンツへの興味行動
- 各興味行動のタイムスタンプ^{*3}

これら上記条件は、ログデータ内に複数のユーザについて各ユーザ別の複数のコンテンツへの興味行動の時系列が必要であると総括できる。この条件を満たすデータセットであれば、汎用的に本提案手法を適用できる。

4.2.2 素性の作成

はじめに、行動ログデータから各ユーザの興味行動を収集する。収集に際し、各ユーザにつき以下の2種類のデータを作成する。前者はユーザの興味カテゴリを特定するためのクラスタリングに、後者は各ユーザの興味カテゴリごとの興味レベルを判定する際に用いる。

興味集合ベクトル: 興味行動をとったコンテンツの集合

興味推移シーケンス: 興味行動があったコンテンツをタイムスタンプ順に並べたデータ

興味集合ベクトルは、ログデータ内に存在する各コンテンツごとに、そのユーザが興味行動がとられている場合を1、取られていない場合を0とした、One-Hotベクトルとする。各ユーザ

^{*3} 発生日時及び時刻のこと。

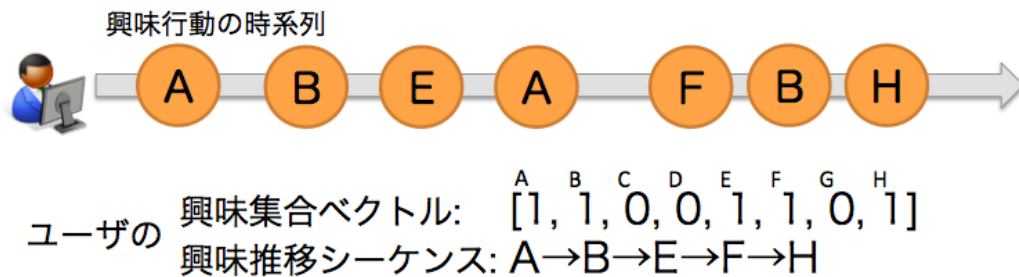


図 4.4 興味行動時系列からのデータ収集の例

から興味推移シーケンスのデータを抽出する際、1つのコンテンツに対し複数回の興味行動をとっていた場合、最初の1回目のみを興味行動とカウントし、そのコンテンツに対するそれ以降の興味行動は無視する。この収集方法の例を図 4.4 に示す。

4.2.3 ユーザの所属クラスタの特定

本項では、ユーザのクラスタリング手法である Fuzzy C-means 法 [Dunn 73] を説明する。Fuzzy C-means 法はソフトクラスタリング手法であり、各ノードを単一のクラスタに割り当てるハードクラスタリングとは異なり、各ノードに対し、各クラスタへの所属度を算出する。

\mathbf{x}_i をノード（合計 N 個）、 μ_k をクラスタ k のセントロイド（重心）としたとき、 \mathbf{x}_i のクラスタ k への所属度 u_{ki} は、式 4.1 で求められる。

$$u_{ki} = \left[\sum_c \left(\frac{\|\mathbf{x}_i - \mu_k\|^{\frac{2}{m-1}}}{\|\mathbf{x}_i - \mu_j\|^{\frac{2}{m-1}}} \right) \right]^{-1} \quad (4.1)$$

m はファジイさを定義するパラメータである。このとき、各ノードについて、所属度の総和が 1 になることに注意する。また、各クラスタのセントロイド μ_k は式 4.2 で計算される。

$$\mu_k = \frac{\sum_i^N (u_{ki})^m \mathbf{x}_i}{\sum_i^N (u_{ki})^m} \quad (4.2)$$

式 4.1 と式 4.2 の計算を交互に繰り返すことでソフトクラスタリングを進め、各ノードの所属度の変化が閾値を下回ったときに処理を終了する。すなわち、以下のステップを繰り返す。

1. 各ノードごとに、 k 個のクラスタごとの所属度を、その合計が 1 になるように無作為に設定する。
2. 各ノードのクラスタへの所属度から、クラスタのセントロイドの位置を式 4.2 によって計算する。

3. 各ノードと各セントロイドとの所属度を、式 4.1 によって計算する。
4. ノードの所属度とセントロイドの位置の変化が閾値を下回るまで手順 2、3 を繰り返す。

以上のステップからなるクラスタリングを、各ユーザごとの興味集合ベクトルに対して行う。最後に、各ノードがそれぞれのクラスタに属するか否かを、その出力で示された所属度の値によって判定する。また、以降のユーザごとの興味レベルの特定については、それぞれのクラスタごとに別々に行う。

4.2.4 ユーザの興味レベルの特定

次に、ユーザの興味の推移を興味推移シーケンスから特定する手法を説明する。ここで述べる手法は McAuley ら [McAuley 13] や Yang らの手法 [Yang 14] においても用いられている。また、クラスタ内の興味レベルの上限値 K は、この手法を行う前に設定する。

興味推移グラフの作成

各ユーザの興味推移シーケンスから、以下の手順でグラフを作成する。

1. スタートノード、ゴールノードを 1 つずつ作成する。
2. 各興味行動につき、それぞれ K 個のノードを作成する。
3. 前の興味行動のノードから後の興味行動のノードへエッジを張る。その際、同一興味レベルの興味行動と、1 つ上の興味レベルの興味行動の 2 方向にエッジを張る。ただし、興味レベルが K である興味行動ノードは、興味レベルが K である後の興味行動ノード 1 つのみにエッジを張る。
4. 最初の各興味レベルの興味行動ノードとスタートノード間と、最後の各レベルの興味行動とゴールノード間にエッジを張る。

例として、 $K = 3$ と設定したときに図 4.4 の興味推移シーケンスから作成されるグラフを、図 4.5 に示す。

コスト関数の設定

次に、各興味行動ノードを通過する際のコストを設定する。そのコストの設定には [Yang 14] 内で用いられる関数を採用する。コンテンツ r を興味レベル q で通過するときのコスト $\Theta(q)_r$

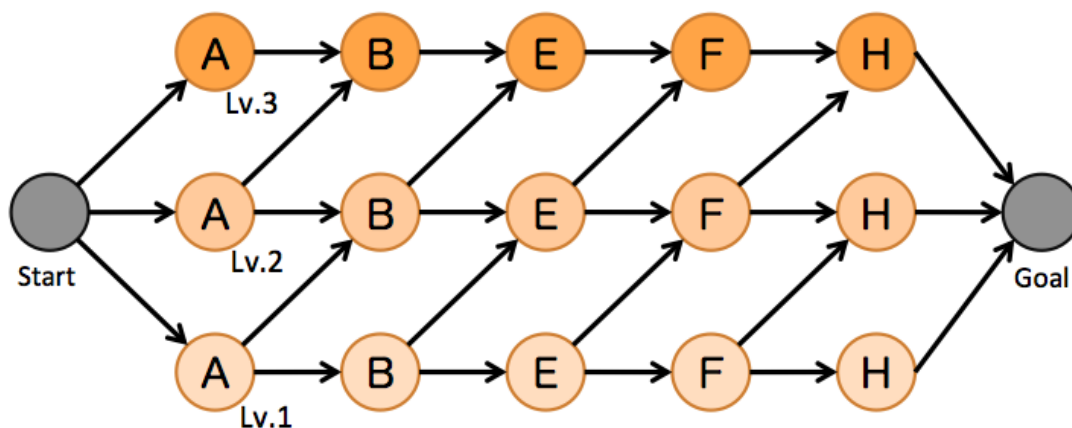


図 4.5 $K = 3$ 時に図 4.4 の興味推移シーケンスから作成されるグラフ例

表 4.1 検証実験用データセットの概要

コンテンツ	データセット	ユーザ数	対象コンテンツ数	平均興味行動回数
映画	MovieLens(1M)	1200	800	72.75
マンガ	Wikipedia 日本語版編集履歴	1871	179	6.34

は、式 4.3 によって計算する。

$$\Theta(q)_r = \frac{\lambda + \sum_i \mathbb{I}\{s_i = q \wedge x_i = r\}}{M\lambda + \sum_i \mathbb{I}\{s_i = q\}} \quad (4.3)$$

式 4.3 における s は興味レベル、 x はコンテンツ、 M は興味推移ベクトル内に登場しうるコンテンツの総種類数であり、 \mathbb{I} は真のとき 1、偽のとき 0 を示す指示関数である。また、 λ はハイパーパラメータであるが、Yang らの研究結果 [Yang 14] において、実験に用いた全種類のデータセットで $\lambda = 1$ が最適であったと述べられているため、それを採用する。式 4.3 で示されるコスト関数は、コンテンツごとに、それを通過するときの興味レベルによってコストが異なり、それは各コンテンツごとにその興味レベルで通過する他のユーザが多ければ多いほど減少する。

最適経路の探索による最適レベルの探索

各ノードのコスト関数が決定された後、各ユーザの興味推移グラフごとに、コストの和が最小となる最適経路を探索する。そして、全シーケンスの最適経路を更新し終えた後、その更新された各ユーザの経路をもとに、各コンテンツにおけるコスト関数を式 4.3 によって再計算する。そして、その新たに設定されたコスト関数の条件下で、再度コストを最小とする経路を探

素により求める。このプロセスを、全グラフの最適経路が変更されなくなるか、全最適経路のコストの和の減少量が閾値を下回るまで繰り返す。

興味レベルごとのコンテンツの集計

Fuzzy C-means 法で得られた各クラスごとに、各ユーザの最適経路を求める。そして、各ユーザごとに、どの興味レベルにおいてどのコンテンツが通過されたかを集計し、その結果をクラスタ内の全ユーザについて合算する。その後、各興味レベルごとに、出現頻度が高い上位のコンテンツを集計する。その上位に含まれたコンテンツを、このクラスタの当興味レベルにおいて興味関心が高くなっている代表的なコンテンツであるとする。この手順を、最適経路探索前に設定する最大興味レベル K を変更しながら、複数回行う。

クラスタごとの結果の結合

最後に、全体で表示されるコンテンツ総数が最大になるように、各クラスタから可視化結果を1つずつ選択する。それらを全て示すことで、市場全体の興味推移の可視化が完了する。

4.3 評価実験

本章では、本提案手法によって、既存手法より実用性の高い消費者の興味推移の可視化ができることを、実用性指標によるコンテンツのカバー状況と可視化結果を用いたユーザの興味行動の予測による、二種類の検証実験によって示す。本手法との比較対象には、Yang らの手法 [Yang 14] を採用し、可視化されるコンテンツは、各クラスタの各興味レベルごとに登場頻度が高かった上位 10 件とした。

4.3.1 データセット

実験用データセットには、MovieLens(1M) データセット^{*4} 及び Wikipedia 日本語版のコンテンツページ編集履歴データを採用した。これらのデータは第 3 章で用いたような顧客が実際に商品を購入したりコンテンツを鑑賞した販売実績を表すものではなく、特定の Web サービス上でその商品に関するページに対して行動を起こしたログのデータである。両データセットの基本情報を、表 4.1 に記す。

MovieLens(1M) データセットは、1,000,209 件の映画に対する 5 点満点の評価データが含

^{*4} <http://grouplens.org/datasets/movielens/>

まれ、ユーザ計 6040 人、映画計 3900 本が含まれている。そして、ユーザが映画へ 4 点以上の評価を付けたことを、ユーザのその映画に対する興味行動であると判断した。しかし、データセット内の全データを実験に用いると、可視化に際し多大な日時を要するため、本実験では、ユーザとコンテンツそれぞれ約 20% に相当する、無作為に選出したユーザ 1200 人についての、興味行動での登場頻度上位 800 位までの映画のみを対象とした。

Wikipedia 日本語版のコンテンツページ編集履歴データは、2014 年 12 月 11 日時点の Wikipedia 日本語版ダンプデータに含まれる編集履歴データとし、ユーザがコンテンツページを編集することを、そのユーザのそのコンテンツに対する興味行動であると判断した。可視化する対象のコンテンツは、オリコン株式会社の提供する 2013 年 12 月におけるコミック月別売上ランキング上位 200 件のうち、Wikipedia 日本語版ダンプデータにコンテンツページが存在したマンガ 179 種類とした。また、興味レベル判定を行うユーザは、上記対象コンテンツを 2 つ以上編集したことのあるユーザ計 1871 人とした。

4.3.2 カバーしたコンテンツによる評価実験

量的な可視化の良さの判定には、以下に説明する実用性指標 U を新たに導入し、それを用いる。

全体コンテンツカバー数 C : その可視化におけるコンテンツの登場総数。この値が大きいほど、可視化におけるもれが少なく、より広く市場の消費者の興味推移が可視化できていることを示す。

クラスタ内コンテンツ重複数 D : 可視化において特定のコンテンツが同一クラスタの複数の興味レベルに登場した回数。この値が大きいほど、可視化結果に重複（だぶり）が多く発生しているため、興味推移における興味レベルの割り当てが正しくなされていないことを示す。

実用性指標 U : $U = C - D$ とし、 U をその可視化の良さを表す点数とする。

ここで、Yang らの手法 [Yang 14] で用いられる可視化の評価方法を用いずに、新たに実用性指標 U を導入して評価に使用した理由を述べる。まず、各コンテンツがそれぞれ適切なカテゴリに正しく分類できているかを評価する際には、各コンテンツがそれぞれに紐づくジャンルなどの正解ラベルに基づいて適切なカテゴリに分類されているかを評価することになる。しかし、その評価は 1 章で述べた分類学的カテゴリへの分類の良さを評価することとなり、本研

表 4.2 Yang らの手法による映画興味推移可視化の指標

クラスタ数	各クラスタ興味レベル数	D	C	U
1	8	26	67	41
2	4	5	62	57
3	4	7	96	89
4	4	15	112	97
5	5	58	168	110
6	5	75	180	105
7	5	52	228	176
8	3	31	146	115

究の目的である消費者の潜在意識に存在するアドホック・カテゴリに適切に分類できているかについての評価にはならない。次に、各カテゴリごとにコンテンツが適切な興味レベルに割り当てられているかどうかを判定する場合には、それぞれのコンテンツごとに興味レベルの正解データが必要になる。そのようなユーザの興味レベルを表すデータは本実験で用意したデータセットに存在せず、[Yang 14] の評価実験においても、レベル分けの評価については、患者の慢性腎臓病の進行具合のデータによる検証のみに留まっている。このことから、ユーザのコンテンツに対する興味レベルの正解データを用意することは困難であり、したがって、各コンテンツが適切な興味レベルに割り当てられているかどうかを直接評価することも困難であると考えられる。

本実験における提案手法では、各興味集合ベクトルの各素性を **Tf-Idf** 法によって **TFIDF** 値に変換した。Tf-Idf 法は、主に文章内でその特徴的な単語を発見するために用いられる手法であるが、興味集合ベクトルについても、その特徴的な素性（コンテンツ）に重み付けをすることで、ソフトクラスタリングの質を高めることを目的としている。登場頻度である **TF**(Term Frequency) を式 4.4、全ユーザ数から **IDF**(Inverse Document Frequency) を式 4.5、**TFIDF** を式 4.6 によって計算する。 $tf_{i,j}$ の分母 $n_{i,j}$ は興味集合ベクトル u_j におけるコンテンツ c_i の登場回数であり、興味集合ベクトルが **One-Hot** ベクトルであるため $n_{ij} = \{0, 1\}$ である。また、分子 $\sum_k n_{k,j}$ は u_j の全ての素性の和となる。 idf_i における分子 $|U|$ は全ユーザ数、分母は興味

