

Knowledge Graph Attention Networkに基づく 購買行動分析モデルに関する一考察

伊藤 史世^{1,a)} 張 志穎^{1,b)} 雲居 玄道^{1,c)} 後藤 正幸^{1,d)}

受付日 2021年4月29日, 採録日 2021年10月8日

概要: 近年, ユーザの購買行動に関するデータは, 購買履歴にとどまらず, 購買アイテムに紐づくカテゴリや店舗情報といった情報 (以下, 補助情報) や, ユーザに関係する性別や年代などの情報 (以下, 属性情報) といったように, 様々な種類のデータが取得できるようになってきている. そのため, これらを統合的に分析し, マーケティング施策へ活かすことへの期待値は高い. このようなモデルとして, ユーザと購買アイテム, ならびに補助情報との関係性をモデル化することでユーザの嗜好を学習する Knowledge Graph Attention Network (以下, KGAT) が提案されている. このモデルでは, アイテムの補助情報を用いることでユーザの嗜好に解釈を与えることができ, この解釈性はマーケティング施策の立案においても有用となりうる. しかし, KGAT ではユーザの属性情報については考慮されておらず, ユーザの購買行動を属性から明らかにする分析を行うことが困難であった. そこで本研究では, アイテムの補助情報を活用した KGAT を, さらにユーザとその属性情報の関係性も考慮したモデルへと拡張することで, ユーザの嗜好を属性と購買アイテム, さらにアイテムの補助情報を用いて分析することが可能なモデルを提案する. また, 実際の EC サイトの評価履歴データに提案手法を適用し, 提案手法の有用性を示す.

キーワード: 知識グラフ, ディープラーニング, 顧客分析, グラフ畳み込みネットワーク, Graph Attention Network

A Study on Analysis Model of Customers' Purchasing Behavior Based on Knowledge Graph Attention Network

FUMIYO ITO^{1,a)} ZHIYING ZHANG^{1,b)} GENDO KUMOI^{1,c)} MASAYUKI GOTO^{1,d)}

Received: April 29, 2021, Accepted: October 8, 2021

Abstract: Recently, it has become possible to make use of not only simple purchase history but also acquire various types of information. For example, we can get the information related to items (auxiliary information) and the information related to users (attribute information). Therefore, it is expected to analyze these various kinds of data in order to investigate customers' purchasing behavior for marketing purposes. For that reason, the KGAT model, which learns user preferences by modeling the relationship between users, purchase items and their auxiliary information, has been proposed. In this model, the users' preference can be interpreted by using the auxiliary information of items and this interpretability can be useful for planning marketing policies. But, the users' attribute information cannot be considered in KGAT, so we cannot interpret users' purchase behavior in the viewpoint of users' attribute information. Therefore, this research proposes a model that enables more diversified analysis by extending KGAT by using not only auxiliary information of items but also the relationship between users and their attribute information. Finally, we apply the proposed method to the evaluation of historical data of actual EC sites and show the usefulness of the proposed method.

Keywords: knowledge graph, deep learning, customer analysis, graph convolutional network, graph attention network

1. はじめに

近年、マーケティング分析に活用されるユーザの購買行動に関するデータには、単なる購買履歴のみでなく、購買アイテムに紐づくカテゴリや店舗情報といった情報（以下、補助情報）や、ユーザに関する性別や年代などの情報（以下、属性情報）といったように、様々な種類のデータを含めることが可能である。そのため、これらのデータを統合的に分析し、ユーザの嗜好について解釈可能な手法の重要性が高まっている。

従来、購買履歴データからユーザの嗜好を分析する手法としてトピックモデル [1] に代表される潜在クラスモデルが多く活用されてきた。潜在クラスモデルとは、観測データの背後に潜在的な変数の存在を仮定したモデルである。購買履歴データに適用し、得られたトピックを解釈することによりユーザの嗜好を分析することが可能であるが、モデルの学習によって得られるトピックの解釈は分析者の主観に基づくものであり、客観的、定量的な評価ができないという問題点が存在する。これに対し、ユーザの購買やサイトの閲覧といった行動履歴に含まれるアイテムやウェブサイトを埋め込み空間上のベクトルとして表現し、それらの関係性を分析する手法も多く提案されている [2], [3]。しかし、これらの手法においては、ユーザの行動系列のみからモデルの学習が行われ、ユーザやアイテムに紐づいている補助情報をモデルに考慮することができない。さらに、ユーザの嗜好は購買アイテムや閲覧サイトの類似性に基づいた分析からしか解釈することができず、ユーザ属性やアイテムの補助情報との関連から多角的な分析を行うことは困難である。

一方、推薦システムの分野においては、ユーザの購買履歴に含まれる様々なデータを統合的に扱いユーザの購買行動を予測するためのモデルが多く提案されている。そのようなモデルの1つに、ユーザと購買アイテム、ならびに補助情報との関係性をモデル化することでユーザの嗜好を学習する Knowledge Graph Attention Network（以下、KGAT） [4] がある。このモデルは、ユーザと購買アイテム、さらにアイテムの補助情報の関係をグラフで表現することによって、それらの間の複雑な関係性を学習することができる。さらに、モデルの学習によってグラフ上の「ユーザと購買アイテム」「アイテムと販売店舗」といった関係性の重要度をエッジの重みとして定量化することが可能である。この重みの大きさを分析することにより、たとえば「過去に購入した店舗と同じアイテムであるから」、「過去

に同じカテゴリのアイテムを購入しているから」といったように、なぜそのアイテムをユーザに推薦するのか、すなわち推薦理由を示すことができる。このエッジの重みは顧客の購買行動を定量的に解釈するために非常に有用であり、これによって、ユーザの嗜好をアイテムの補助情報との関係性などから定量的に解釈可能になることが期待される。そこで本研究では、推薦システムのためのモデルとして提案された KGAT モデルを、マーケティング施策立案のための顧客分析に活用することを考える。しかし、従来の KGAT モデルにおいてはアイテムに関する情報（補助情報）の活用のみが焦点が当てられ、ユーザの属性情報については考慮されておらず、ユーザの購買行動を属性情報の観点から分析することは困難であった。しかし、マーケティングの分野においては、ユーザ属性に基づいて施策を検討および実施するセグメントマーケティングが行われることも多く、ユーザの属性情報が重要な観点である。

以上により、本研究では、推薦システムで用いられていた KGAT をベースとし、ユーザの属性情報を考慮した顧客分析も可能とする拡張型 KGAT ベース分析モデルを提案する。さらに、提案モデルの学習によって得られるエッジの重みを活用した分析方法についても提案を行う。本提案手法により、ユーザの購買行動について、ユーザの属性情報、購買アイテム、アイテムの補助情報を用いて多角的、かつ定量的に分析を行うことが可能となる。最後に、実際の EC サイトの評価履歴データに提案手法を適用し、提案手法の有用性を示す。