表 4.3 提案手法による映画興味推移可視化の指標

クラスタ数	クラスタ別興味レベル数	D	C	U
2	(8, 8)	15	115	100
3	(8, 8, 8)	24	155	131
4	(6, 8, 8, 8)	29	178	149
5	(6, 8, 8, 8, 8)	31	194	163
6	(4, 6, 8, 8, 8, 8)	37	207	170
7	(4, 4, 5, 8, 8, 8, 8)	29	213	184
8	(6, 6, 8, 8, 8, 8, 8, 8)	53	254	201

集合ベクトルにコンテンツ c_i を含むユーザ数を示す。

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (4.4)$$

$$idf_i = \log \frac{|U|}{|\{u : u \ni c_i\}|} \quad (4.5)$$

$$tfidf_{i,j} = tf_{i,j} \cdot idf_i \quad (4.6)$$

さらに、興味集合ベクトルをそのままクラスタリングに用いると、計算量が膨大になることで処理に多大な時間を要してしまう [Sen 09] ため、Tf-Idf 法の処理を実施した各興味集合ベクトルを、主成分分析 (Principal Component Analysis、PCA) によって $k = 20$ 次元に削減した。

Fuzzy c -means クラスタリングにおけるパラメータ m には、Bin らの手法 [Bin 12] で用いられる $m = 2$ を採用し、セントロイドの変化合計距離が 0.005 を下回ったときか、ステップを 1000 回繰り返したときに処理を終了した。

クラスタへの所属の判定は、クラスタ総数が n のときに、各ユーザはそれぞれ所属度が $1/n$ 以上であるクラスタに所属するものとした。興味レベル判定については、各ユーザがコストを最小とするグラフ上の経路が変化しなくなったときか、ステップを 50 回繰り返したときに処理を終了した。また提案手法は、クラスタ数 $C = (2, 3, 4, 5, 6, 7, 8)$ と最大興味レベル上限数 $K = (2, 3, 4, 5, 6, 7, 8)$ における組み合わせ計 49 通り、比較対象とする Yang らの手法はクラスタ数 $C = (1, 2, 3, 4, 5, 6, 7, 8)$ と最大興味レベル上限数 $K = (2, 3, 4, 5, 6, 7, 8)$ における組み合わせ計 56 通りを試行した。

実験結果を、映画については表 4.2 及び表 4.3 に、マンガについては表 4.4 及び表 4.5 に記す。映画についての詳細な結果は付録 A に、漫画については図 4.6 に記す。また、各表で U

表 4.4 Yang らの手法によるマンガ興味推移可視化の指標

クラスタ数	各クラスタ興味レベル数	<i>D</i>	<i>C</i>	<i>U</i>
1	6	11	53	42
2	3	3	51	48
3	3	7	63	56
4	2	0	55	55
5	2	4	57	53
6	3	25	92	67
7	2	13	68	55
8	2	10	76	66

表 4.5 提案手法によるマンガ興味推移可視化の指標

クラスタ数	クラスタ別興味レベル数	<i>D</i>	<i>C</i>	<i>U</i>
2	(3, 8)	14	81	67
3	(2, 3, 8)	20	79	59
4	(2, 3, 3, 8)	18	84	66
5	(2, 2, 3, 3, 8)	20	84	64
6	(2, 2, 2, 3, 4, 8)	19	89	70
7	(2, 2, 2, 2, 3, 3, 8)	20	84	64
8	(2, 2, 2, 2, 3, 3, 3, 8)	20	84	64

が最も高い結果を太字で示してある。実験の結果、提案手法によって、Yang らの手法より高い *U* が得られる可視化が実現されていることが確認できる。映画についてはクラスタ数を増やすことで *U* が増加する傾向にあるが、マンガの場合は必ずしもそうではないことが分かる。このような結果となったことは、表 4.1 で示されるように、マンガのデータセットの対象コンテンツと平均興味行動回数が映画のデータセットのそれらと比べて少なかったことによって、クラスタ数の増加による *C* の増加の余地が小さかったことが原因であると考えられる。また、Yang らの手法による結果では、クラスタ数が増加してもエラー数やカバー数が減少する場合があるが、その理由として、手法を行った結果ユーザが 1 人も発生しないクラスタが発生したことや、コンテンツが 1 つも集計されない興味レベルが発生したことが原因にあげられる。

C Group 1	LV.1 NARUTO-ナルト 北斗の拳 ONE PIECE ポケットモンスター BLEACH D.Gray-man トリコ 銀魂 Axis.powers.ヘタリア 名探偵コナン	LV.2 進撃の巨人 黒子のバスケ 銀の匙 Silver Spoon ニセコイ 暗殺教室 トリコ IS (インフィニット・ストラトス) 東京喰種トーキョーグール はたらきマン 鬼灯の冷徹	Group 2	LV.1 こちら葛飾区亀有公園前派出所 名探偵コナン NARUTO-ナルト ポケットモンスター 北斗の拳 ハヤテのごとく! 美少女戦士セーラームーン あたしンち D.Gray-man 間文字D	LV.2 魔法少女まどか☆マギカ Axis.powers.ヘタリア Angel Beats! FAIRY TAIL ソウルイーター 日常 (漫画) トリコ ヴァンパイア騎士 銀魂 べるぜバブ								
	Group 3	LV.1 こちら葛飾区亀有公園前派出所 名探偵コナン NARUTO-ナルト ポケットモンスター 北斗の拳 美少女戦士セーラームーン ハヤテのごとく! D.Gray-man あたしンち ONE PIECE		LV.2 Axis.powers.ヘタリア 魔法少女まどか☆マギカ FAIRY TAIL トリコ 銀魂 ソウルイーター べるぜバブ ヴァンパイア騎士 日常 (漫画) ONE PIECE	Group 4	LV.1 To LOVEる-とらぶる- 神のみぞ知るセカイ FAIRY TAIL 君に届け ソウルイーター トリコ Axis.powers.ヘタリア セキレイ (漫画) ハヤテのごとく! 天体観戦サンレッド	LV.2 魔法少女まどか☆マギカ 進撃の巨人 僕は友達が少ない 黒子のバスケ 日常 (漫画) IS (インフィニット・ストラトス) Angel Beats! 半田兄弟 べるぜバブ ちはやふる	LV.3 東京喰種トーキョーグール マギ (漫画) 蒼き狼のアルペジオ 艦隊これくしょん-艦これ- 鬼灯の冷徹 テラフォーマーズ 東京レイヴンズ 魔法ペナル 七つの大罪 (漫画) のんのんびより					
Group 5	LV.1 僕は友達が少ない IS (インフィニット・ストラトス) 魔法少女まどか☆マギカ 神のみぞ知るセカイ 日常 (漫画) Angel Beats! 君に届け 進撃の巨人 黒子のバスケ To LOVEる-とらぶる-	LV.2 テラフォーマーズ 東京喰種トーキョーグール ノラガミ のんのんびより 動物のお医者さん 大奥 (漫画) アイアンヒーロー リアル (漫画) バガボンド 山田くんと人の魔女	LV.3 進撃の巨人 マギ (漫画) ダイヤのA FAIRY TAIL 聖☆おにいさん 銀の匙 Silver Spoon 史上最強の弟子ケンイチ ソウルイーター 半田兄弟 鬼灯の冷徹	LV.4 暗殺教室 ニセコイ 黒子のバスケ 青木椿雄の♀難 食戟のソーマ 艦隊これくしょん-艦これ- 銀魂 七つの大罪 (漫画) ワールドトリガー ジョジョの奇妙な冒険	Group 6	LV.1 僕は友達が少ない 神のみぞ知るセカイ 録音探偵録カエジ IS (インフィニット・ストラトス) 日常 (漫画) 魔法少女まどか☆マギカ アカギ ~闇に降り立った天才~ べるぜバブ Angel Beats! おおきく振りかぶって	LV.2 暗殺教室 鬼灯の冷徹 ソウルイーター テラフォーマーズ ノラガミ 東京喰種トーキョーグール 山田くんと人の魔女 神の雫 薬師寺流子の怪奇事件簿 のんのんびより	LV.3 進撃の巨人 黒子のバスケ マギ (漫画) ニセコイ トリコ 半田兄弟 ジョジョの奇妙な冒険 聖☆おにいさん 艦隊これくしょん-艦これ- 間金ウジマくん	LV.4 IS (インフィニット・ストラトス) 食戟のソーマ 暗殺教室 七つの大罪 (漫画) はたらきマン 東京喰種トーキョーグール ワールドトリガー 魔法少女まどか☆マギカ 君に届け 君のいる町	LV.5 銀の匙 Silver Spoon 艦隊これくしょん-艦これ- 僕は友達が少ない セキレイ (漫画) のんのんびより 蒼き狼のアルペジオ モンスタースターの日常 トリコ 天体観戦サンレッド 半田兄弟	LV.6 ダイヤのA 新宿スワン WORLD (漫画) ファンタジスタ (漫画) PandoraHearts 龍狼伝 最強伝説 黒沢 あまつき 暗殺教室 Angel Beats!	LV.7 ダイヤのA センセイ君主 こちら葛飾区亀有公園前派出所 女神の鬼 電波教師 日々娘々 進撃の巨人 P2:K7 GIANT KILLING 北斗の拳 ベイビーステップ	LV.8 神のみぞ知るセカイ 東京喰種トーキョーグール テラフォーマーズ 月光条例 トリニティセブン 7人の魔書使い 境界のRHINE ムシブキオー 永遠の0 ねじまきガキョウ 蒼き狼のアルペジオ

図 4.6 提案手法による可視化結果の例 (マップ C)

表 4.6 各可視化手法による映画・マンガの興味推移予測実験精度

手法	映画	マンガ
[Yang 14]	0.024	0.069
提案手法	0.033	0.081

また、この可視化結果に示されたクラスターを定性的に分析することで、マーケティング施策を立案する際に活用できる洞察を得ることができる。たとえば、図 4.6 で可視化されているクラスターに注目すると、Group 1 は主に週刊少年ジャンプに連載されていた作品から、主に他の週刊少年漫画雑誌に連載されていた作品に興味を推移していると推察され、Group 2 は主にゴールデンタイムで放映されていたアニメから、主に深夜に放送されていたアニメに興味を推移していると考えられる。他にも、Group 6 では時期ごとに流行していた漫画が順番に並んでいるため、その時々に関連され注目されていた漫画を時系列順に追いつけているクラスターであると考えられる。このように、本手法による可視化結果をドメイン知識を活用しながら解釈することによって、人間の目視だけでは得られないような発見ができると考えられる。

4.3.3 ユーザの興味行動の予測による評価実験

前項ではコンテンツのカバー状況を評価するために独自の指標を導入して提案手法の評価を行ったが、それだけでは本提案手法の応用への有効性が不明確であり、他の手法と比べた際の比較が困難である。そのため、[Yang 14]でも行われているような、可視化結果を用いたユーザの興味行動の予測の精度による定量的評価も実施した。

具体的には、興味行動シーケンス長が3以上の各ユーザから最後の1つの興味行動を取り除いたものを入力データとし、それから興味推移を可視化した。その後、各ユーザのクラスタとレベルを推定し、そのレベルにおける可視化結果の上位10件を予測候補とし、取り除いた最後の1つのコンテンツがそれに含まれているか否かによって予測の精度を評価した。すなわち、最後の2つの興味行動間でレベルの変化がないことを仮定した上で次にユーザの興味行動をとる対象コンテンツの予測を行い、その精度を評価した。提案手法については、そのユーザが K 個のクラスタに属する場合、それぞれのクラスタの該当レベルの上位 $10/K$ 個（小数点以下切り捨て）を予測候補とした。さらに、候補としたコンテンツに重複があった場合にも、候補を追加はしていないため、提案手法によって提示される予測の候補はユーザ次第では10個より少なくなることがある。

以上の方法で予測を行った結果を表4.6に記す。予測精度そのものの値は小さいものの、[Yang 14]で行われた同様の実験結果より予想される結果からは大きく乖離しておらず、[Yang 14]の手法より提案手法のほうが予測精度の面でも高くなっていることが示された。ただし、[Yang 14]よりは良いものの予測精度そのものは小さいため、本提案手法は推薦システムとしては実用的であるとは言えない。推薦システムとしてより高い精度を出す研究については、次の第5章にて述べる。

4.4 まとめ

本研究では、Webサービスなどのユーザの行動が記録されたログデータを用いて、ユーザの興味対象がどう推移するかを表すカテゴリを可視化する手法を提案した。あるコンテンツに対し興味を示している証拠となる行動を興味行動と定義し、ログデータからユーザの興味行動の時系列を抽出、ユーザをソフトクラスタリング、そして各クラスタ内のユーザの興味推移状況をレベルで表すことで、クラスタごとに興味対象がどう推移するかを可視化した。検証実験の結果、本手法によって消費者の興味推移動向が従来手法よりも詳細に可視化でき、商品の推薦

にも活用できる可能性があることを示した。ただし、提案手法に用いた [Yang 14] 中の仮定である人間のコンテンツへの趣味嗜好は深化するが退化することはないという制約は、人間の直感に反し不自然なものであるとも考えられるので、その制約を廃した場合に可視化がよりよいものになるかは、今後の研究課題である。

第 5 章

行動時刻を考慮した条件付き変分オートエンコーダによる推薦システム

近年、深層学習に基づく推薦システムの研究では、変分オートエンコーダ (Variational Auto-encoder; VAE) に基づく手法が注目されている。VAE ベースのレコメンデーションシステムは、ユーザの行動データと VAE エンコーダの出力との差分に基づいて商品を推薦する手法である。VAE ベースの推薦システムは、従来の深層学習を用いた推薦システムと比較して、より小さなデータセットでも精度が向上し、ロバスト性が向上することが報告されている。また、深層学習に基づく推薦システムに関する研究の中には、ユーザが購入した商品の情報に加えて、ユーザの性別や年齢などの属性情報や、ユーザが使用している端末の情報を利用することが提案されているものもある。これらの付加情報は、単に入力時に付加するのではなく、どのように入力するかについて様々な研究が行われている。このような追加情報の一つとして、商品の購入やレビューなどの行動の時期を考慮することで、レコメンドの精度を向上させることができる。これにより、行動に反映されていない、行動の順序や時間経過によるユーザの好みの変化など、ユーザの内部状態の遷移を考慮することができ、推薦の精度を向上させることができる。したがって、これらの時間情報を推薦時に利用できれば、VAE に基づく推薦システムの精度を向上させることが可能である。本論文では、入力情報に各アクションの時間情報を付加して推薦精度を向上させる VAE ベースの推薦システム (Time-Sequential VAE; TSVAE) を提案する。

5.1 はじめに

推薦において行動時刻は精度向上に寄与する重要な情報であるので、これまでに商品の購入や商品のレビューなどの行動をした時刻の情報を推薦に用いる手法が提案されている [Zhu 17, Pavlovski 20]。一般に、推薦システムの研究ではユーザの購入した商品の情報にユーザの性別や年代といった属性情報や使用している端末の情報などを条件として用いることで推薦精度を高める文脈考慮型推薦システム (Context-Aware Recommender Systems) がこれまでに数多く提案されている [Adomavicius 11]。さらに、これらの情報は入力方法によって大きく推薦精度が変化するため、学習のどのタイミングでどのように付与すればよいかについても研究が行われている [Beutel 18]。

本研究では、Variational Auto-encoder (VAE)[Kingma 13] ベースの推薦システムに行動の時刻情報を付与する方法を提案する。VAE は深層学習による生成モデルの一種であり、エンコーダによって入力データから潜在変数の分布の平均や分散を学習し、デコーダによってこの分布からサンプリングした潜在変数から再び入力データと似たデータを復元する手法である。近年では深層学習をベースとした推薦システムの研究においても VAE をベースとした手法が注目されつつある。VAE を用いた推薦システムは、ユーザの商品に関する行動履歴データをエンコーダに入力し、それとデコーダからの出力との差分をとることで、出力には復元されているが入力自体には含まれていない商品を推薦するという方法であり、協調フィルタリングから発展した推薦手法として研究が進められてきた。VAE ベースの推薦システムは、従来の深層学習を用いた推薦システムに比べ、より小さいデータセットに対する精度が向上し、頑健性も高まることが報告されている [Liang 18]。しかし、[Liang 18] など既存の VAE ベースの推薦手法では、好んだ商品の情報が順序や時刻を無視した Bag-of-Words 形式で与えられるため、その時刻や時間間隔は無視されてしまっている。そこで、本研究では VAE ベースの推薦システムの精度を時刻情報を考慮することで改善することを目的とする。

時刻情報を VAE に織り込む既存研究として、RNN によるエンコード層を通すことによって商品を好んだ順序の情報を織り込む手法 [Sachdeva 19] が挙げられる。しかし、[Sachdeva 19] では、その商品を好んだ時刻情報は考慮されていない。推薦時に絶対時刻の情報が考慮されていないと、たとえば、映画や音楽といったコンテンツを推薦対象とした際に、一時的なブームによって特定の商品が多くのユーザによって好まれる事態が生じた場合、そのユーザの嗜好による選好であると判断され、推薦結果に悪影響を及ぼしてしまう可能性がある。他にも、複数

行動の時刻の差である時間間隔情報が考慮されていないと、特に酒類やコンテンツのような時間の経過とともに成熟度が変わることによって趣味嗜好の対象が変化すると考えやすい種類の商品カテゴリを推薦する場合 [McAuley 13] に、観測された行動に反映されていないユーザの内部状態の変遷を追うことができないため、その影響が適切に推薦に反映されない可能性がある。

以上の問題を踏まえ、本研究では入力とする情報に各行動の時刻情報を付与することで推薦精度を高める VAE ベースの推薦システムを検討し、最終的に (Time-Sequential VAE; TSVAE) を提案し、その精度を確認する。また、時刻情報をどのように付与していけば精度の向上に寄与するかを検証するために、絶対時刻情報の有無や時間間隔情報を織り込むエンコーダの違いによる複数のモデルを用いた実験を行った。本研究では、提案手法の有効性を確認するために、以下の検証実験を行った。

- [Sachdeva 19] では系列情報を VAE に入力する前のエンコーダとして RNN を用いるが、そこに時間情報を考慮する RNN である Time-LSTM[Zhu 17] を採用し、行動間の時間情報を RNN エンコード前の表現に織り込む手法を検討した。検証実験の結果、データセットの時間情報がより正確に収集できているほど推薦精度が向上することが確認された。
- VAE を用いた推薦システムを拡張した CVAE[Sohn 15] による推薦システムによる手法を導入し、その入力条件に最新の行動の時刻情報を用いる手法を検証した。検証実験の結果、単純に時刻情報を条件とした CVAE を用いることによって、単純な VAE による推薦システムに比べて精度が向上することが確認された。
- 以上で検討した手法でよりよい結果を残した構造を組み合わせ、精度を向上する手法 (TSVAE) を提案する。この結果、上記の既存手法に比べてより良い精度を記録することが確認された。
- VAE エンコーダの出力に、商品情報だけではなく時刻情報を同時に再構成する手法を検証した。実験の結果、この方法では推薦精度が上がらないことが確認された。

5.2 前提知識

本章では、本研究で提案する手法の一部として用いられている手法について解説する。5.3 項では、これらを前提とした手法を提案することになる。また、5.4 項では、提案手法の優位

性を確認するために本章で紹介した手法との比較実験を行う。

5.2.1 Sequential Variational Auto-encoder

変分オートエンコーダ (Variational Auto-encoder; VAE) とは、通常のオートエンコーダと同様のエンコーダとデコーダの構造を持つ深層学習を用いた生成モデルである [Kingma 13]。通常のオートエンコーダは、エンコーダにより入力 \mathbf{x} から潜在変数 \mathbf{z} に変換し、デコーダによりそれから元データ $\hat{\mathbf{x}}$ を復元、それらの再構成誤差が小さくなるように学習をすすめる。しかし、VAE では、その潜在変数が確率分布 $\mathcal{N}(0, 1)$ に分布に従うことを仮定し、潜在変数の平均 $\boldsymbol{\mu}$ および分散 $\boldsymbol{\sigma}$ を入力データより学習する。そして、VAE は生成モデルであるのでデータの分布 $p(\mathbf{x})$ を推定するよう学習をすすめるが、その尤度を直接最大化することはできないため、以下のような式変形を行う。

$$\begin{aligned} \log p(\mathbf{x}) &= \int q(\mathbf{z} | \mathbf{x}) \log p(\mathbf{x}) d\mathbf{z} \\ &= \int q(\mathbf{z} | \mathbf{x}) \log \frac{p(\mathbf{x} | \mathbf{z})p(\mathbf{z})}{p(\mathbf{z} | \mathbf{x})} d\mathbf{z} \\ &= \int q(\mathbf{z} | \mathbf{x}) \log p(\mathbf{x} | \mathbf{z}) d\mathbf{z} \\ &\quad + \int q(\mathbf{z} | \mathbf{x}) \log \frac{q(\mathbf{z} | \mathbf{x})}{p(\mathbf{z} | \mathbf{x})} d\mathbf{z} \\ &\quad - \int q(\mathbf{z} | \mathbf{x}) \log \frac{q(\mathbf{z} | \mathbf{x})}{p(\mathbf{z})} d\mathbf{z} \\ &= E_{\mathbf{z} \sim q}[\log p(\mathbf{x} | \mathbf{z})] + D_{KL}(q(\mathbf{z} | \mathbf{x}) \| p(\mathbf{z} | \mathbf{x})) \\ &\quad - D_{KL}(q(\mathbf{z} | \mathbf{x}) \| p(\mathbf{z})) \end{aligned}$$

最終的にこの式を整理することによって、以下の式を得られる。左辺に表されている尤度の最大化問題は、右辺の変分下界 (Evidence Lower Bound; ELBO) の最大化問題に帰着できる。VAE の学習は、この変分下界の最適化問題として実行される。

$$\log p(\mathbf{x}) - D_{KL}[q(\mathbf{z} | \mathbf{x}) \| p(\mathbf{z} | \mathbf{x})] = E_{\mathbf{z} \sim q}[\log p(\mathbf{x} | \mathbf{z})] - D_{KL}[q(\mathbf{z} | \mathbf{x}) \| p(\mathbf{z})]$$

その入力データが系列データとなった Sequential Variational Auto-encoder (SVAE)[Sachdeva 19] も、VAE と同様にデコーダによって復元された系列データとの復元誤差が小さくなるようにエンコーダとデコーダの学習をすすめる。ただし、系列はエンコーダに一度に直接入力するのではなく、系列の順で逐次的に RNN によるエンコーダを通し、その出力を VAE エンコーダ

に入力していく。その後、VAEと同様にVAEデコーダから逐次的に出力していく。そのため、系列データの順序に関する情報がRNNエンコーダを通じた際に織り込まれることによって、VAEによって復元された系列にも反映されると考えられる。ただし、通常のVAEとさらに異なる点として、 μ と σ で表される潜在変数が正規分布ではなく、多項（マルチヌーイ）分布に従うと仮定することが挙げられる。また、[Sachdeva 19]ではRNNエンコーダにGated Recurrent Unit(GRU)を採用している。

$$\begin{aligned}\mu(t), \sigma(t) &= \varphi(\mathbf{h}_t) \\ \mathbf{h}_t &= RNN(\mathbf{h}_{t-1}, \mathbf{x}_{(t-1)})\end{aligned}$$

5.2.2 Conditional Variational Auto-encoder

通常のVAEでは入出力の復元対象とする x は $x \sim p_\theta(z)$ として表されるが、それを条件 y を付けた条件付き尤度 $p(x|y)$ に拡張したものを考える。それがConditional Variational Auto-encoder(CVAE)[Sohn 15]と呼ばれる。具体的には、条件を y として以下の式で表される。

$$\begin{aligned}\log p(y|x) &\geq \int \log p(y|x,z)p(z|x)dz \\ &= -D_{KL}(q(z|x,y)|p(z|x)) + E_q[\log p(y|x,z)] \\ &\equiv \mathcal{L}_{CVAE}(x,y)\end{aligned}$$

学習の対象が条件付き尤度に拡張されても、エンコーダやデコーダの形状は入力に条件を付与する変更のみでよく、その学習もVAEと同様に変分下界の最大化によって進めることができる。

5.2.3 Time-LSTM

通常の再帰的ニューラルネットワーク(Recurrent Neural Network; RNN)は系列データをその順番に入力しているが、その系列データが順序のみを考慮すればよい自然言語文ではなく、行動履歴のようなそれぞれの要素に時刻が定義でき、それらに異なる時間間隔があるようなデータの場合、その情報は無視されてしまう。そこで、その時間間隔を同時に入力データとして考慮するゲート付きRNNであるTime-LSTM[Zhu 17]が提案された。[Zhu 17]の論文ではTime-LSTMのいくつかのバリエーションが提案されているが、本研究では[Zhu 17]にて最も

良い精度を示していた Time-LSTM(3) を用いる。

これより、本提案手法で使用する Time-LSTM(3) の概要を紹介する。まず、重みを W 、バイアス項を b 、セルを c 、セルの重みを w 、入力ゲートを i 、出力ゲートを o と表記すると、Time-LSTM(3) は以下の式で表される。

$$i_m = \sigma_i (x_m W_{xi} + h_{m-1} W_{hi} + w_{ci} \odot c_{m-1} + b_i) \quad (5.1)$$

$$\begin{aligned} T1_m &= \sigma_1 (x_m W_{x1} + \sigma_{\Delta t} (\Delta t_m W_{t1}) + b_1) \\ \text{s.t. } W_{t1} &\leq 0 \end{aligned} \quad (5.2)$$

$$T2_m = \sigma_2 (x_m W_{x2} + \sigma_{\Delta t} (\Delta t_m W_{t2}) + b_2) \quad (5.3)$$

$$o_m = \sigma_o (x_m W_{xo} + \Delta t_m W_{to} + h_{m-1} W_{ho} + w_{co} \odot \tilde{c}_m + b_o) \quad (5.4)$$

$$\tilde{c}_m = (1 - i_m \odot T1_m) \odot c_{m-1} + i_m \odot T1_m \odot \sigma_c (x_m W_{xc} + h_{m-1} W_{hc} + b_c) \quad (5.5)$$

$$c_m = (1 - i_m) \odot c_{m-1} + i_m \odot T2_m \odot \sigma_c (x_m W_{xc} + h_{m-1} W_{hc} + b_c) \quad (5.6)$$

$$h_m = o_m \odot \sigma_h(\tilde{c}_m) \quad (5.7)$$

ただし、 σ は活性化関数を表し、 $\sigma_i, \sigma_{\Delta t}, \sigma_o$ はシグモイド関数、 σ_c, σ_h は \tanh 関数である。また、 \odot は要素積（アダマール積）を表す。

Time-LSTM では、一般的な LSTM で入力される x_m, h_m の他に、ユーザの各行動ごとにそのユーザの直前のアクション x_{m-1} からの時間間隔 Δt を導入する。そして、 Δt をより効果的に出力に反映させるために、Time-LSTM(3) では Δt の直近の推薦への影響を制御する時間ゲート $T1_m$ と、 Δt の影響を長期に渡って保持するための時間ゲート $T2_m$ を導入する。ユーザの直近（短期）の興味が x から入力されるのに対し、 Δt が入出力ゲートに作用させるとともにそれぞれのセル c_m や \tilde{c}_m にも作用されるため、 Δt の影響が長期間にわたって保持できるようになっている。そして、一般的な LSTM において忘却ゲートとして用いられる機構は、Time-LSTM(3) では、GRU と類似した入力ゲートを用いた機構である $1 - i_m$ に置き換えられている。これを模式図で表すと、図 5.1 となる。

5.3 提案手法

前章の内容を踏まえて、本章では時刻に関する情報を推薦時に織り込む VAE ベースの推薦システム (Time-Sequantional Variational Auto-Encoder; TSVAE) を提案する。モデル全体の概要を図 5.2 に示す。具体的には、TSVAE は SVAE[Sachdeva 19] をベースとし、時間情報 Δt を織り込むために RNN エンコーダに Time-LSTM(3) を、時刻情報 t を織り込むために、VAE

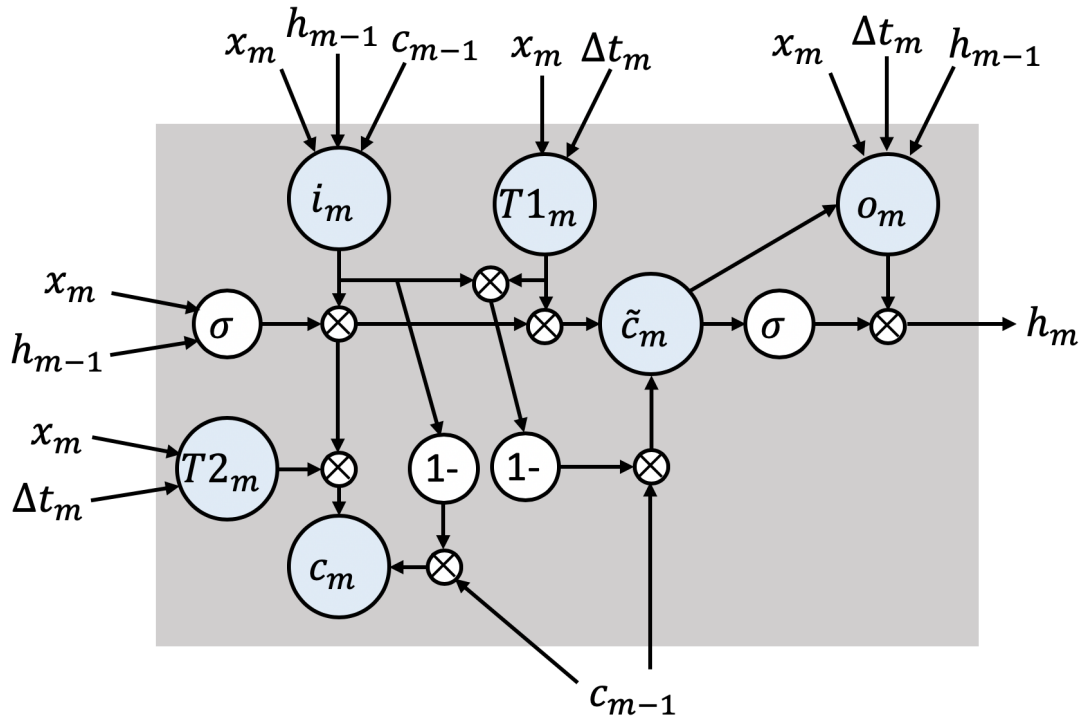


図 5.1 Time-LSTM 3 の模式図

エンコーダおよびデコーダに Conditional VAE を採用したモデルである。実際の推薦タスクを行う際も [Sachdeva 19] と同様に、VAE エンコーダの出力では評価されているが入力（実データ）では評価されていない商品をランキング化し、それらを上位から推薦する。