2. 関連研究

2.1 行動履歴データを対象とした顧客分析

顧客の嗜好は多様であり、異なる好みや購買行動を有したグループが存在することから、潜在クラスモデルは広く活用されている。潜在クラスモデルとは、観測されたデータの背後に観測不能な潜在的変数の存在を仮定したモデルであり、その中でも特に、文書と単語の共起関係を表現したモデルはトピックモデルと呼ばれる [1]。潜在クラスモデルは、非常に汎用性の高い手法であり、文書データ以外に様々なユーザの購買履歴データへの応用されており、ユーザの行動履歴の背後にある潜在的な情報、つまり潜在トピックの抽出を可能としている [5], [6], [7]。しかしながら、これらの手法によって得られた各トピックの解釈は分析者の主観に基づくものであり、客観的な評価が難しい場合もある。また、これらの手法においては、対象とするデータに応じて異なる生成過程を仮定する必要がある場合も多く、データの変更に対して柔軟でないという問題点も存在する。

一方で、近年ではユーザやアイテムの潜在的な特徴を低次元空間のベクトルとして表現する手法が多く提案されている。このようなアプローチの代表的なモデルとし

¹ 早稲田大学
Waseda University, Shinjuku, Tokyo 169-8555, Japan
a) fumiyo0607@fuji.waseda.jp
b) choshiei@akane.waseda.jp
c) moto-aries@ruri.waseda.jp
d) masagoto@waseda.jp

て、ユーザの購買履歴に含まれるアイテムの関係性を学習することが可能な Item2Vec がある [2]. このモデルでは、Word2Vec [8] を応用し、各アイテムの共起関係から各アイテムを潜在意味空間上のベクトルとして表現する。これにより、アイテム間の関係性のベクトル演算による表現やアイテムどうしの類似性を評価することが可能となる。ここで、Word2Vec とは、自然言語データを対象として単語間の関係性をモデル化する手法であり、ある文中に出現する単語は同じ文脈の単語から予測することが可能であるという仮定に基づいている。この手法では、ニューラルネットワークを用いて各単語のベクトル表現を学習する。Item2Vec のほかに、ユーザの購買以外の行動履歴データに Word2Vec のアプローチを適用した研究として、保坂らの研究 [3] がある。保坂らは、ユーザのサイトの閲覧履歴データに対して、Word2Vec のアプローチを援用し、サイト間の関係性をモデル化することによって、ユーザのサイト閲覧傾向の分析を可能にしている。しかし、サイトの閲覧傾向には、ユーザの属性やサイトのカテゴリなどが影響を与えると考えられるが、これらの手法ではサイトやユーザに関する補助的な情報を考慮することができない。また、ユーザの閲覧傾向に関する分析は、サイトどうしの類似度に基づいた議論にとどまり、多角的な観点から分析を行うことが困難である。

2.2 アイテムの補助情報を活用した推薦モデル

近年では、アイテムに関する様々な情報を活用することが可能なため、推薦システムの分野においても、アイテムの補助情報を活用するモデルが多く提案されている。その中の代表的な推薦モデルとして、Factorization Machines (以下、FM) [9] が知られている。この手法では、アイテムに関する補助情報を説明変数に含めることができ、それらの間の交互作用をモデル化することが可能である。汎用性の高い手法であることから広く用いられており、近年ではこの手法にディープラーニングを適用した手法 [10], [11] も提案されている。一方で、これらの手法では事前に定義されているデータ間の関係性については考慮することができないことから、近年では、アイテムに関する補助情報をグラフ構造で表現する手法が提案されている [4], [12], [13]. これらの手法は、非常に高い推薦精度を誇るとともに、アイテムの補助情報を表現したグラフを活用することで、推薦理由に対して解釈を与えることが可能であり、精度の高さと解釈性を両立させるための手法として注目を集めている。従来手法である KGAT [4] も、購買アイテムに関する補助情報を知識として活用することによって、精度と推薦理由の解釈性の向上させることに成功している。

2.3 Attention 機構を活用した解釈性の向上

Attention 機構は、自然言語処理の翻訳タスクの分野に

おいて、Vaswani ら [14] によって提案され、単語の訳出を行う際に、文脈に含まれる重要な単語に注目 (Attention) するという行動をモデル化したものである。各単語の注目度合い (Attention Weight) は、単純な行列演算によって定式化され、従来の手法と比較して計算量を抑えながら飛躍的に精度を向上させたとして非常に注目を集めた。以来、Attention 機構は自然言語処理以外の様々な分野でも応用がなされている。推薦システムの分野においては Attention 機構の導入により、精度を向上させるとともに、Attention Weight を活用することによって、その推薦理由について解釈を与えるという試みがなされている。

Hu ら [15] は、ユーザ、アイテム、さらにアイテムの補助情報を含む異種グラフを入力としてモデルの学習によってノードのベクトル表現を獲得し、このベクトル表現によってユーザの各アイテムの購買確率を出力するモデルを提案している。この手法では、定義されたメタパスに対する Attention Weight を用いて推薦理由に対して解釈を与えることが可能である。ここで、メタパスとは2つの異なる実体を接続する複合関係であり、映画の視聴履歴の場合 (Movie, Actor, Movie) といったメタパスが考えられ、この場合、メタパス中の2つの映画には同じ俳優 (Actor) が出演しているといった情報を表現することができる。また、(User, Movie, User, Movie) などのように、より高次の複合関係を定義することも可能である。Hu らの手法においては、定義したメタパスに対して畳み込み演算を適用し、1つのベクトルに集約したのちにユーザのベクトルに対して Attention Weight を算出することによって、各ユーザにおける各メタパスの重要度を定量化することが可能である。しかしながら、この手法においてはメタパスの設計にモデルの推薦精度が依存するため、異種グラフに含まれる実体の種類数が多い場合には、それらの組合せを考慮し、適切なメタパスを設計および選択することが困難であるという問題点がある。また、重要度の解釈はメタパスレベルにとどまり、グラフ上の2つのノードの直接的な関係性の重要度について分析することができない。

また、Seo ら [16] は、EC サイトのレビューのテキストから各アイテムの購買確率を予測するモデルを提案しており、Attention 機構を導入することによって従来のモデルと比較して評価値予測のタスクにおいて良い精度を示している。この手法を用いることにより、ユーザのレビューの文書中のどの単語がユーザの購買において重要であるかを Attention Weight を活用することによって解析することが可能である。しかし、膨大な文書データと文書に含まれる単語の重要度を分析することは実務上困難であると考えられ、マーケティング施策への応用などに対してこの手法は適切でないと考えられる。

KGAT [4] は、アイテムの補助情報を知識グラフとして表現し、この知識を活用して推薦を行うモデルである。こ

の手法においては知識グラフ上のノードのベクトル表現を獲得する際の畳み込み演算に Attention 機構を導入しており、モデルの学習によって重み付き有向グラフを得ることができる。そして、このエッジの重みを活用することによって推薦理由について解釈を与えることが可能である。この研究では、個別の推薦に対する解釈の方法については提案がなされているが、顧客分析への応用などについては方法論が確立されていない。

いずれの手法においても、研究の主眼はあくまで推薦精度の向上であり、従来 Attention Weight を活用した分析に主眼が置かれた研究は行われてきていない。