まず、入力データの前処理としての変換 (Embedding) を行う。ID 形式で入力されている商品の系列を、ベクトル表現に変換する。また、時間情報 Δt には、[Zhu 17] でも提案されている通り、アクション x の間隔時間差を直接用いるのではなく、より推薦精度を高めることができるよう以下の式で表される時刻の差に対数をとったものに変換する。ただし、 $\Delta t_1 = 0$ とする。

$$\Delta t_n = \log(t_n - t_{n-1}) \quad (5.8)$$

5.3.1 Time-LSTM による Δt のエンコーディング

次に、変換された系列データを SVAE と同じ要領で RNN エンコーダに入力する。SVAE では RNN エンコーダに GRU を採用していたが、本提案手法ではそれを Time-LSTM(3) に置き換える。そうすることで、RNN エンコーダを通じた際に順序情報のみならず、 \mathbf{x} 中に含まれる時間情報 Δt の情報もエンコードされ、推薦時に考慮されるようになると考えられる。具体的には、以下の式で表される。

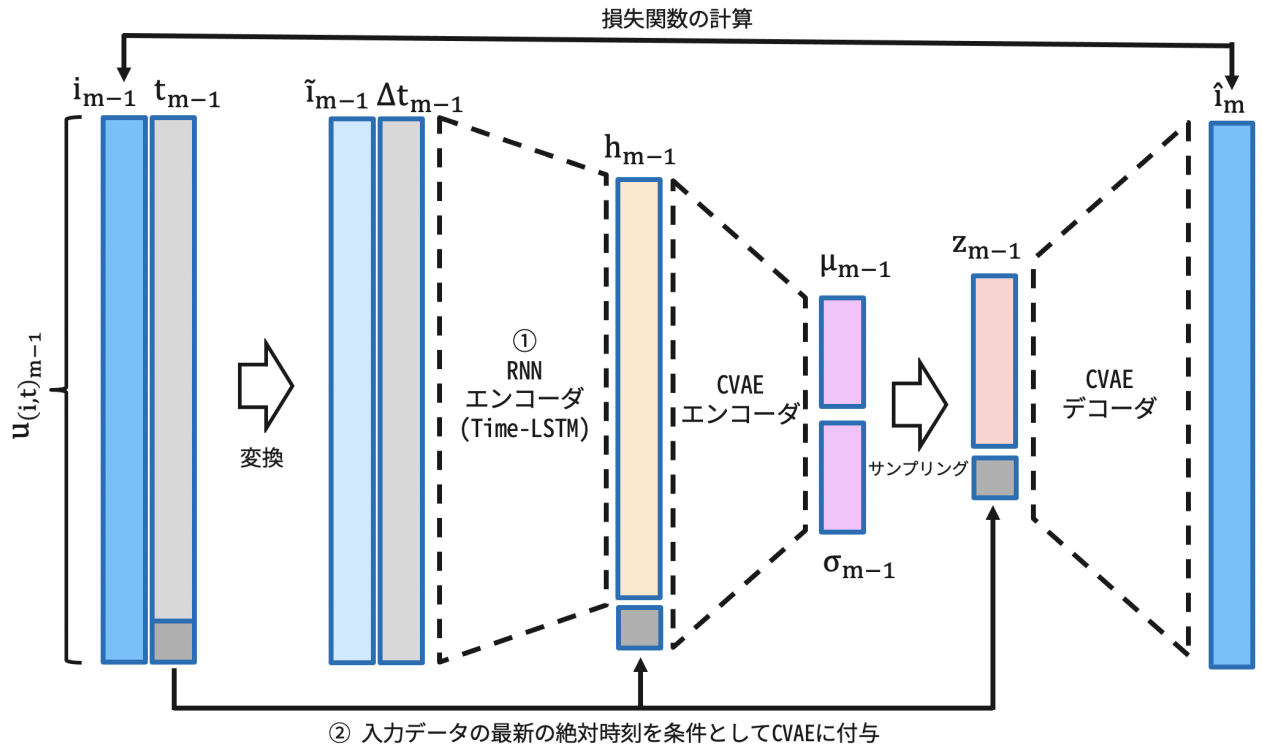


図 5.2 提案手法の模式図

$$\mathbf{h}_{m-1} = \text{Time-LSTM}_3(\mathbf{h}_{m-1}, \mathbf{x}_{(m-1)}, \Delta t_{m-1}) \quad (5.9)$$

5.3.2 t を条件とした CVAE による入力データの復元

SVAE では VAE エンコーダおよびデコーダが通常の VAE であったが、これを時刻を条件とした CVAE に置き換える。RNN エンコーダを通じた入力 \mathbf{h}_n と VAE エンコーダ φ を通じて得られた潜在変数 \mathbf{z} に VAE にユーザの行動履歴の最新の行動時刻 t_n を条件として付与し、それを VAE デコーダ ϕ に入力することによって最終的にユーザの行動履歴を復元する。そうすることで、推薦システム全体で見た際に、その入力の絶対時刻の情報が推薦時に反映されるようになると考えられる。具体的には、以下の式で表される。ただし、*Multi* は多項分布である。

$$\boldsymbol{\mu}_{m-1}, \boldsymbol{\sigma}_{m-1} = \varphi(\mathbf{h}_{m-1}, t_{m-1}) \quad (5.10)$$

$$\mathbf{z}_{m-1} \sim \text{Multi}(\boldsymbol{\mu}_{m-1}, \boldsymbol{\sigma}_{m-1}) \quad (5.11)$$

$$\hat{\mathbf{i}}_m = \phi(\mathbf{z}_{m-1}, t_{m-1}) \quad (5.12)$$

表 5.1 データセットの概要

データセット	ユーザ数	作品数	平均データ長
MovieLens	2597	705	39.86
LastFM	149	3898	1224

以上の方法で t や Δt に関する情報を付与することによって、VAE ベースの推薦システムにユーザの各行動の時間間隔やその絶対時刻といった情報が反映され、従来の SVAE より推薦精度の向上につながると考えられる。

5.4 実験

5.4.1 実験設定

本章では、ユーザの行動時刻情報や提案手法によるその付与の方法の有効性を検証するために、実際行動履歴データを用いた推薦タスクに、RNN エンコーダに GRU ([Sachdeva 19] の元々の設定)、LSTM、LatentCross、Time-LSTM(3) 三種類を採用した推薦システム、VAE エンコーダに通常の VAE および CVAE を用いた推薦システムを用意し、それらを用いて実際の推薦タスクを実行した。また、提案手法である TSVAE については、VAE デコーダの出力に時刻の系列 t_n も復元するように誤差の計算方法を変更し、それらも正しく再構成されるようにしたモデルも作成し、精度を比較検証した。どのモデル、データセットも共通して、それぞれ 30 epoch ずつ学習させた。最適化アルゴリズムには Adam を採用し、VAE エンコーダ、デコーダはそれぞれ三層で、25% の DropOut を実行した。

5.4.2 データセット

本検証実験で用いるデータセットには、映画のレビューサイトのデータセットである MovieLens-1M^{*1}、およびインターネットラジオをベースとした SNS のデータセットである Last.fm の Million Song Dataset^{*2}を用いた。これらはどちらも在庫の概念がないコンテンツを対象としているデータセットである。データセットの概要を表 5.1 に、データセットごとの $\Delta t = 0$ の分布を図 5.3 および図 5.4 に記す。対象とする作品数の多さや、使用できるデータ数

^{*1} <https://grouplens.org/datasets/movielens/1m/>

^{*2} <http://millionsongdataset.com/lastfm/>

や長さの少なさから、Last.fm の楽曲推薦タスクがのほうがより難易度が高いと言える。

MovieLens

このデータセットの出典は第 4 章で用いたものと同じである。ユーザが 5 段階評価で 5 を付けた作品を本研究の対象とする行動であるとみなし、そのような行動を起こす作品を推薦するタスクとした。総評価数が 50 以上のユーザのみを対象とし、それらユーザから 100 人以上に評価された作品のみを推薦の対象とした。以上のフィルタリングの結果、各推薦モデルは 705 本の作品から各ユーザが次にどの作品を評価するかを予測して推薦することになった。訓練データセットは 2197 件、バリデーションおよびテストデータセットはそれぞれ 200 件である。また、図 5.3 から分かるように、データセット内の多くのレコードで $\Delta t = 0$ となっており、 Δt の分布がべき乗則に従っているようにも見えることが分かる。これは MovieLens データセットでの t 自体は `unixtime` (秒単位) で記録されているものであり、同時に多数の作品を評価をした時刻が記録されている。このことから、これはユーザが実際に作品に評価をしたタイミングではなく、その評価情報が収集された時刻であると考えられる。そのため、同一ユーザの履歴の中で $\Delta t = 0$ で記録されている映画の順序については、暫定的に各映画に付与されている ID 順にした。

Last.fm

このサービス上で実際に楽曲を聴くことを本研究の対象とする行動であるとみなし、そのユーザが次に聴く楽曲を推薦するタスクとした。実際に視聴した履歴のデータセットであるため、先述した MovieLens より第 3 章で用いた小売データセットに近い、実際の行動実績を表すデータセットであると言える。総評価数が 5 以上 2000 以下のユーザのみを対象とし、それらユーザから 5 人以上に評価された作品のみを推薦の対象とした。以上のフィルタリングの結果、各推薦モデルは 3898 本の作品から各ユーザが次にどの曲を聴くかを予測して推薦することになった。訓練データセットは 89 件、バリデーションおよびテストデータセットはそれぞれ 30 件である。このデータセットでは MovieLens のようなデータのタイミングによって $\Delta t = 0$ となったデータは存在せず、MovieLens とは異なる分布を構成している。

5.4.3 評価指標

評価実験において、各推薦モデルはユーザごとに作品を上位 $n = 10$ か $n = 100$ 本推薦し、それらに対して精度を評価する。精度の評価には、以下に挙げる DCG、Precision@ n 、Recall@ n

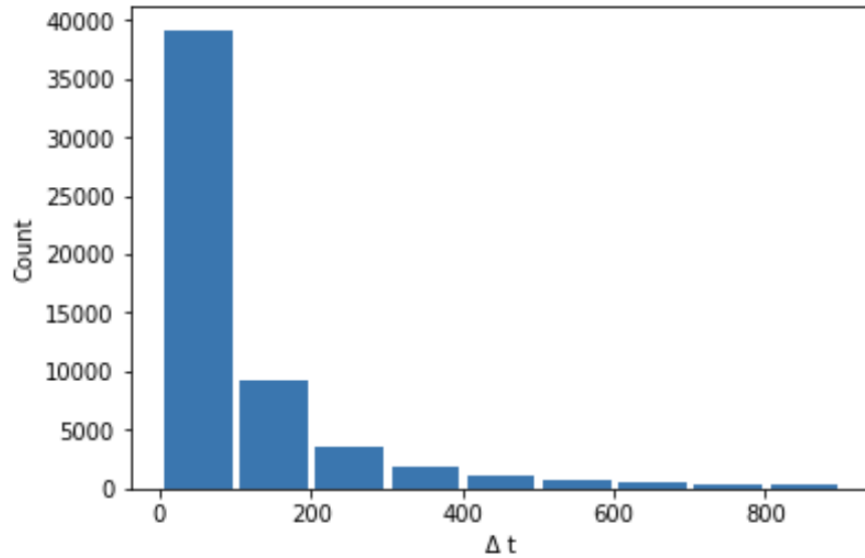


図 5.3 MovieLens データセット内での Δt の分布

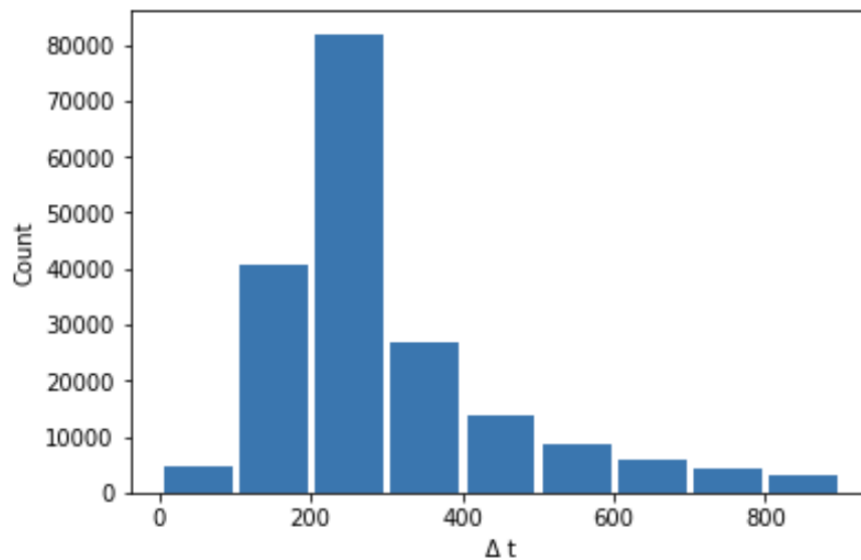


図 5.4 Last.fm データセット内での Δt の分布

を採用した。

Discounted Cumulative Gain (DCG)

Discounted Cumulative Gain (DCG) とは、推薦システムの提示した順位が実際の順位とどれだけ近いかを示す指標である。 r_i は i 番目に推薦した作品を評価したか否かを 0、1 で表す。DCG は大きい値をとるほど性能が良いことを示している。

表 5.2 MovieLens 実験結果

RNN エンコーダ	DCG@10	Recall@10	Precision@10	DCG@100	Recall@100	Precision@100
GRU	10.09	9.59	7.95	26.38	51.89	4.49
LSTM	10.70	10.59	8.85	25.96	51.26	4.35
Time-LSTM 3	10.98	10.44	8.50	26.17	50.86	4.31
CVAE	10.01	9.59	7.95	25.93	51.34	4.39
Latent Cross	8.50	8.49	6.75	24.30	50.60	4.28
TSVAE	11.31	11.19	8.95	29.24	52.29	4.90

表 5.3 Last.fm 実験結果

RNN エンコーダ	DCG@10	Recall@10	Precision@10	DCG@100	Recall@100	Precision@100
GRU	2.38	1.07	2.33	3.12	4.27	1.10
LSTM	1.93	0.62	2.00	3.20	4.91	1.67
Time-LSTM 3	1.89	0.66	1.33	2.29	2.50	0.76
CVAE	1.31	0.30	1.40	1.32	2.46	0.70
Latent Cross	4.46	1.62	3.00	3.72	4.85	1.07
TSVAE	4.31	2.30	3.40	4.24	5.29	1.90

$$DCG@n = \sum_{i=1}^n \frac{r_i}{\log_2(i+1)} \quad (5.13)$$

Precision、Recall、F-Value

推薦した作品がどれだけ実際にユーザが評価したかを表す Precision (適合率) と、ユーザが推薦リスト中の作品を実際にユーザが評価した割合を表す Recall (再現率) を評価指標として用いる。言い換えるなら、Precision は正解率であり、Recall はどれだけ作品をどれだけカバーできたかを表す指標である。|R| は推薦したアイテム数を表す。どちらもより大きい値をとるほど性能が良いことを示している。

$$Hits = \sum_i r_i \quad (5.14)$$

$$Precision@n = \frac{Hits@n}{n} \quad (5.15)$$

$$Recall@n = \frac{Hits@n}{|R|} \quad (5.16)$$

表 5.4 TSVAE 時刻情報の再構成の有無による比較実験 (Last.fm)

再構成対象	DCG@10	Recall@10	Precision@10	DCG@100	Recall@100	Precision@100
x	4.31	2.30	3.40	4.24	5.29	1.90
x, t	2.21	1.62	2.40	3.64	3.36	0.70

5.4.4 実験結果

実験結果を、MovieLens については表 5.2、Last.fm については表 5.3 に記す。既存手法はデータセットが異なると精度の良し悪しに違いが現れているが、提案手法である TSVAE については、Last.fm における DCG@10 のみを除き、他のいずれの既存手法に比べても今回設定した評価指標でより高い精度を示す結果が出たことが確認された。

比較対象とした手法について注目すると、Time-LSTM(3) は、MovieLens データセットでは精度が高いが、Last.fm データセットについては、 Δt をエンコードしない GRU や LSTM に比べて精度が低い結果となっている。このことから、元々の Time-LSTM(3) では、 Δt の情報が精度に活かしていない可能性がある。その原因としては、実験に使用できたデータセット内のユーザの総数が少なかったことが考えられる。提案手法である TSVAE では小さいデータセットに対しても精度を出せる VAE による推薦を取り入れているため、良い精度が出せたと考えられる。

また、Latent Cross[Beutel 18] は、MovieLens データセットでは他の手法と比べて精度が同程度か低い結果となっているが、Last.fm データセットでは、TSVAE に近い高い精度を示している。これは、Latent Cross が Δt の情報が正確にデータに反映されている場合には有効であるが、そうでない場合には精度を下げてしまうことを示唆していると言える。それに対して TSVAE では、どちらのデータセットでも高い精度を示している。

次に、TSVAE を用いた推薦において、CVAE デコーダで再構成する対象として商品 i_n の系列だけではなくその時刻の系列 t_n も対象とし、その再構成誤差の損失関数が小さくなるようモデルを学習させた。そして、Last.fm データセットにおいて他のモデルと同様の推薦実験を行った。その実験結果を表 5.4 に記す。結果から見て分かる通り、CVAE によって時刻を再構成しないモデルのほうが精度が高くなることが確認された。再構成誤差を小さくする対象を商品のみ限定したほうが、最終的に商品の再構成の精度が高まることが示唆される。

5.5 まとめ

本研究では、入力とする情報に各行動の時刻情報を付与することで推薦精度を高める VAE ベースの推薦システムを検討し、最終的に TSVAE を提案し、その精度を確認した。商品系列情報を VAE に入力する前のエンコーダとして Time-LSTM[Zhu 17] を採用し、VAE を用いた推薦システムを拡張し時刻情報を条件とした CVAE を導入することによって、単純な VAE による推薦システムや条件付き推薦システムに比べて精度が向上することが確認された。また、VAE エンコーダの出力に、商品情報だけではなく時刻情報を同時に再構成する手法を検証したが、実験の結果、この方法では推薦精度が上がらないことが確認された。ただし、本研究の提案システムはベースとした手法に比べてパラメータが増加することで計算量が大きくなり、より大規模なデータセットに対して適用しづらくなっている。これは、[Liang 18] が主張する、より小さいデータセットでの推薦精度の確保というメリットを相殺していると言える。このため、今後は計算量を削減しつつ、TSVAE と同様の推薦精度を示せるよう、特に RNN エンコーダ部分についての改善を検討したい。

第 6 章

考察

本章では、第 3 章、第 4 章、および第 5 章の成果を整理した上で議論し、本研究の全体の貢献をまとめる。そして、本研究の成果の限界について論じた上で、今後の課題について議論する。最後に、本研究の成果が本研究の対象とするマーケティング以外の応用可能性についても議論する。

6.1 各論の考察

6.1.1 複数周期に着目した畳み込みネットワークによる小売時系列データの予測

3 章では、複数周期に関するドメイン知識を踏まえ、CrCNN 及び MWCNN を用いた予測手法および小売時系列予測に特化した予測手法である CPNet を提案し、他の時系列予測手法と比較することによって、本提案手法の優位性を確認することができた。古典的な時系列予測手法を用いる場合は、予測対象とする時系列データの特徴を捉えた上でそのモデリングを行うが、それを行うためには、専門知識や試行錯誤のための工数が必要になる。本手法は小売時系列の特徴を前提として設計された手法であり、予測に際し深層学習を検討する際にベースラインとして比較的容易に採用できる手法であると言える。今回の実験では、予測対象が複数の周期性を持つことを考慮している MWCNN を用いたモデルのほうがより精度の高い結果となったが、データセット内の商材のバリエーションが少ない場合や、学習に用いることができるデータセットが小さくなってしまいうような場合には、MWCNN の代わりに CrCNN による畳み込み周期を固定したモデルを用いるほうがより高い予測性能を示す可能性があるとも考えられる。

6.1.2 Web 上でのユーザ行動履歴データを用いた消費者の興味推移の可視化

4章では、行動ログデータからユーザの興味の推移動向やその段階ごとに好むコンテンツを可視化することができ、さらに、アンケート評価によって他者の目からその質が担保されていることを確認できた。比較対象とした [Yang 14] も一種類の興味推移という仮定と制約をかけてタスクを実行していたが、[Yang 14] はビールやワインといった単一カテゴリの推薦タスクに向けて設計されたものであった。今回の予測対象とした映画などのコンテンツは、複数種類の興味が存在するとみなすことがふさわしく、その結果、[Yang 14] に比べてより質の高い可視化を実現することができた。この結果は、汎用的なモデルに対して人間の帰納によって得られる知見による制約をかけることが、より人間にとって納得がいくアウトプットを生み出すことができる可能性を示唆している。一方で、提案手法にも用いた [Yang 14] の仮定である人間のコンテンツへの趣味嗜好は深化する一方であり退化や循環をすることはないという制約はそのような知見に反しているため、その制約を廃した場合に可視化がよりよいものになるかや、そうした際にモデル自体が学習できるかどうかは、今後の研究課題となる。

6.1.3 行動時刻を考慮した条件付き変分オートエンコーダによる推薦システム

5章では、検証実験を通じて、RNN エンコーダに Time-LSTM を採用し、VAE に時刻を条件づけた CVAE を用いた提案手法である TSVAE のほうが他の VAE ベースの既存手法に比べて精度が高くなることを示せた。時間情報が正確にサンプリングされていないデータセットである MovieLens における実験では、RNN エンコーダに時間情報を用いない単純な GRU や LSTM のほうが、時間情報を用いる Time-LSTM に比べて精度が高かった。このことから、正確な行動時刻情報をサンプリングできていないデータセットについては、学習時に時間情報がノイズになってしまっている可能性が考えられる。ただし、時刻情報を用いた本提案手法 TSVAE では、そのいずれの手法よりも精度が勝っていることから、時刻情報を学習時に取り入れることができれば、推薦精度の向上に繋がることが考えられる。また、Last.fm データセットにおける実験結果が他と比べて良かったことから、データセット内の時刻のばらつきがあるほど本提案手法の有効性があるが示唆される。特に、時間情報を直接学習データに掛け

合わせることで学習に反映させる Latent Cross の性能が他のデータに比べて高かったことから、時間情報がより正確に収集できている場合にはより単純な Latent Cross も有効であることが示唆される。

6.2 本研究全体の考察

これらの研究により、需要予測には時間情報を入力し活用することが有効であり、精度の向上に結びつけることができると考えられる。そして、正確な時間情報が得られない場合においても、順序や時間間隔といった時間に類する情報を活用することによっても、精度を高めることができると考えられる。また、時間情報を予測時に単純に条件として付与するだけでは精度を向上させるとは限らず、精度を高めるためにはその付与の仕方に工夫が必要であり、一方でその方法によって予測精度を大きく変化させることができるということも示唆される結果となった。このような指摘は [Beutel 18] でも行われていたが、本研究はそれを複数の実験で裏付けたことになる。

また、本論文内での複数の研究では、顧客や市場の次回経過に伴う変化に対してそのパターンに明示的な仮定を置いた上でそれを学習できるようにモデルを設計する手法と、そのような明示的な仮定は行わず、想定される変化を学習できるようにモデルを設計した上でモデルの学習によってパターンを予測できるようにする手法という、複数のアプローチを検討した。前者では観察などに基づいてより適切な仮定を置くこと、後者ではより幅広い変化のパターンに対応できるようなモデルを設計することが、より精度の高い予測につながるということが示唆された。さらに、これら 2つのアプローチを併用して予測を行うことが有効である可能性を示す結果も得られた。これは、明示的な仮定を置いたモデルによって、それで表される時間経過による変化パターンを学習し、その一方でそのモデルでは想定していなかった変化パターンを、併用したモデル側で学習することによって、精度の高い予測が行えるようになったということが考えられる。

また、各章で示した内容はいずれも同時並行して用いることができるため、図 1.1 に示すようなバリューチェーン全体の最適化に貢献することができる。第 3 章の提案手法で商品の需要を予測し、その結果に応じて商品を生産することで市場レベルの最適化を行い、第 4 章および

第5章の提案手法によって想定される顧客に商品を購入されるようにすることで、個人レベルの最適化を実行することができる。そして、実際の売上状況やレビュー状況のデータを収集してそのフィードバックとすることで、マーケティング・プロセスを改善していく。以上のようなサイクルを繰り返すことによって、提案手法の予測結果が改善されるだけでなく、その結果の活用方法も洗練されていき、マーケティング全体の効果が高まっていくと考えられる。

さらに、本研究で提案した手法は汎用的に適用できるように設計されているため、商材を問わず、たとえ海外市場でのマーケティングであっても活用できるものである。本研究の提案手法を活用するには、顧客個人単位や市場単位の売上数値データの他に、それに用いるデータ収集時のタイムスタンプ情報を付与するだけで良いため、他の商品や顧客の属性データに比べて収集は容易であり、その活用先の制約も小さい。ゆえに、様々な需要予測手法において本研究成果は十分その基礎と位置づけられるものであると言える。

6.2.1 変数としての時間情報 t の取り扱い

本研究では、いずれも時間情報 t を導入していたが、ここで各研究の t の取り扱い方について考察する。

観測データを推定に用いるには、そのデータに合った尺度を導入し、そのもとで活用していくことになる。その際に用いられる変数尺度は主に4つに大別することができ、表6.1にまとめられる。この観点から本研究を振り返ると、3章および5章では t を間隔尺度として、4章では t を順序尺度として導入していたことになる。

他の多くの機械学習を用いた研究において、名義尺度はカテゴリやラベル、条件として条件付き確率の推定などで利用されることが多く、比率尺度は多くの入力データ内の説明変数に用いられている尺度であると言える。一方で、本研究で用いた順序尺度、間隔尺度という観点は、他の機械学習の研究では一般的に導入されるようなものではなく、特殊な事例であると言える。この尺度水準としての性質が通常用いられる特徴量と異なるという点が、 t を既存の予測モデルに単純に条件として付与しても精度の向上につながらず、本研究で示したような工夫が必要になる原因の一つであると考えられる。このことから、本研究の t の取り扱いは、よ

表 6.1 尺度水準とその性質

尺度水準	大小がある	単位がある	絶対原点がある	用途
名義尺度	-	-	-	頻度を数えて集計する
順序尺度	○	-	-	対象を順序付ける
間隔尺度	○	○	-	和や差を計算する
比率尺度	○	○	○	四則演算等を行う

り幅広いデータを取り扱うための機械学習手法の発展にも貢献していると考えられる。

6.3 本研究結果の制約と限界

6.3.1 本提案手法による予測可能な期間範囲

前項で述べたとおり、本研究で取り入れた時間情報は理論上では無限に大きい値を取ることが可能である。しかし、本研究で提案された手法によって無限に先の未来を予測することが可能であるとは言えない。本項では、本研究の提案手法によって予測が可能な時間の幅について考察する。

本研究で提案された手法は、理論上、無制限に長期のデータを入力し、無制限に先の未来の予測を行うこと自体は可能である。実際は、入力するデータサイズが大きくなることで計算リソースが足りなくなることから予測ができなくなってしまう可能性があるため、その際は、データのサンプリング周期を変更したり、適宜入力するデータを制限したりすることで対応することになる。たとえば、第 3 章の提案手法で 10 年単位の予測を行う際、1 時間毎にサンプリングされたデータを入力すると、その入力データが大きくなりすぎてしまうため、粒度を月単位にすることによって現実的なデータサイズに変更することになる。

しかし、理論的には提案手法によって無限に未来を予測可能ではあるものの、その予測は現実的ではなく、予測精度が担保されないと考えられる。本研究が提案する予測手法は、入力データの時期と予測対象とする時期において、顧客の行動パターンに大きな違いがないということが前提となっている。実際は、データセットからは予測ができないような行動変容が生じ、それがデータセットに影響を与え、入力データの時期と予測対象とする時期のデータの性

質の差が大きくなってしまい、未来の予測になればなるほどその精度が落ちていくことが考えられる。たとえば、2000年代までのデータセットを用いて音楽CDの売上を予測する際には、後年に音楽のダウンロード販売やストリーミングサービスが流行してくことによって音楽CDの売上自体が低下すること自体を、本研究が提案する手法では予測に反映させることができない。また、Webサービスなどプラットフォームのログデータを予測に使用する際にも、そのユーザは流動的であり、数年をかけたユーザの入れ替わりや、提供サービスの内容の変化が発生すると言われている。他にも、Webサービス上の行動変化が、それに紐付いた商品への興味に変化していることではなく、そもそものユーザ自体が離脱してしまっている可能性が十分考えられる。このような行動変容は、Webサービスではないデータセットにも発生しうると考えられる。たとえば、小売店における商品の売上データの予測では、競合店が出現することや、ECサイトやフリーマーケットといった異なる購入チャネルが登場すること、その商品の代替品となりうる新製品が登場することで顧客が乗り換えるといった事象が考えられる。

以上のような理由で、本予測手法が適用できる将来予測の範囲は長くても数年単位になってしまうと考えられる。また、技術的イノベーションやパンデミックによる行動変容のような、顧客の行動が非連続的に大きく変化する自体が発生してしまった際には、そもそも本研究の提案手法を用いた予測自体が困難になってしまっていると考えべきである。