2.4 知識グラフの埋め込みに関する研究

知識グラフ (Knowledge Graph) とは、現実世界に定義されている関係性をグラフ構造で表現するものである。知識グラフでは、エンティティ集合を \mathcal{Q} 、リレーション集合を \mathcal{R} とし、それぞれのノードがエンティティ $q \in \mathcal{Q}$ 、2つのエッジの間に付与された辺ラベルがリレーション $r \in \mathcal{R}$ (関係性) を表し、有向グラフの形で表現される。また、知識グラフは、「先頭エンティティ (Head entity)」、「リレーション (Relation)」、「末尾エンティティ (Tail entity)」の3つ組であるトリプルの集合 $\mathcal{G} = \{(h, r, t) | h, t \in \mathcal{Q}, r \in \mathcal{R}\}$ として定義される。このトリプルによって、エンティティ間の関係性を表現する。

そして、知識グラフ上のエンティティやリレーションのベクトル表現を獲得するために、様々な手法が提案されている。そのうちの1つに、先頭エンティティと末尾エンティティの間に任意の変換操作 (Translation) を定義することにより、これらのエンティティの関係性をモデル化する手法として TransE [17] をはじめとする Translation-based Model が知られている。

2.4.1 TransE

TransE [17] においては、知識グラフ上で (h, r, t) が定義されている場合、末尾エンティティ t が、先頭エンティティ h におけるリレーション r による変換後のベクトル表現の近傍に位置するようにモデルの学習を行う。すなわち、知識グラフ上のトリプル (h, r, t) について、 $e_h + e_r \approx e_t$ という近似式を満たすように、各エンティティおよび各リレーションを意味空間上のベクトルとして表現する。ベクトル表現がどのくらい事実関係に従っているのかを測るためのスコア関数 $g(h, r, t)$ として、通常 $e_h + e_r - e_t$ の L_1 ノルムまたは L_2 ノルムが用いられ、 L_2 ノルムを用いた場合のスコア関数は式 (1) のように定義される。

$$g_e(h, r, t) = \|e_h + e_r - e_t\|_2 \quad (1)$$

ただし、 $\|\cdot\|_2$ は L_2 ノルムを表し、 $g_e(h, r, t) = 0$ であることは、知識グラフで定義されているトリプルに基づく事実関係を正確に表していることを意味する。TransE で

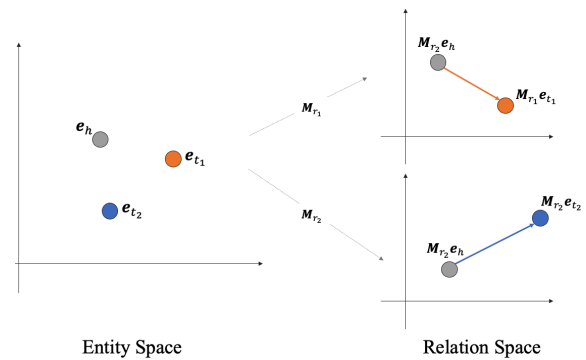


図 1 TransR の概念図

Fig. 1 Conceptual diagram of TransR.

は、リレーションを先頭エンティティと末尾エンティティの変換と見なしてそれらのベクトル表現を学習し、すべてのエンティティとリレーションのベクトル表現を同一の埋め込み空間上で表現する。しかし、実際にはエンティティは様々な意味的な側面を持ち、それらの意味の違いをリレーションによって表現する必要があると考えられる。この手法では、エンティティの持つ意味の多面性を表現することができないという問題点が存在する。

2.4.2 TransR

TransR [18] では、リレーションごとに異なる埋め込み空間を定義し、エンティティをこれらの異なる意味空間上に写像することで、リレーションの違いを考慮したベクトル表現を獲得することが可能である。ここで、TransR の概念図を図 1 に示す。TransR では、エンティティとリレーションを Entity space と Relation space という異なる埋め込み空間上でモデル化し、先頭エンティティと末尾エンティティの変換を Relation space 上で行う。知識グラフ上のトリプル (h, r, t) について、エンティティを $e_h, e_t \in \mathbb{R}^k$ 、リレーションを $e_r \in \mathbb{R}^d$ としてそれぞれ埋め込む空間上で表現する。それぞれのリレーションについて、写像行列を $M_r \in \mathbb{R}^{k \times d}$ として、この行列を用いてエンティティを Relation space へ写像し、写像先の空間において変換後の先頭エンティティと末尾エンティティが近傍に位置するように学習が行われる。TransR のスコア関数は式 (2) によって定義される。

$$g_r(h, r, t) = \|e_h M_r + e_r - e_t M_r\|_2 \quad (2)$$

3. 従来手法 (KGAT)

3.1 概要

KGAT [4] は、ユーザと購買アイテム、購買アイテムの補助情報からなるグラフを入力として、ユーザの各アイテムに対する購買確率を出力する推薦モデルである。このモデルでは、グラフ上の高次の関係性をとらえながら、推薦において重要である関係を学習し、ユーザやアイテム、アイテムの補助情報をベクトルで表現する。KGAT を学習

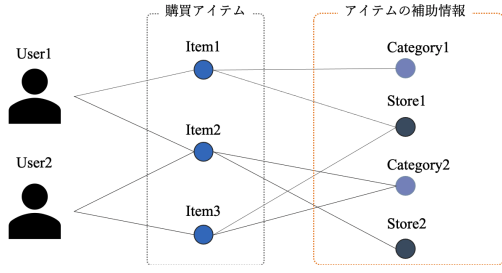


図 2 入力とするグラフのイメージ
Fig. 2 Image of graph for input.

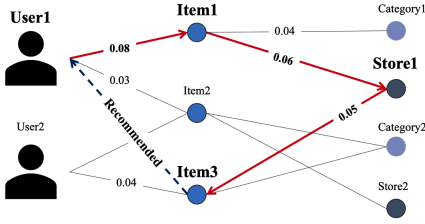


図 3 学習済み有向グラフを用いた推薦理由の解釈

Fig. 3 Interpretation of the recommendation reason using the trained directed graph.

させることによって重み付き有向グラフを得ることができ、このグラフを用いることによって各ユーザに対する各アイテムの推薦理由について解釈することができる。具体的には、ユーザから推薦アイテムまで接続されているエッジのうち、エッジの重みが最大なものを選択し、これを推薦パスとする。この推薦パスに含まれている補助情報を分析することによって、推薦理由の合理性を考察することが可能となる。たとえば、図 3 に示した例の場合、User1 に Item3 が推薦された理由は、推薦パスより、「過去に購入した Item1 と同じ店舗 Store1 で販売されているアイテムであるから」という解釈が可能である。

3.2 モデル式

KGAT ではエンティティとリレーションをベクトルで表現する。エンティティまたはリレーション $p \in \mathcal{Q} \cup \mathcal{R}$ のベクトル表現を $e_p \in \mathbb{R}^d$ とし、TransR [18] に基づき、知識グラフ上のエンティティのベクトル表現 e_h, e_t を、リレーションごとに異なる意味空間上に写像する。ここで、リレーション $r \in \mathcal{R}$ による写像行列を $\mathbf{W}_r \in \mathbb{R}^{d \times d}$ としたとき、グラフのスコア関数 g は、式 (3) として定義される。

$$g(h, r, t) = \|\mathbf{W}_r e_h + e_r - \mathbf{W}_r e_t\|_2 \quad (3)$$

また、グラフ構造に対する損失は、式 (4) で表される。

$$Loss_{KG} = \sum_{(h,r,t,t') \in \mathcal{T}} -\ln \sigma(g(h, r, t') - g(h, r, t)) \quad (4)$$

ただし、 σ をシグモイド関数、 \mathcal{T}_r をリレーション r について定義されている末尾エンティティの集合として、 $\mathcal{T} = \{(h, r, t, t') \mid t, t' \in \mathcal{T}_r, (h, r, t) \in \mathcal{G}, (h, r, t') \notin \mathcal{G}\}$ であ

る。ここで、 (h, r, t') は、グラフ上のトリプル $(h, r, t) \in \mathcal{G}$ について $t' \in \mathcal{T}_r$ を t 以外のエンティティからランダムにサンプリングすることによって作られるトリプルである。さらに、グラフに存在するトリプル (h, r, t) のエッジの重み $\pi(h, r, t)$ を以下の式で求める。

$$\tilde{\pi}(h, r, t) = (\mathbf{W}_r e_t) \tanh(\mathbf{W}_r e_h + e_r)^\top \quad (5)$$

$$\pi(h, r, t) = \frac{\exp(\tilde{\pi}(h, r, t))}{\sum_{(h,r',t') \in \mathcal{N}_h} \exp(\tilde{\pi}(h, r', t'))} \quad (6)$$

ただし、 \mathcal{N}_h は先頭エンティティ h に対してグラフ上で定義されているトリプルの集合であり、 \top は転置を表す。また、 $\tanh(x)$ を $x = (x_1, x_2, \dots, x_d)$ に対して $\tanh(x) = (\tanh(x_1), \tanh(x_2), \dots, \tanh(x_d))$ と定義する。

続いて、 $\pi(h, r, t)$ により、 e_h に対する周辺のエンティティのベクトル表現の重み付け和を計算し、その表現を自身に取り入れることを L 回繰り返すことにより、グラフ上の L 次近傍の関係性を考慮したベクトル表現を獲得する。

$$e_{\mathcal{N}_h} = \sum_{(h,r,t) \in \mathcal{N}_h} \pi(h, r, t) e_t \quad (7)$$

この繰り返し操作は、 l 回目の量込み演算を終えた後のノード h のベクトル表現を $e_h^{(l)} \in \mathbb{R}^{d^{(l)}}$ として、以下のよう定式化される。

$$e_{\mathcal{N}_h}^{(l-1)} = \sum_{(h,r,t) \in \mathcal{N}_h} \pi(h, r, t) e_t^{(l-1)} \quad (8)$$

$$e_h^{(l)} = f_{GA}^{(l)}(e_h^{(l-1)}, e_{\mathcal{N}_h}^{(l-1)}) \quad (9)$$

ただし、 $f_{GA}(\cdot)$ は任意の非線形関数である。

式 (8)–(9) の処理を L 回繰り返すことで、 L 次近傍までの周辺ノードの接続情報を考慮することができる。このとき、 e_h は $\pi(h, r, t)$ の大きい e_t に依存した表現となる。そのため、知識グラフで関係性が定義されている 2 つのエンティティの関係性の強さをエッジの重みとして定量化することが可能となる。

最後に、ユーザ $u \in \mathcal{U}$ とアイテム $i \in \mathcal{I}$ の推薦スコア $\hat{y}(u, i)$ を、更新した L 個のベクトル表現を用いて以下のよう求める。このスコアが高いほど、ユーザ u とアイテム i の共起関係が強いことを表す。ここで、 $\mathbf{a}|\mathbf{b}$ は、ベクトル \mathbf{a}, \mathbf{b} の連結を表す。

$$e_u^* = e_u^0 | e_u^1 | \dots | e_u^L, \quad e_i^* = e_i^0 | e_i^1 | \dots | e_i^L \quad (10)$$

$$\hat{y}(u, i) = e_u^* e_i^{*\top} \quad (11)$$

ユーザとアイテムの共起に関する損失は、式 (12) で表される。

$$Loss_{CF} = \sum_{(u,i,j) \in \mathcal{O}} -\ln \sigma(\hat{y}(u, i) - \hat{y}(u, j)) \quad (12)$$

ここで、 \mathcal{D}^+ を実際にデータ中に観測されたユーザとアイ

テムのペア集合, D^- を観測されていないのペア集合とし, $\mathcal{O} = \{(u, i, j) \mid (u, i) \in D^+, (u, j) \in D^-\}$ である.

3.3 モデルの最適化

モデルの全体の損失 $Loss$ は知識グラフの損失 $Loss_{KG}$ と推薦の損失 $Loss_{CF}$ を用いて以下のように定義される.

$$Loss = Loss_{KG} + Loss_{CF} + \lambda \|\Theta\|_2 \quad (13)$$

ここで, $\Theta = \{e_p, \mathbf{W}_r, \mathbf{W}_{GA}^{(l)} \mid \forall p \in \mathcal{Q} \cup \mathcal{R}, \forall r \in \mathcal{R}, \forall l \in \{1 \dots, L\}\}$ とし, $\mathbf{W}_{GA}^{(l)}$ は非線形関数 f_{GA} におけるパラメータ, $\lambda > 0$ は正則化の強度を決定する正則化パラメータである. Adam [19] に基づいて, 式 (13) を最小化するように各パラメータを更新する.

4. 提案手法

4.1 概要

提案手法では, 従来の KGAT をユーザの属性情報を考慮したモデルへと拡張する. ここでは, ユーザ, ユーザの属性, アイテム, そしてアイテムの補助情報からなるグラフを構築し, これをモデルの入力とし学習を行い, これにより得られたエッジの重みを用いた分析方法を提案する. 提案手法により, ユーザとその属性, アイテムとその補助情報間の関係性の強さを定量化することができ, それらの関係性について分析を行うことが可能となる.

4.2 グラフの構築

従来研究においては, FreeBase や YOGO といった膨大な知識を含んだ知識グラフをモデルの入力に用いることもある [4], [20] が, この知識グラフには分析者にとって有用でない情報も含まれていることが考えられる. そのため, 本研究においては学習の効率性の観点から独自に知識グラフを構築し, これをモデルの入力に用いることとする. ここで, グラフ上のエンティティを $q \in \mathcal{Q}$, リレーションを $r \in \mathcal{R}$ とする. このエンティティとして提案手法では, ユーザ集合 \mathcal{U} , アイテム集合 \mathcal{I} , アイテムの補助情報として店舗集合 \mathcal{A}_1 , カテゴリ集合 \mathcal{A}_2 に加えてユーザの属性集合 \mathcal{A}_3 も考慮する. すると, エンティティ集合は, 式 (14) として定義される.

$$\mathcal{Q} = \mathcal{U} \cup \mathcal{I} \cup \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3 \quad (14)$$

また, グラフ \mathcal{G} は, 「先頭エンティティ」, 「リレーション」, 「末尾エンティティ」の3つの組であるトリプル (h, r, t) の集合として, 式 (15) のように定義される.

$$\mathcal{G} = \{(h, r, t) \mid h \in \mathcal{Q}, r \in \mathcal{R}, t \in \mathcal{Q}\} \quad (15)$$

ここで, 最終的に構築されるグラフのイメージを図 4 に示す.