6.3.2 商品価格が需要予測に与える影響

実際の商品の需要は、その価格によっても大きな影響を受ける。本項ではその影響について述べる。

商品は基本的に、価格を下げると需要が増える傾向がある。そのため、実際にミクロ経済学では需要曲線と供給曲線を用いて商品の需要を推定したり、売上を最大化する最適価格を決定したりすることが行われる。ただし、商材の性質によって商品価格が需要に与える影響は異なる。

たとえば、日用品や生活必需品は、価格の影響を受けやすい。なぜなら、顧客はその商材を購入することは決まっているが、競合他社より代替可能な製品が販売されており、それらの中でより価格の低い商品を選択する傾向があるからである。一般に生活必需品は価格弾力性が小

さいと言われているが、特定の商品の需要はその商材の中での相対価格によって決定される。言い換えると、たとえ商品そのものの価格自体は変更しなくても、競合商品の価格の変化によって需要が変動してしまうということである。それは、需要が外部環境の影響を強く受けやすいということの意味している。

一方、嗜好品や創作品は、商品の絶対価格が需要に影響を与える。これらの商材は先述した日用品などに比べ、類似する競合商品のほうがより安価だからという理由で購入対象を変更するということが考えづらい。その価格は、顧客が想定していた予算や、顧客がそれぞれ持つ購買を判断する価格点（留保価格）と比べるための材料であり、その大小によって購入するか否かが決定される。言い換えると、商品そのものの価格を変更しない限り、価格要因の需要の変動の影響を受けにくいということの意味する。

本研究では、価格が需要に与える影響については考慮されていない。その一番の理由として、価格が需要にどのように影響を与えているか、その関連性を表す十分なデータが存在しないことにある。実務上では、例えば EC サイトなどでは、A/B テストなどを通じて実証的にデータを採取し、売上を最大化できるような価格を動的に設定する、ダイナミック・プライシングといった手法が考えられる。

本研究における予測手法は価格が需要に与える影響を考慮していないので、商品の価格が一定であることを前提としていると言える。そのため、セールやキャンペーンによって予測対象商品の絶対価格が変更された場合、これまでの予測精度を担保するためには予測モデルにその内容を反映させる必要がある。それを反映させる方法としては、キャンペーン実施時点を欠損値として予測データとしてに用いないことや、キャンペーンの実施時期を追加条件として入力データに付与する、といったものが考えられる。また、ダイナミック・プライシングによる販売を継続するなど、幅広い商品価格に対する需要のデータを収集することができていれば、価格変化の影響を織り込んだ需要予測モデルが設計できるようになり、セールなど絶対価格の変更を行った際に考えられる需要の変化については、より正確な予測がされやすくなると考えられる。

6.3.3 本研究による予測手法が適する商材

前項で述べたように、商品の価格はその需要に大きな影響を与える。そのため、本研究の予測対象とする商材は、価格変動の影響を受けにくいことが望ましい。具体的には、商品の価格変動自体が小さいか、商品の価格弾力性が小さいということが挙げられる。このような商品の性質を言い換えると、以下のように言える。

嗜好品、創作品である

たとえば、映画や音楽のようなコンテンツは、競合となる製品が価格が低いからという理由で代替品として購入されるということが考えづらいため、相対価格による需要の変動を考慮する必要性が小さくなることから、本手法の予測が適していると言える。逆に、日常的に消費される食品や日用品は競合商品も多く、相対的な価格の差によって需要が変化しやすい傾向があるため、そのような商品の需要予測は困難であると言える。

標準となる価格が存在する

たとえば、漫画は一冊あたりの価格が、映画は標準的な料金が慣習的に定められており、売上を増やすために値下げをするような、マーケティング戦略として価格戦略が採用されづらい商材であると言える。このような商品は絶対的にも相対的にも価格変動による需要の変化が起こりづらいため、本研究の需要予測手法が適していると考えられる。一方で、標準小売価格が定められておらず、完全競争の状況下にあると考えられる日用品などは価格が変動しやすく、相対価格の変化によって需要が変動しやすくなるため、需要の予測が難しいという傾向がある。

在庫が大量にあるか、在庫を考慮する必要がない

たとえば、大量生産される日用品や音楽や映画のようなダウンロード・コンテンツは、在庫切れの影響やそれに伴う価格変化を考慮する必要がないため、本研究による需要予測が適している商材であると言える。逆に、上限が決まっている航空機の座席や、販売数が固定されている限定品のような商品は、在庫数によって価格を柔軟に変化させることが最終的な売上の向上に結びつくと考えられるため、価格を固定した上で需要を予測することよりも、在庫を全て売り切ることを前提とした上でダイナミック・プライシングなどの手法で価格を調整することが有効である商材であると言える。

全体として、第4章や第5章で予測の対象とした、映画や漫画、音楽といったコンテンツの需要予測に、本研究の手法は適していると考えられる。

6.3.4 その他の時間に関する諸問題と限界

1章では、需要予測の際に捨象されることの多い時間という概念について説明し、その問題点を提示した。本項では、本研究がそれに対しどのようなアプローチを行い、どのような点が解決できなかったかを述べる。

未来情報の観測不可能性

予測モデルの学習に用いるデータからは、それが収集された（サンプリングされた）時点の状況しか観測できないという問題を抱えている。本研究で提案した手法は、いずれも入力データからその傾向を外挿したり、入力時に欠けている情報を予測時に補完する機構が導入されている。3章で提案したCPNetは、CNNの他に線形スキップコネクションという外挿機構を組み合わせている。5章は入力データが予測対象に関してすべて観測されていることは前提としておらず、各種エンコーダを通して復元した際に復元される機構を導入している。そして、以上の補完は強制的に行われるものではなく、必要に応じて実行される。ただし、ここで行われる補完はあくまでも予測モデルによる推定であるため、サンプリング時に落とされている情報が多ければ多いほど、その推定の質も低下してしまう。ゆえに、サンプリング時点で発生するこの問題を解決する完璧な手法であるとは言えない。

需要発生要因の観測不可能性

本研究で提案した手法はいずれも、要因がそれ自体は観測不可能であるが存在はするという点を前提としつつ、もしそれが存在するとデータから学習できる場合には、それを予測に反映させられるような手法として構成されている。たとえば、3章では様々な周期を捉えられるような機構をあらかじめ設定し、その捉えられる範囲については自動で学習できるようにしている。4章のユーザレベルモデリングも、ユーザが様々な嗜好をそれぞれ異なる変化パターンで持つことを考慮し、既定の一定の変化モデルにならないように学習を進めている。本研究で提案した手法いずれにも共通する特徴は、あらかじめ周期や傾向をデータから学習できるような学習モデルを設計し、必要に応じて学習できるようにしてあるという点である。ただし、これらの変化がモデルで学習できることを前提としているため、そのモデルで学習できる範囲を

超えるものは当然学習できない。そして、あらゆる周期や傾向を学習できるようにするには、モデル容量を無制限に大きくしなければならないため、そのモデルの学習限界は別途設定しなければならない。

さらに、時間に関する概念について、従来より存在し、本研究の提案手法においても対処することができない、以下のような問題も存在する。

本質的に予測不可能な事象による影響

もし需要の予測が完璧にできるのであれば、その予測される量の商品を生産し、想定される顧客全員に販売することによって在庫コストや機会損失を最小化されることによって期待される利益を最大化することができるが、世界には技術的イノベーションによる不可逆的で大きな市場の変化や、世界的な恐慌、パンデミックといった、いつか来ること自体は予測できても、その正確な時期までは現代の科学では予測ができない事象が存在する。このような予測不可能な事態はブラック・スワンと呼ばれる。マーケティングが対象とする人間の購買活動は、現代社会で人間が生活する上でほぼ必須な行動であり、根幹的なものであるため、これら予測不可能なあらゆる事象の影響を受けてしまう。そして、これらをすべての影響を予測に入れ込むことは難しい。言い換えるならば、この要素に関して2章で売上を決定する関数 R で代表させていたが、 R に関して指標 y を増加させることによって増加できるといった単純な性質であれば推定できるものの、 R のより詳細な形状自体は推定することができないということを意味する。他にも、機械学習ではベイズ誤差と呼ばれる、予測モデルを用いた際の予測性能の限界が存在する。このため、理論的にも完璧な予測モデルを設計することはできない。

また、突発的なイベントによる予測精度の低下については、本研究で提案する手法ではいずれも対応できない。なぜなら、本研究で主張する需要予測手法はいずれも過去の履歴データに基づいて予測を行うものであり、その過去に含まれないケースについてはそもそも予測ができないからである。それに対応するために、より長期間のデータを学習に使用することによって特殊な事例についても学習できるようにするといった方法も考えられるが、その特殊なケースがデータに含まれていなかった場合には、過去に基づく予測を強固なものとしてしまい、本来対応しなかったイレギュラーな場合に対し悪影響を及ぼす可能性もある。そのため、実務利用を考える上では、このような特殊なケースの予測自体を諦め、「予測モデルでは予測が不可能な事態が必ず発生しうる」ということを想定した上で、マーケティング・プロセスの実施や意

思決定の段階において、そのリスクをあらかじめ織り込んでおくことが求められる。特に市場レベルの需要予測の場合に想定されるが、もし背景のモデル自体が変化してしまうようなことがあった場合には、変化前のデータを学習に使わず、変化後のより新しいデータのみを用いて予測を行ったほうが最終的な予測結果が良くなるケースもあると考えられる。

6.3.5 データドリブン・アプローチ自体の限界

本研究で述べた提案手法の限界には、これまでに述べた時間に関するものだけでなく、データを活用して需要を予測を行おうとする時点で生じるものがある。

たとえば、4章において、データ収集の観点では興味の推移を正確に反映した興味推移シーケンスを作成しづらいという問題も抱えている。たとえば、あるユーザが現在たまたま世間で人気になっているからとりあえず周囲に合わせて興味行動を発生させたため、実際は興味を持っていないにもかかわらず興味行動を起こしたという場合が考えられる。他にも、コンテンツの登場がログデータ収集を行う期間の後半で登場した場合には、このユーザの興味レベルが高まったことでこのコンテンツに興味をもったのか、それともコンテンツの市場への登場が後半であったことによるものなのか、どちらであるかの判定はできない。

他にも、顧客の興味動向を定量的に的確に示す正解データが存在しないことが、手法のさらなる精度向上を難しくするという問題がある。しかし、市場全体におけるコンテンツのトレンドを定量的に集めることは困難である。たとえアンケートを実施した場合でも、世の中の全コンテンツに対するトレンドを収集することはできず、さらに、消費者がまだ表立って出していない潜在的な人気を測定することもできない。このような状況下では、近似的に正解データとして扱えそうなデータを用いて手法を検証し、専門家との議論を重ねつつ、妥当性のある精度向上方法を検討することによって対応することが求められると言える。

他にも、顧客の持つ嗜好や興味は、行動にその一部分しか反映されないため、すべての嗜好や興味に対応する行動を観測するということとはできない。MovieLens レビューのような能動的な行動の場合、その顧客が興味を持つあらゆる作品に評価が行われるという可能性は低い。ゆえに、評価が行われていないということが、必ずしもその作品を評価していないということ

を意味しない。また、顧客に関するあらゆる行動のデータを収集できることが望ましいが、それも不可能である。

より需要の予測精度を高めるために顧客の行動データをさらに集めるという目的で、幅広い行動ログデータを収集し、需要予測に活用するアプローチが考えられる。ただし、興味があった起こしたと仮定した行動が、必ずしもそれに対し興味を持っていることを意味しないことには注意しなければならない。たとえば、お気に入り登録がその対象に対して必ずしも好意を持って行われた行動ではないことを指摘する研究がある [Meier 14]。

さらに、個人の行動ログデータを活用した需要予測は、簡単に攻撃され、その影響を受けてしまうという問題を含んでいる。たとえば、SNS で口コミを意図的に発生させて、それを見た人に人気があるように思わせるステルスマーケティングという手法がある。商品の購買履歴データのような実行時に金銭を必要とする行動はこれを行うための敷居は高いが、SNS 上での情報発信や商品へのレビューのような行動は誰でも簡単に実行できるため、その行動が顧客の真の心理による行動であると仮定している本研究の予測手法は、その攻撃的な影響を受けてしまう可能性があると考えられる。その悪影響を回避するためには、以上のような行動を検知して除外、禁止するシステムを構築するなど、ログデータ収集の段階でそのようなデータが入り込まないような設計を行うことが求められる。

関連して、本研究で提案した一部の手法では、結果を解釈してその原因を考察したり、なぜこのような出力となったのか説明したりすることができないという問題がある。特に第3章や第5章で提示した深層学習を用いた手法は、説明可能性を犠牲にした上で精度を高めた手法であると言える。このような結果が解釈できなかつたり説明がつかなくなかつたりする手法には、上記したような予測結果を狂わせる事態が発生した際にそれに気づくことが出来ないということや、予測結果を踏まえて相乗効果を狙ったマーケティング・プロセスを考えることが困難であるという欠点がある。近年では説明可能な AI(XAI) の技術が注目されているが、本研究で提案したような手法がマーケティングの現場でより信頼を獲得しつつ用いられるようにするためには、単純に予測精度を高めていくだけではなく、因果推論を行うなど、説明可能性を高めるといったアプローチもあると考えられる。

他にも、本研究で述べた予測アプローチに共通する限界として、影響を与えるすべての要因

となるデータを取り込むことはできないという点が挙げられる。IoT(Internet of Things)の普及が進むことで、より多くのデータが収集できるようになるが、GDPR(一般データ保護規則)など個人情報保護の観点から、収集や活用が許されないデータも存在する。実務上では、このような収集が許されていないデータに関して収集できるデータから推定することになるが、その推定には限界がある。本研究で行った実験は、比較的入手がしやすいデータのみを用いて予測を行っているが、様々なデータを条件として付与することができれば精度を向上させることができると考えられる。

また、仮に、すべての要因となるデータを予測時に取り込むことができたとしても、それらをすべて活用しようとする、計算リソースが不足することが考えられる。本研究における各実験も、マシンリソースと時間の制約上、いずれもデータを一部削減した上で実行している。実務での利用においては、これらの予測は高い精度が出たほうが良いものの、このような限界があるため、予測を踏まえた以降のマーケティング・プロセス全体においてカバーすることになる。

6.4 本研究の他ドメインへの応用可能性

本研究で対処を試みた時間という概念は、マーケティング以外の分野にも影響する汎用的な概念である。これを踏まえて、本研究の対象とした需要予測の考え方をより一般的に拡張し、どのように他の分野に適用できるかを議論する。

市場レベルの需要予測の一般化

本研究で取り扱った市場レベルの需要予測とそれに対する提案手法は、言い換えると、周期性を確定できない時系列に対するトップダウンな予測手法である。時系列予測の研究自体は2章で述べたように様々な分野において、それぞれ異なる特徴を持つ手法が提案されており、その適用範囲も経済学、工学、理学など多岐に及んでいる。

発生源のモデルが定義できない、もしくはその必要がないものの、その出力は予測したいという動機は、機械学習の文脈ではたびたび発生するものである。機械学習による時系列モデルについても2章で様々な例を挙げているが、本研究で提案した手法のように複数の内部パター