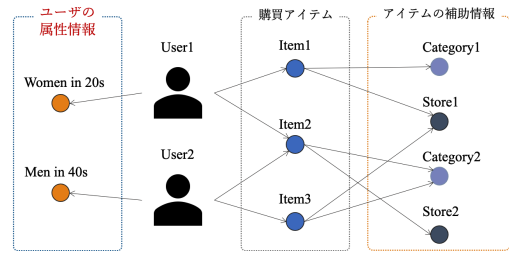


図 4 提案手法により構築されるグラフ

Fig. 4 The graph created by the proposal method.

4.3 提案モデルを用いた分析方法

提案手法では KGAT と同様に, グラフ上に定義されている関係性を正確に表現しているかを表す知識グラフの損失 $Loss_{KG}$ とユーザの購買行動をどの程度正確に表現できているのかを表す損失 $Loss_{CF}$ を同時に最適化することにより, モデルの学習を行う. これにより, 実際にグラフで定義されている関係のうち, ユーザの購買行動を表現するうえで重要である関係が式 (6) によって算出される正規化後のエッジの重み π により強調される. この値を用いることで, ユーザの購買行動に関する分析が可能となる. 提案手法で構築されるグラフにおいては, ユーザは「購買アイテム」と「属性」に接続されており, アイテムは, 「アイテムを購入したユーザ」, 「販売店舗」および「アイテムカテゴリ」に接続されている. これを用いて, グラフとして表現されたユーザの購買行動において重要な関係性を検出することが可能となる. さらに, 年代・性別からなるユーザ属性, アイテム, また店舗について分析を行うために, ある末尾エンティティに対して定義されている先頭エンティティのエッジの重みの平均値を算出することによって, 末尾エンティティの「性質」について分析する方法を提案する. 具体的には, 末尾エンティティ t に対して定義されている先頭エンティティの集合を \mathcal{H}_t , ある任意のエンティティ集合を \mathcal{A} とし, 以下のように平均値を算出する.

$$\mu(\mathcal{A}, r, t) = \frac{\sum_{h' \in \mathcal{A} \cap \mathcal{H}_t} \pi(h', r, t)}{|\mathcal{A} \cap \mathcal{H}_t|} \quad (16)$$

ここで, $\pi(h', r, t)$ が大きいエンティティ h' は末尾エンティティ t に依存した表現になっている. そのため, この依存度が末尾エンティティに対して平均的に大きいということは, 末尾エンティティ t は, $h' \in \mathcal{H}_t$ にとって重要な関係にあると解釈することができる. この値を用いることによって, 顧客にとって重要なアイテムや, 販売アイテムとの関連度が高い店舗を分析することが可能となる.

5. 実データを用いた分析

提案手法の有用性を示すため, 国立情報学研究所の IDR データセット提供サービスより提供を受けた 2012 年の楽天市場の評価履歴データ [21] に対して提案手法を適用し, 分析を行う.

表 1 処理後のデータセット基本情報

Table 1 Basic information of data set after processing.

ユーザ数	27,641
アイテム数	123,751
店舗数	13,123
カテゴリ数	13,658
定義されているトリプルの数	275,143

5.1 分析条件

データ期間は、2012年の1年間であり、デモグラフィック属性として、年代が10代から70代までの7段階、性別が男女の2属性の計 $7 \times 2 = 14$ カテゴリを使用した。また、1) 被購買件数が3件未満のアイテム、2) 被購買件数が25件未満の店舗、3) 購買数が3件未満のユーザ、という条件に該当するデータを削除後、各属性から2,000人ずつサンプリングした。ただし、70代女性に関しては、上記1)~3)の処理後のユーザ数が1,641人であったため、サンプリングは行わなかった。ここで、処理後の対象データセットの基本情報を以下の表1に示す。

また、Wangら[4]が提供しているパラメータの設定*1に則り*2、グラフの埋め込み表現の次元数は $(d^{(0)}, d^{(1)}, d^{(2)}) = (64, 32, 16)$ 、正則化の強度を調節する正則化パラメータは $\lambda = 0.0001$ 、LeakyReLUのハイパーパラメータは $\alpha = 0.2$ とし、グラフの更新回数は $L = 2$ とした。さらに、エポック数は300として学習を行った。

5.2 実験環境

モデルの学習は、NVIDIA Tesla T4 GPUを1個、vCPUを4個搭載したマシンで実行した。提案手法の学習においては、1エポックあたり約35秒を要した。

5.3 分析結果と考察

5.3.1 提案モデルの予測性能

提案モデルの妥当性を検証するため、同様のデータを取り扱い、ユーザの購買予測を行うことのできるモデルとの予測精度の比較を行う。ここでは、比較手法として、ユーザの属性情報を用いない従来手法(KGAT)と、ニューラルネットワークの下でFMを包含する因子分解モデルであるNeural Factorization Machines (以下、Neural FM) [10]を用いた。なお、Neural FMにおいては、モデル化による差異のみを明らかにするためにユーザの属性情報も使用した。また本研究では、評価指標としてNDCG@20, NDCG@100を使用した。ここで、各手法におけるNDCGの比較を以下の表2に示す。表2より、提案手法は従来手法(KGAT)

表 2 NDCG の比較

Table 2 Comparison of NDCG.

	NDCG@20	NDCG@100
従来手法 (KGAT)	0.0233	0.0364
Neural FM	0.0340	0.0424
提案手法	0.0375	0.0562

とNeural FMよりも良い性能を示しており、ユーザの属性情報を活用し、それらを含むデータをグラフとして表現することによって、ユーザの嗜好をより正確にとらえることができていていること分かる。

5.3.2 ユーザと属性に関する分析

各属性の嗜好の傾向に関する分析を行うために、ユーザ $u \in \mathcal{U}$ と属性 $a \in \mathcal{A}_3$ のエッジの重み $\pi(u, r, a)$ を用いて各属性の比較と考察を行う。ここで、 $\pi(u, r, a)$ が大きいことは、ユーザ u のベクトル表現が属性 a に依存した表現となっていることを意味しており、属性 a について $\mu(\mathcal{U}, r, a)$ の値が大きいことは、属性 a への依存が大きいユーザの割合が高いことを意味する。よって、 $\mu(\mathcal{U}, r, a)$ を比較することによって、属性ごとにどの程度属性に依存している、つまり個人の嗜好性が低いユーザが存在するのかを明らかにすることができる。ここで、年代・性別ごとの $\mu(\mathcal{U}, r, a)$ の値と、 $\pi(u, r, a)$ の分布を図5に示す。図5より、 $\pi(u, r, a)$ はいずれの属性においても0付近に多く分布しているが、属性ごとに分布の広がり方が異なることが分かる。まず性別で比較すると、男性の方が全体的に $\mu(\mathcal{U}, r, a)$ の値が低く、 $\pi(u, r, a)$ が0付近であるユーザの割合が高い傾向が見られる。このことから、男性は女性と比較してユーザ固有の嗜好性が強く、属性に依存しない購買を行うユーザが多いことが分かる。また、同一の性別における年代間の比較をすると、男性においては、 $\mu(\mathcal{U}, r, a)$ の値は30代、40代で大きくなり、年代が上昇するにつれてその値は小さくなっていく一方で、女性は、年代が上昇するにつれて大きくなる傾向がある。このことから、男性は50代以降に、ユーザ固有の嗜好を強く持つユーザの割合が多くなり、趣味などの個人の嗜好に基づく購買を行うユーザの割合が多くなることが推測される。反対に、女性においては、年代が上昇するとユーザ固有の嗜好を強く持つユーザが減少していくことが分かる。この結果を用いて、「 $\mu(\mathcal{U}, r, a)$ の値の大きい50代以降の女性、かつ $\pi(u, r, a)$ の大きいユーザには売れ筋商品のクーポンの発行など画一的な施策を行う、一方で、 $\mu(\mathcal{U}, r, a)$ の値の小さい10代、20代の男女には過去の購買履歴を活用し、パーソナライズした推薦を行う」といった、マーケティング施策に関する提案を行うことが可能である。

5.3.3 ユーザの嗜好に関する分析

エッジの重み π の有用性について検討するため、ユーザの1次近傍のエンティティに関して、ユーザと属性のエッ

*1 <https://github.com/xiangwang1223/knowledge-graph-attention-network>

*2 3つの異なるデータセット (Amazon-book, Last-FM, Yelp2018) において、最適なパラメータ設定として同値のパラメータ設定が提供されていることから、今回のデータセットにおいてもWangらの設定とした。

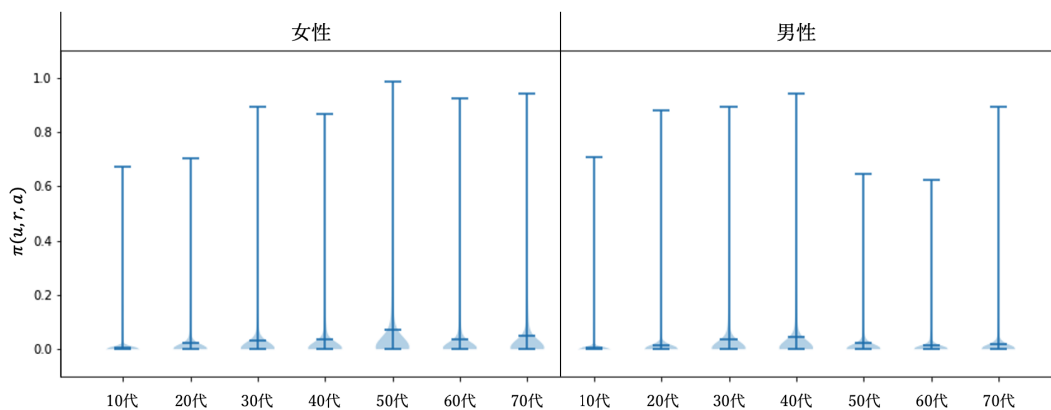


図 5 属性ごとの $\pi(u, r, a)$ の分布と $\mu(\mathcal{U}, r, a)$

Fig. 5 The distribution of $\pi(u, r, a)$ and the value of $\mu(\mathcal{U}, r, a)$ of every users' attribute.

表 3 u_1 の購買履歴

Table 3 Purchase history of u_1 .

エンティティ名	π
a (40代男性)	9.43×10^{-1}
生パスタセット	2.67×10^{-2}
座いす	1.09×10^{-2}
腕時計	8.01×10^{-3}
懐中電灯	6.82×10^{-3}
DVD プレーヤ	5.03×10^{-3}

表 4 u_2 の購買履歴

Table 4 Purchase history of u_2 .

エンティティ名	π
ペット用パーカ	9.99×10^{-1}
ペット用ボロシャツ	2.35×10^{-8}
ブランケット	7.40×10^{-9}
犬用サプリメント	2.42×10^{-9}
犬用デンタルケア	2.06×10^{-9}
a (40代女性)	1.64×10^{-13}

ジの重み $\pi(u, r, a)$ とユーザの購買アイテムのエッジの重み $\pi(u, r, i)$ について分析する. これらの π の値はユーザの 1 次近傍において正規化された値であるため, 合計すると 1.0 となる. すなわち, $\pi(u, r, a)$, $\pi(u, r, i)$ は各ユーザの嗜好に対する影響の割合であると解釈することができる. ここでは, 評価件数が 5 件のユーザのうち, $\pi(u, r, a)$ が最も高かったユーザ u_1 , および最も低かったユーザ u_2 を抽出し, その購買傾向について考察する. 該当するユーザの π の値を表 3, 表 4 に示す. ここで, 表 5 より, 一般に 40 代男性では, 食料品や電化製品などの日用品が多く購買されており, 表 3 のユーザ u_1 は, 40 代男性の標準的な購買行動をしているといえる. 一方で, $\pi(u, r, a)$ が最も低かった 40 代女性のユーザ u_2 は表 4 に示したようにペット商品を多く購入しており, 表 5 から読み取れるファッションアイテムや美容アイテムといった購買は見られない. このように, ユーザ u_2 では, ユーザ固有の嗜好に基づく一

表 5 被購買件数上位 5 アイテム

Table 5 The 5 items in the number of number of items purchased.

属性	40代男性	40代女性
1	飲料水	レギンス
2	シャンプー	ワンピース
3	乾電池	まつげ美容液
4	うどんセット	お米
5	SD カード	裏起毛レギンス

貫した購買行動が確認できる. このように, $\pi(u, r, a)$ の値が大きく属性との関連度の高いユーザは属性内において標準的な嗜好を有し, 反対に, この値が小さく属性との関連度の低いユーザは, ユーザ特有の嗜好を有すると解釈することができる.

5.3.4 アイテムの特徴に関する分析

アイテムの特徴を明らかにするため, ユーザ集合 \mathcal{U} とアイテム $i \in \mathcal{I}$ に対して式 (16) より $\mu(\mathcal{U}, r, i)$ を算出し分析を行う. $\pi(u, r, i)$ が大きいアイテム i は, ユーザとアイテム間の関係性が強く, ユーザの嗜好を表現するうえで重要なアイテムであることを意味する. 一方で, $\pi(u, r, i)$ が平均的に小さいアイテム i は, ユーザとアイテムの関係性が弱く, 購買ユーザに嗜好を表現するうえで重要なアイテムではないと解釈することができる. そのため, $\pi(u, r, i)$ を用いて $\mu(\mathcal{U}, r, i)$ を算出することによって, 各アイテムのユーザの嗜好に対する重要度を定量化することが可能となる. ここで, 被購買件数が 16 件以上, 25 件未満のアイテムのうち $\mu(\mathcal{U}, r, i)$ の上位 10 件と下位 10 件のアイテムを図 6 に示す. 図 6 より, ファッション系のアイテムにおいて $\mu(\mathcal{U}, r, i)$ の値が大きくなっており, $\pi(u, r, i)$ は 1.0 付近に多く分布している一方で, この値が小さいアイテムは, $\pi(u, r, i)$ がほとんど 0 付近に分布している. このように同程度の購買件数で同程度のアイテムであっても, ユーザの嗜好における重要度が異なることが分かる. ここで, あるアイテム i について, ユーザの嗜好における重要度 $\pi(u, r, i)$