ンを想定し、それらを重ね合わせる予測手法はバギングの考え方にも類似していると言える。3章で示した深層学習や単純な線形モデルを組み合わせる方法は、容量が大きく複雑な表現をより厳密に再現できるモデルと、自由度は小さいものの望ましい制約をかけることができるモデルの組み合わせであると言え、このような組み合わせは様々な応用可能性があると考えられる。

また、今回予測を試みた需要の時系列は、様々な要因が複雑に絡んだ結果として出力され、かつ外部の予測不可能なノイズの影響を強く受けがちな時系列であると言える。そのような時系列の例として、マーケティングに近い分野としては、株価や先物価格といった金融分野の時系列が挙げられる。他のドメインでは、それは測定時にノイズの影響を強く受けがちな生物学における時系列と類似点があると言える。本研究での提案手法やそれに向けた考え方は、生物の内部活動も様々な要因が影響しているという点も共通しているため、データが十分に用意でき、内部構造のモデリングを行う必要性がない場合には、本研究の成果が活用できる可能性があると言える。

個人レベルの需要予測の一般化

個人レベルの需要予測で予測対象とした顧客の購買行動を、より一般的な枠組みから考えると、顧客をエージェントとみなし、顧客はそれぞれ情報処理モデルやメンタルモデルを持つ存在であるとみなせる。そして、顧客は外界から得られる情報を処理し、メンタルモデルの出力を基に購買の意思決定を行う。また、各エージェント内のモデル内のパラメータが異なることから、たとえ外界からの同様の入力があったとしても、各エージェントごとに行動結果（購買するか否か）が異なることになる。エージェントの持つメンタルモデルは外部の影響を受けるので、マーケティング・プロセスで行われる各種施策は、各エージェントのメンタルモデルのパラメータを変更するように働きかけることによって、最終的な購買確率を向上させる行為であると言える。

すなわち、個人レベルの需要予測を一般的に言い換えると、エージェントの意思決定モデルの推定であると言える。モデルの形式やパラメータが推定し、そのメカニズムを少しでも理解することができれば、それに働きかける施策を行うことで、エージェントに対して実行してほしい行動を起こす確率を高めることができると考えられる。4章の提案手法はモデルの形状を

トップダウンで明示的に定義した部分的なモデリング、5章の提案手法はそのような明示的なモデリングを行わず購買確率を推定する手法であると言える。このように考えると、本研究の成果は、マルチエージェントの研究が対象とするような様々なドメインのタスクに応用できると言える。

第7章

結論

本研究では、市場レベル及び個人レベルの需要予測アプローチにおいて時間の要素を入力データから学習して取り入れるアーキテクチャを導入し、予測精度を向上させるアプローチを複数提案した。予測対象データにおける時間の影響そのものを直接モデリングしたことなく、その影響を学習できるモデルを用いることによって、既存の予測アプローチよりより高い精度を出すことができるということを示していく。

はじめに、市場レベルの需要予測のアプローチとして、複数の周期を自動的に取り入れられる売上時系列予測手法を提案した。これまでに商品の売上予測のために様々な時系列予測のアプローチが提案されているが、それに生じている周期を把握した上でモデリングを行い、その上でパラメータを調整しないと精度が出せないという問題を抱えている。この研究では、深層学習に畳込みネットワークを活用して、店舗や商材ごとの複数の周期を同時に学習し、予測結果に織り込むアプローチを検討した。検証実験の結果、深層学習を用いた他の時系列予測手法と比べて、高い精度を記録したことを示した。

次に、個人レベルの需要予測のアプローチとして、ユーザが複数の趣味嗜好を持ち、かつそれが時間とともに成熟したモデリングを行うことで推薦したモデルを提案した。従来の商品推薦モデルはユーザの趣味嗜好が過去の行動履歴によって推定され、それが時間に対して不変であるということを前提としていた。しかし、商材によっては、その対象への成熟度が変化したことで趣味嗜好が変わっていくことが仮定できる場合も考えられる。そこで、この研究ではユーザが対象商材に対し複数種類の趣味嗜好のカテゴリを持ち、それらが時間によって成熟していくということを仮定した上で、商品を推薦したアプローチを検討した。対象として成熟度

が変化したコンテンツに着目し、本提案手法によってより高い精度でより多くの商品を推薦の対象とできるようになることを示した。

最後に、先ほどの需要予測モデルの発展型として、ユーザの行動時刻を付与条件としてコンテンツ推薦したモデルを提案した。近年、より少ないデータセットに対し、より頑健性を持った推薦モデルとして、変分オートエンコーダを用いたものが注目されている。しかし、このモデルはユーザの行動順序のみに着目し、行動した時刻やその時間間隔を考慮に入れられていない。本研究では、このモデルに対してもユーザの行動時間を適切に条件として入力し、推薦精度を高めることができることを示した。時間情報を条件として付与した複数のアプローチを検証したことによって、行動時間情報をより適切に取り扱う手法を提案した。

以上の複数の研究を通じて時間の要素を取り入れることによる予測モデルの精度の向上を示し、これらの研究の共通点および相違点を整理した。それを踏まえて、市場レベル及び個人レベルの需要予測アプローチにおいて時間の要素を考慮したモデルを用いることが、そうでないアプローチに比べてどのような利点があると考えられるか、精度以外の面からも考察した。そして、本研究による予測アプローチや、それ以外の市場や需要の予測アプローチ自体の限界を考察した。さらに、本研究から得られる知見が今後どのように発展してマーケティングプロセスの改善に貢献していくかを検討し、それに加えて、異なるドメインの予測タスクに対しても本研究がどのように貢献できるかを議論した。

謝辞

本博士論文は、研究室の方々を始め、多くの方々のご協力がなければ完成できませんでした。この場をお借りして感謝を申し上げます。

指導教員である松尾 豊 教授には、自分が学部、修士課程の頃からの長い間、ご指導・ご鞭撻を賜りました。機械学習や人工知能の理論のみならず、これからの社会を生き抜くための物事の捉え方や考え方、研究者のとしての心構えといった面でも大変勉強になりました。また、経済産業省や企業との共同研究の機会をいただけたことも、よい社会勉強になりました。深層学習が世界を変えていく様子を研究室の最前線で見られたことは、とても幸運なことでした。振り返ると、松尾研究室に在籍した期間は総じてとてもエキサイティングであり、とても有意義な人生経験だったと思います。本当にありがとうございました。

本論文の副査を担当していただきました東京大学大学院 工学系研究科 坂田 一郎 教授、武田 史子 准教授、田中 謙司 准教授、鳥海 不二夫 准教授には、貴重なお時間を割いて本論文の審査および改善のためのご指摘をしていただきました。この様々な方面からのご指導は、この論文の完成度を高めるのみならず、自身の研究を深める上でも有益なものでした。ありがとうございました。

研究室の皆様にも、大変お世話になりました。本論文は、研究会や輪読会での日々の学びが基になっております。優秀な先輩・同期・後輩に囲まれて研究活動を進めることはプレッシャーでもありましたが、良い刺激となり研究を進める推進力になったように思います。また、自分の研究活動に集中できたのは、裏で事務作業やハードウェアの整備などを進めてくださった秘書やスタッフの方々のお陰であり、大変感謝しております。特に、岩澤 有祐 特任講師には、各研究について定期的に個別に相談する時間を割いていただきつつ、研究がうまく行

かず苛立ちや不安を覚えた際にも心の支えとなっていただきました。岩澤氏のご指導がなければ、自身の研究を博士論文の水準に磨き上げて博士課程を修了することはできなかったと思います。本当にありがとうございました。

最後に、この博士課程の修了までに経済面や精神面で支えていただきました家族に感謝いたします。博士課程での研究や論文執筆が順風満帆には進まなかったため、これまでに多くの不安を与えてしまったと思いますが、最後まで暖かく見守っていただいたことに深く感謝いたします。

東京大学大学院 工学系研究科
技術経営戦略学専攻
保住 純

参考文献

- [Abel 10] Abel, Fabian, et al. "Analyzing the blogosphere for predicting the success of music and movie products." *Advances in Social Networks Analysis and Mining (ASONAM)*, 2010 International Conference on. IEEE, 2010.
- [Adomavicius 11] Adomavicius, Gediminas, and Alexander Tuzhilin. "Context-aware recommender systems." *Recommender systems handbook*. Springer, Boston, MA, 2011. 217-253.
- [Antin 12] Antin, Judd, Coye Cheshire, and Oded Nov. "Technology-mediated contributions: editing behaviors among new wikipedians." *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 2012.
- [Arunaj 15] Arunaj, Nari Sivanandam, and Diane Ahrens. "A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting." *International Journal of Production Economics* 170 (2015): 321-335.
- [Beutel 18] Beutel, Alex, et al. "Latent cross: Making use of context in recurrent recommender systems." *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 2018.
- [Bian 09] Bian, Jiang, et al. "Learning to recognize reliable users and content in social media with coupled mutual reinforcement." *Proceedings of the 18th international conference on World wide web*. ACM, 2009.
- [Bin 12] Xu, Bin, et al. "An exploration of improving collaborative recommender systems via user-item subgroups." *Proceedings of the 21st international conference on World Wide Web*. ACM, 2012.
- [Bińkowski 17] Bińkowski, Mikołaj, Gautier Marti, and Philippe Donnat. "Autoregressive convolutional neural networks for asynchronous time series." *arXiv preprint arXiv:1703.04122* (2017).

- [Borovykh 17] Borovykh, Anastasia, Sander Bohte, and Cornelis W. Oosterlee. "Conditional time series forecasting with convolutional neural networks." arXiv preprint arXiv:1703.04691 (2017).
- [Chen 16] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. ACM, 2016.
- [Cheng 16] Cheng, Heng-Tze, et al. "Wide & deep learning for recommender systems." Proceedings of the 1st workshop on deep learning for recommender systems. ACM, 2016.
- [Chevillon 07] Chevillon, Guillaume. "Direct multi - step estimation and forecasting." Journal of Economic Surveys 21.4 (2007): 746-785.
- [Choi 12] Choi, Hyunyoung, and Hal Varian. "Predicting the present with google trends." Economic Record 88.s1 (2012): 2-9.
- [Drucker 73] Peter Drucker, "Management: Tasks, Responsibilities, Practices," New York: Harper & Row, 1973.
- [Dunn 73] Dunn, Joseph C. "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters." (1973): 32-57.
- [Goel 10] Goel, Sharad, et al. "Predicting consumer behavior with Web search." Proceedings of the National Academy of Sciences 107.41 (2010): 17486-17490.
- [Goel 17] Goel, Hardik, Igor Melnyk, and Arindam Banerjee. "R2n2: Residual recurrent neural networks for multivariate time series forecasting." arXiv preprint arXiv:1709.03159 (2017).
- [Harvey 90] Harvey, Andrew C., and Simon Peters. "Estimation procedures for structural time series models." Journal of Forecasting 9.2 (1990): 89-108.
- [He 17] He, Xiangnan, et al. "Neural collaborative filtering." Proceedings of the 26th international conference on world wide web. 2017.
- [Kaneko 16] Kaneko, Yuta, and Katsutoshi Yada. "A deep learning approach for the prediction of retail store sales." 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). IEEE, 2016.
- [Kang 17] Kang, Wang-Cheng, et al. "Visually-aware fashion recommendation and design with generative image models." 2017 IEEE International Conference on Data Mining (ICDM). IEEE, 2017.
- [Kastov 18] Katsov, Ilya. AI アルゴリズムマーケティング 自動化のための機械学習/経済モデ

- ル, ベストプラクティス, アーキテクチャ. インプレス, 2018.
- [Ke 17] Ke, Guolin, et al. "Lightgbm: A highly efficient gradient boosting decision tree." *Advances in Neural Information Processing Systems*. 2017.
- [Kingma 13] Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." *arXiv preprint arXiv:1312.6114* (2013).
- [Kingma 14] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).
- [Kotler 16] Kotler, Philip, et al. "Marketing management. 15th Global Edition." England: Pearson Educationn Limited 17 (2016).
- [Lai 18] Lai, Guokun, et al. "Modeling long-and short-term temporal patterns with deep neural networks." *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 2018.
- [Lauschke 12] Lauschke, Claudia, and Eirini Ntoutsi. "Monitoring user evolution in twitter." *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*. IEEE Computer Society, 2012.
- [Levitt 86] Levitt, I. M., and Theodore Levitt. *Marketing Imagination: New*. Simon and Schuster, 1986.
- [Liang 18] Liang, Dawen, et al. "Variational autoencoders for collaborative filtering." *Proceedings of the 2018 World Wide Web Conference*. 2018.
- [Liu 17] Liu, Xin, and Ryutaro Ichise. "Food sales prediction with meteorological data—a case study of a japanese chain supermarket." *International Conference on Data Mining and Big Data*. Springer, Cham, 2017.
- [Liu 19] Liu, Hao, et al. "Joint representation learning for multi-modal transportation recommendation." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 2019.
- [McAuley 13] McAuley, Julian John, and Jure Leskovec. "From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews." *Proceedings of the 22nd international conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2013.
- [Meier 14] Meier, Florian, David Elsweiler, and Max L. Wilson. "More than Liking and Bookmarking? Towards Understanding Twitter Favouriting Behaviour." (2014).
- [Mestyán 13] Mestyán, Márton, Taha Yasseri, and János Kertész. "Early prediction of movie box

- office success based on Wikipedia activity big data.” PloS one 8.8 (2013): e71226.
- [O’ Connor 99] O’ Connor, Mark, and Jon Herlocker. ”Clustering items for collaborative filtering.” Proceedings of the ACM SIGIR workshop on recommender systems. Vol. 128. UC Berkeley, 1999.
- [Oord 17] Van Den Oord, Aron, et al. ”WaveNet: A generative model for raw audio.” SSW 125 (2016).
- [Pavlovski 20] Pavlovski, Martin, et al. ”Time-Aware User Embeddings as a Service.” Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020.
- [Sachdeva 19] Sachdeva, Noveen, et al. ”Sequential variational autoencoders for collaborative filtering.” Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. 2019.
- [Sarwar 02] Sarwar, Badrul M., et al. ”Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering.” Proceedings of the fifth international conference on computer and information technology. Vol. 1. 2002.
- [Sen 09] Sen, Shilad, Jesse Vig, and John Riedl. ”Tagommenders: connecting users to items through tags.” Proceedings of the 18th international conference on World wide web. ACM, 2009.
- [Shaw 12] Shaw, Arch Wilkinson. ”Some problems in market distribution.” The Quarterly Journal of Economics 26.4 (1912): 703-765.
- [Sohn 15] Sohn, Kihyuk, Honglak Lee, and Xinchun Yan. ”Learning structured output representation using deep conditional generative models.” Advances in neural information processing systems. 2015.
- [Taylor 18] Taylor, Sean J., and Benjamin Letham. ”Forecasting at scale.” The American Statistician 72.1 (2018): 37-45.
- [Tsoumakas 18] Tsoumakas, Grigorios. ”A survey of machine learning techniques for food sales prediction.” Artificial Intelligence Review (2018): 1-7.
- [Xue 17] Xue, Hong-Jian, et al. ”Deep Matrix Factorization Models for Recommender Systems.” IJCAI. Vol. 17. 2017.
- [Yang 14] Yang, Jaewon, et al. ”Finding progression stages in time-evolving event sequences.” Proceedings of the 23rd international conference on World wide web. International World

Wide Web Conferences Steering Committee, 2014.