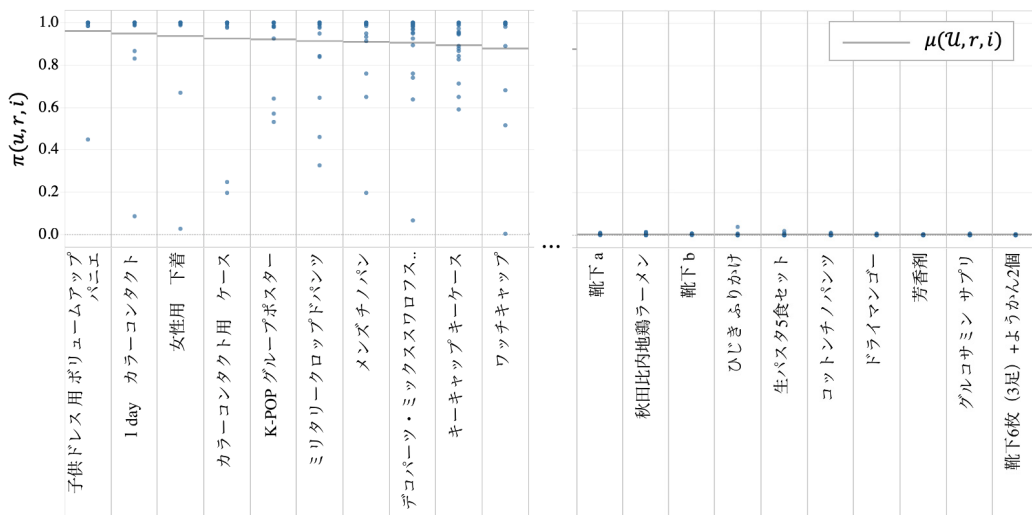


図 6 該当アイテムの $\pi(u, r, i)$ の分布と $\mu(U, r, i)$

Fig. 6 The value of $\mu(U, r, i)$ and distribution of $\pi(u, r, i)$ of selected items.

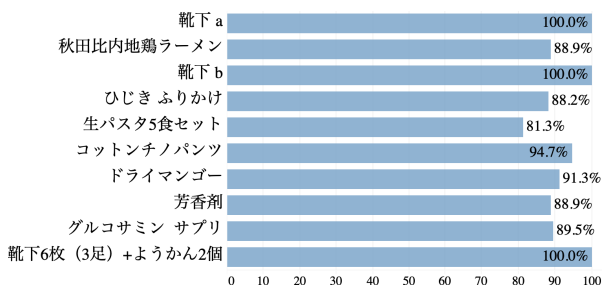


図 7 各該当アイテムの購買ユーザのうち $\pi(u, r, i')$ が最大であるユーザの割合

Fig. 7 Percentage of users for whom $\pi(u, r, i')$ is maximal in users purchasing each selected item.

が小さい要因として、アイテム i を購買しているユーザの嗜好における属性の重要度 $\pi(u, r, a)$ の値が大きい、もしくは、アイテム i 以外のアイテム i' に対する重要度 $\pi(u, r, i')$ の値が大きいことが考えられる。図 6 に示した $\mu(U, r, i)$ の下位 10 位のアイテムにおいて、 $\pi(u, r, i')$ の値が最大のユーザの割合を図 7 に示す。この結果より、アイテム i を購買しているユーザの嗜好において、アイテム i 以外のアイテム i' に対する重要度 $\pi(u, r, i')$ の値が大きいことが分かる。よって、 $\mu(U, r, i)$ の小さいアイテムは、ユーザの嗜好を説明するアイテムではないため、ユーザに宣伝対象となりうるアイテムでないと考えられる。一方で、このようなアイテムにおいてはセット売りなどの施策によって、販売数を増加させるなどのマーケティング戦略が考えられる。

5.3.5 アイテムと店舗に関する分析

アイテムと店舗の関係性を明らかにするため、アイテム集合 \mathcal{I} と店舗 $s \in \mathcal{A}_1$ に対して、式 (16) を適用し分析を行う。ここでは、取扱いアイテムの被購買件数の合計が 500 以上 1,000 未満の店舗を抽出した。ここで該当店舗について、縦軸に $\mu(\mathcal{I}, r, s)$ 、横軸に取扱いアイテム数をとった散

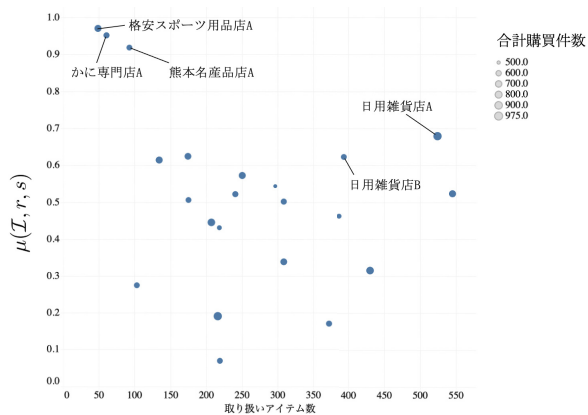


図 8 該当店舗の $\mu(\mathcal{I}, r, s)$ と取扱いアイテム数

Fig. 8 $\mu(\mathcal{I}, r, s)$ and the number of handling items.

布図を図 8 に示す。図 8 より、アイテム数が増加すると $\mu(\mathcal{I}, r, s)$ が小さくなっていることが分かる。アイテム数が多い店舗においては、一般的に様々なアイテムを取り扱っていることが想定されるため、 $\mu(\mathcal{I}, r, s)$ の大きい店舗においては、取扱いアイテムの多様性が低く、 $\mu(\mathcal{I}, r, s)$ の小さい店舗においては、取扱いアイテムの多様性が高いと考えられる。実際に、図 8 より、 $\mu(\mathcal{I}, r, s)$ の大きい店舗は、取扱いアイテム数が少なく、専門店などの特定の食材やカテゴリのアイテムを販売する店舗である。全体的には取扱いアイテム数が増加すると $\mu(\mathcal{I}, r, s)$ の値が小さくなるという傾向が見られるが、日用品専門店 A については、取扱いアイテム数が多いものの、 $\mu(\mathcal{I}, r, s)$ の値が比較的大きくなっている。ここで、日用雑貨店 A と日用雑貨店 B における各属性ごとの購買割合を図 9 に示す。これらの店舗は、同様に日用雑貨を取り扱っているが、日用雑貨店 A の方が取扱いアイテム数に対して $\mu(\mathcal{I}, r, s)$ の値が大きくなっている。図 9 より、日用品雑貨店 B は属性ごとの購買割合に顕著な傾向は見られないが、日用雑貨店 A は、70 代女

性における購買割合が非常に高く、特定の顧客層からの支持が高いアイテムを取り扱う店舗であることが分かる。このことから、顧客の嗜好においてアイテムが類似しているため、 $\mu(I, r, s)$ の値が大きくなったと考えられる。このように、 $\mu(I, r, s)$ の値の値を分析することによって、各店舗の取扱アイテム数の多様性について分析することが可能である。

5.3.6 エンティティのベクトル表現に関する分析

提案モデルでは関係性の強いエンティティ、すなわち $\pi(h, r, t)$ が大きいエンティティ h, t のベクトルどうしの表現が近くなるようなベクトルの更新方法をとる。さらに、各エンティティについてを L 近傍の周辺エンティティの情報を畳み込み演算により取り入れることで、グラフ上のエッジでつながれた L 次のエンティティについてその重みが大きいものどうしのベクトル表現が類似したものとなる。これによって、実際に事実関係が定義されているエンティティすべてのベクトル表現が類似したものではなく、関係性の重要度に基づいてそれらの類似性をモデリングすることが可能となる。