[Zhao 17] Zhao, Zhou, et al. "Social-aware movie recommendation via multimodal network learning." *IEEE Transactions on Multimedia* 20.2 (2017): 430-440.

[Zheng 17] Zheng, Lei, Vahid Noroozi, and Philip S. Yu. "Joint deep modeling of users and items using reviews for recommendation." *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. 2017.

[Zhu 17] Zhu, Yu, et al. "What to Do Next: Modeling User Behaviors by Time-LSTM." *IJCAI*. Vol. 17. 2017.

[伊藤 05] 伊藤 剛. "テヅカ・イズ・デッド ひらかれたマンガ表現論へ." NTT 出版, 2005.

[新倉 05] 新倉 貴士. "消費者の認知世界—ブランドマーケティング・パースペクティブ." 千倉書房, 2005.

[松島 09] 松島 克守. "東京大学の俯瞰経営学—技術経営戦略学専攻講義." 工業調査会, 2009.

発表文献

論文誌

- [1] 保住 純, 岩澤 有祐, 松尾 豊 ”複数周期に着目した畳み込みネットワークによる小売時系列データの予測手法.” 情報処理学会論文誌 (投稿中) .
- [2] 保住 純, 岩澤 有祐, 松尾 豊 ”ユーザの行動履歴データを用いたコンテンツ興味推移カテゴリの可視化手法.” 人工知能学会論文誌 (投稿中) .
- [3] 保住 純, 岩澤 有祐, 松尾 豊 ”行動時刻を考慮した条件付き変分オートエンコーダによる推薦システム.” 人工知能学会論文誌 (照会) .

国内学会発表 (査読なし)

- [1] 保住 純, 松尾 豊 ”ユーザの行動履歴データを用いたコンテンツ興味推移カテゴリの可視化手法”, 2017 年度 人工知能学会全国大会, 2017.
- [2] 保住 純, 松尾 豊 ”深層敵対的生成ネットワークを用いたマンガイラスト自動生成の試み”, 2018 年度 人工知能学会全国大会, 2018.
- [3] 保住 純, 松尾 豊 ”線形回帰モデルを活用した深層時系列予測モデルの比較検討”, 2019 年度 人工知能学会全国大会, 2019.
- [4] 保住 純, 松尾 豊 ”行動時刻を考慮した VAE による推薦システム”, 2020 年度 人工知能学会全国大会, 2020.

訳書

- [1] Ian Goodfellow, Yoshua Bengio, Aaron Courville 著, 深層学習, 分担翻訳 (第 5 章), KADOKAWA.

付録 A 提案手法による可視化マップ

表 1 提案手法による MovieLens ユーザ興味推移可視化の結果 (1)

レベル	Group.1	Group.2
Level.1	Godfather, The American Beauty Star Wars: Episode IV - A New Hope Star Wars: Episode VI - Return of the Jedi One Flew Over the Cuckoos Nest Schindlers List Saving Private Ryan Raiders of the Lost Ark Star Wars: Episode V - The Empire Strikes Back Jurassic Park	Star Wars: Episode IV - A New Hope Star Wars: Episode V - The Empire Strikes Back Raiders of the Lost Ark Saving Private Ryan Godfather, The Silence of the Lambs, The American Beauty Star Wars: Episode VI - Return of the Jedi Sixth Sense, The Schindlers List
Level.2	Fargo Pulp Fiction Monty Python and the Holy Grail Back to the Future GoodFellas American Beauty Star Wars: Episode V - The Empire Strikes Back M*A*S*H Blade Runner Shakespeare in Love	Matrix, The Terminator 2: Judgment Day Terminator, The Aliens Fugitive, The Alien Indiana Jones and the Last Crusade Braveheart Back to the Future Hunt for Red October, The
Level.3	Crying Game, The Apollo 13 Swingers Full Monty, The Jerry Maguire Fugitive, The Sling Blade Three Kings Leaving Las Vegas Boogie Nights	Jurassic Park Men in Black Independence Day (ID4) Star Trek: First Contact Fifth Element, The Total Recall Stargate Indiana Jones and the Temple of Doom Rock, The Star Wars: Episode I - The Phantom Menace
Level.4	Men in Black Forrest Gump Aladdin Blues Brothers, The Big Ferris Buellers Day Off Caddyshack Splash Good Morning, Vietnam Fast Times at Ridgemont High	Groundhog Day Forrest Gump Toy Story Ferris Buellers Day Off Aladdin When Harry Met Sally... Shakespeare in Love Grosse Pointe Blank Lion King, The Austin Powers: International Man of Mystery
Level.5	Exorcist, The Godfather, The Shining, The Psycho Birds, The Godfather: Part II, The Rosemarys Baby Unforgiven Alien Poltergeist	Tombstone Shining, The Gladiator Misery Sixth Sense, The Exorcist, The Pretty Woman Blair Witch Project, The Ghost Scream
Level.6	One Flew Over the Cuckoos Nest Taxi Driver Apocalypse Now Graduate, The Chinatown Rear Window Casablanca Close Encounters of the Third Kind Amadeus Silence of the Lambs, The	Platoon Patton Star Wars: Episode V - The Empire Strikes Back Monty Python and the Holy Grail Graduate, The Fargo Silence of the Lambs, The Full Metal Jacket Apocalypse Now One Flew Over the Cuckoos Nest

表2 提案手法による MovieLens ユーザ興味推移可視化の結果 (2)

レベル	Group.3	Group.4	Group.5
Level.1	Star Wars: Episode VI - Return of the Jedi Jurassic Park Saving Private Ryan Matrix, The Braveheart Star Wars: Episode IV - A New Hope Terminator 2: Judgment Day American Beauty Star Wars: Episode V - The Empire Strikes Back Men in Black	Star Wars: Episode VI - Return of the Jedi Star Wars: Episode V - The Empire Strikes Back Star Wars: Episode IV - A New Hope Matrix, The Terminator 2: Judgment Day Jurassic Park Raiders of the Lost Ark Terminator, The Alien Braveheart	Star Wars: Episode VI - Return of the Jedi Matrix, The Star Wars: Episode IV - A New Hope Terminator 2: Judgment Day Star Wars: Episode V - The Empire Strikes Back Jurassic Park Braveheart Saving Private Ryan Raiders of the Lost Ark Men in Black
Level.2	Forrest Gump Little Mermaid, The Shakespeare in Love Sleepless in Seattle Beauty and the Beast Aladdin Gone with the Wind Mask of Zorro, The Independence Day (ID4) True Lies	Aladdin Toy Story Lion King, The Ghostbusters Beauty and the Beast Shakespeare in Love Groundhog Day When Harry Met Sally... Little Mermaid, The Forrest Gump	American Beauty Forrest Gump Toy Story Groundhog Day Shakespeare in Love Beauty and the Beast Aladdin Little Mermaid, The American Pie As Good As It Gets
Level.3	American Beauty Fargo Schindlers List Silence of the Lambs, The Pulp Fiction L.A. Confidential Shawshank Redemption, The GoodFellas Usual Suspects, The Saving Private Ryan	Fargo Star Wars: Episode IV - A New Hope Schindlers List Star Wars: Episode V - The Empire Strikes Back Raiders of the Lost Ark Casablanca Chinatown L.A. Confidential Blade Runner Silence of the Lambs, The	Sixth Sense, The Being John Malkovich Austin Powers: The Spy Who Shagged Me Gladiator Ghostbusters Erin Brockovich Galaxy Quest Sleepy Hollow X-Men Patriot, The
Level.4	Terminator 2: Judgment Day Star Wars: Episode V - The Empire Strikes Back Fight Club Seven (Se7en) Election Matrix, The Green Mile, The Raiders of the Lost Ark Princess Bride, The Groundhog Day	American Beauty Jurassic Park Men in Black Patriot, The Gladiator Terminator 2: Judgment Day Fly, The Frequency Rosemarys Baby Airplane!	Psycho Jaws Fly, The Rosemarys Baby Blade Runner Scream L.A. Confidential Shining, The Misery Evil Dead II (Dead By Dawn)
Level.5	As Good As It Gets Apollo 13 Good Will Hunting Theres Something About Mary Babe Hamlet Dances with Wolves Fugitive, The Few Good Men, A Clerks	Forrest Gump Apollo 13 Field of Dreams Star Trek: The Wrath of Khan Crimson Tide Rain Man Dead Poets Society Braveheart Professional, The (a.k.a. Leon: The Professional) Its a Wonderful Life	Princess Bride, The Star Wars: Episode IV - A New Hope Star Wars: Episode V - The Empire Strikes Back Silence of the Lambs, The Big Fugitive, The Jumanji Schindlers List Fargo Raiders of the Lost Ark
Level.6	Men in Black Waynes World Galaxy Quest Clueless Austin Powers: International Man of Mystery Total Recall Eyes Wide Shut Bugs Life, A My Cousin Vinny Double Jeopardy	Aliens Air Force One Last of the Mohicans, The Full Metal Jacket Mad Max Guns of Navarone, The Beetlejuice Thomas Crown Affair, The Lion King, The Alice in Wonderland	Braveheart Stand by Me Twelve Monkeys Contact Jurassic Park Matrix, The Reservoir Dogs Groundhog Day Star Trek: The Wrath of Khan Shawshank Redemption, The
Level.7	Ghostbusters Austin Powers: The Spy Who Shagged Me American Pie Gladiator Office Space Big Daddy Galaxy Quest Shakespeare in Love Stir of Echoes Run Lola Run (Lola rennt)	Starman Top Gun Lethal Weapon Crocodile Dundee Batman Legends of the Fall Robocop Vacation Alien? Karate Kid, The	Philadelphia Story, The Roman Holiday Star Trek: Generations Laura Cape Fear American Graffiti Keeping the Faith Lion King, The Lord of the Rings, The Taxi Driver
Level.8	Sixth Sense, The Talented Mr. Ripley, The X-Men Toy Story 2 Who Framed Roger Rabbit? Sleepy Hollow South Park: Bigger, Longer and Uncut Thomas Crown Affair, The What Lies Beneath Meet the Parents	First Blood Superman II NeverEnding Story, The Planet of the Apes Gladiator King Kong Star Trek: The Wrath of Khan Galaxy Quest Gremlins Animal House	Key Largo Hunt for Red October, The Angel Heart 2001: A Space Odyssey Arsenic and Old Lace Dune As Good As It Gets Back to the Future U-571 Star Wars: Episode IV - A New Hope

表3 提案手法による MovieLens ユーザ興味推移可視化の結果 (3)

レベル	Group.6	Group.7	Group.8
Level.1	Star Wars: Episode V - The Empire Strikes Back Star Wars: Episode IV - A New Hope Raiders of the Lost Ark Matrix, The Terminator, The Star Wars: Episode VI - Return of the Jedi Terminator 2: Judgment Day Alien Princess Bride, The American Beauty	American Beauty Sixth Sense, The Fargo Silence of the Lambs, The Saving Private Ryan Shawshank Redemption, The Being John Malkovich Pulp Fiction Schindlers List Braveheart	American Beauty Fargo Star Wars: Episode IV - A New Hope Being John Malkovich Saving Private Ryan Schindlers List Star Wars: Episode V - The Empire Strikes Back L.A. Confidential Pulp Fiction Princess Bride, The
Level.2	Ferris Buellers Day Off Men in Black Lethal Weapon Big Batman Breakfast Club, The Romancing the Stone Blues Brothers, The Ghostbusters When Harry Met Sally...	Stand by Me Forrest Gump When Harry Met Sally... Airplane! Back to the Future Princess Bride, The Ferris Buellers Day Off Sleepless in Seattle Big Four Weddings and a Funeral	Men in Black Forrest Gump Big Ferris Buellers Day Off Ghostbusters Full Monty, The Lethal Weapon Breakfast Club, The Austin Powers: International Man of Mystery Blues Brothers, The
Level.3	Fargo Graduate, The Shawshank Redemption, The Wizard of Oz, The Godfather, The Casablanca Citizen Kane Silence of the Lambs, The One Flew Over the Cuckoos Nest Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb	Almost Famous American Beauty Schindlers List L.A. Confidential Fargo Pulp Fiction Godfather, The Shawshank Redemption, The Erin Brockovich Silence of the Lambs, The	Fargo Graduate, The Godfather, The One Flew Over the Cuckoos Nest Casablanca Citizen Kane Amadeus Shawshank Redemption, The Silence of the Lambs, The Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb
Level.4	Field of Dreams Rain Man Good Will Hunting Few Good Men, A Jerry Maguire Apollo 13 Edward Scissorhands Boogie Nights Witness Truman Show, The	Patriot, The Gladiator Shanghai Noon Perfect Storm, The Legends of the Fall World Is Not Enough, The Matrix, The Twister Lethal Weapon Green Mile, The	Jerry Maguire Good Will Hunting Dances with Wolves Crying Game, The Forrest Gump Leaving Las Vegas Boogie Nights Whats Eating Gilbert Grape Edward Scissorhands Chasing Amy
Level.5	Fly, The Thomas Crown Affair, The Poltergeist Shining, The Rocky Horror Picture Show, The Cape Fear Jurassic Park Back to the Future Blues Brothers, The Scream	Babe Contact Apocalypse Now Hamlet Truman Show, The Whole Nine Yards, The Good Will Hunting What Lies Beneath Driving Miss Daisy Braveheart	Almost Famous Best in Show Gladiator High Fidelity Cape Fear Erin Brockovich Chicken Run Pulp Fiction Run Lola Run (Lola rennt) Talented Mr. Ripley, The
Level.6	Requiem for a Dream Pulp Fiction Aladdin Almost Famous Lion King, The Sound of Music, The Beauty and the Beast Jungle Book, The Mary Poppins Fantasia	River Runs Through It, A Mr. Hollands Opus Rocky Crimson Tide Dances with Wolves U-571 Powder Thelma Malcolm X Full Monty, The	Big Lebowski, The Field of Dreams Few Good Men, A Lion King, The Toy Story Rain Man Mummy, The Louise & Crimes and Misdemeanors Producers, The Sleepy Hollow
Level.7	Talented Mr. Ripley, The Alice in Wonderland Cinderella Mulan Wizard of Oz, The One Flew Over the Cuckoos Nest Elephant Man, The Bambi Snow White and the Seven Dwarfs Pinocchio	Big Lebowski, The Hurricane, The Happy Gilmore Dumb & Dumber Mystery Men Men in Black Godfather: Part II, The Analyze This True Romance Sleepless in Seattle	Young Frankenstein West Side Story Austin Powers: International Man of Mystery Shining, The Poltergeist Full Monty, The Airplane! Psycho Ghostbusters American Werewolf in London, An
Level.8	Gladiator Almost Famous Pale Rider Best in Show Unforgiven Monty Python and the Holy Grail Great Escape, The Tombstone Good Will Hunting Meet the Parents	Austin Powers: The Spy Who Shagged Me Ace Ventura: Pet Detective Contender, The Sound of Music, The Midnight in the Garden of Good and Evil Graduate, The Ghostbusters Abyss, The Being There Full Metal Jacket	Sound of Music, The Hard Days Night, A Total Recall Sixth Sense, The No Way Out Dirty Dancing Rocky Untouchables, The Meet the Parents Dr. No