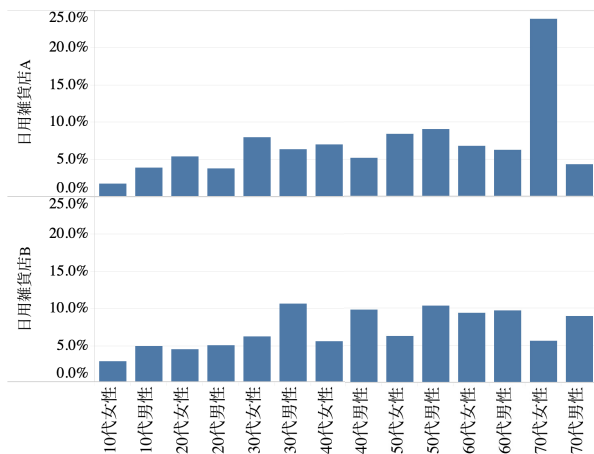


図 9 日用品雑貨店 A, 日用品雑貨店 B の属性ごとの購買割合の比較
 Fig. 9 Comparison of percentage of purchases of each attribute for Daily Necessities Store A and Daily Necessities Store B.

ここでは、エンティティのベクトル表現に関する分析を行うために、以下の2つのアイテムについて、Entity space においてそれらと Cos 類似度が最も大きかった上位 10 件のアイテムを算出し、考察を行う。ここで抽出した2つのアイテムの概要を表 6、該当アイテムと店舗の間のエッジの重みを表 7 に示し、さらに該当アイテムと Cos 類似度の大きかった上位 10 件のアイテムを表 8、表 9 に示す。

表 8 より、ニットカーディガン A との類似度が高いアイテムにはファッション系のアイテムが多く出現し、それらは同一の店舗で販売されているアイテムであることが分かる。一方で、ワンピース A については表 9 より、類似度の高いアイテムとしてレディースのファッション系のアイテムが出現しているが、ペット用品や食品などのアイテムも出現している。さらにニットカーディガン A と異なり、ワンピース A では異なる店舗で販売されているアイテムとの類似度が高くなっている。この違いは、主にアイテムの被購買件数に依存していると考えられる。アイテムはグラフ上で購買されたユーザと、販売店舗、アイテムカテゴリのエンティティと接続されており、これらの周辺のエンティティの情報を関係性の強さ π に基づいて取り入れながら、そのベクトル表現を学習する。さらに、このベクトルの学習はグラフの損失 $Loss_{KG}$ とユーザとアイテムに関する損失 $Loss_{CF}$ を同時に最小化するように更新する。その

表 6 抽出したアイテム

Table 6 Selected items.

アイテム名	店舗名	被購買件数
ニットカーディガン A	AquaGarage	10
ワンピース A	Dark Angel	288

表 7 該当アイテムの店舗とのエッジの重み $\pi(i, r, s)$

Table 7 Wedge wight $\pi(i, r, s)$ between selected item and store.

アイテム名	$\pi(i, r, s)$
ニットカーディガン A	3.119×10^{-1}
ワンピース A	8.255×10^{-6}

表 8 ニットカーディガン A と類似度の高いアイテム

Table 8 Items with a high degree of similarity to Knit Cardigan A.

	アイテム名	店舗名	cos 類似度
1	ボーダー柄ワンピース	AquaGarage	0.912
2	女性用下着	小悪魔ドレス DiamondHearts	0.906
3	ブーツ	AquaGarage	0.904
4	マキシ丈ワンピース	AquaGarage	0.903
5	長袖ワンピース	AquaGarage	0.901
6	ジェルネイル	ネイル工房	0.887
7	フレアスカート	AquaGarage	0.887
8	シャツワンピース	AquaGarage	0.885
9	もこもこパーカ	AquaGarage	0.885
10	マキシ丈ワンピ	AquaGarage	0.882

表 9 ワンピース A と類似度の高いアイテム
Table 9 Items with a high degree of similarity to One Piece A.

	アイテム名	店舗名	cos 類似度
1	トリミングシザ	ペット用品ペットの道具屋さん	0.932
2	レディースインナートップス	Dark Angel	0.932
3	保湿ケアクリーム	ゲルクリームの【レステモ】	0.912
4	ドルマンカーディガン	Pierrot 【ピエロ】	0.886
5	レディースパーカ	神戸レタス (KOBE LETTUCE)	0.879
6	マキシワンピース	ワンピース	0.870
7	女性用下着	三軒茶屋通信インナーウェア三恵	0.867
8	ルイボスティ	ティーライフ shop 健康茶自然食品	0.863
9	コットンスキニーパンツ	kormarch	0.863
10	レギンス	ギャルスター (GALSTAR)	0.862

ため、アイテムが多くユーザに購入されている場合は、 $Loss_{CF}$ で発生する損失が大きくなり、知識グラフ定義されている店舗との関係性の重要度 $\pi(i, r, s)$ の値が小さくなるのが考えられる。実際にワンピース A とニットカーディガン A の店舗とのエッジの重み $\pi(i, r, s)$ を比較すると明らかにニットカーディガン A の方がその重要度が高いことが分かる。

このように、このモデルでは、購買件数が少ないアイテムに関してグラフで定義されている「知識」を用いることによって、適切なベクトル表現を得ることができているといえる。このことは、推薦システムの研究の分野において、モデルの「解釈性」とともに主張されている知識グラフをモデルに用いることの有用性である。一方で、ワンピース A に関しては、ユーザとアイテムの購買関係がベクトルの表現に大きく影響しており、直感的に類似しているとは考えにくいペット用品や食品などのアイテムとも類似度が高くなっている。特に、本研究で対象とした EC サイトにおいては、ユーザが様々なジャンルの商品を購入するため、ユーザとアイテムの購買関係を用いてモデルの学習を行うと、潜在的な関係性が学習できる反面、ノイズのようなデータの影響を受けてしまうことも考えられる。

6. 考察

6.1 モデルの拡張性に関する考察

本稿では、ユーザの属性として「年代・性別」を用いたが、提案手法の入力とするグラフは分析者によって任意に変更が可能であり、ユーザの属性として、ソーシャルネットワーク上のつながりを考慮することも可能である。実際にソーシャルネットワーク上でつながれているユーザ同士はその嗜好が類似している場合が多く、実際にこのつながりを考慮した推薦モデルも提案されている。多くのユーザとのつながりがあり、かつそれらの関係性が強いユーザは、多くのユーザの嗜好に影響を与えているユーザである。このようなユーザを検出し、この購買行動を分析することで、ユーザの嗜好の特性を把握することが可能であると考えら

れる。また、アイテムに関しても同様に任意の補助情報を考慮することができる。提案モデルを映画の視聴履歴に適用した場合、映画とその補助情報との関係性の強さを分析することによって、映画の視聴理由に対する考察が可能であり、映画と関係性の強い役者がいる場合には、ユーザがその役者に注目して視聴に至っている可能性が高いという解釈ができる。また、この結果を用いて、そのような役者に焦点を当てた施策の実施などの提案につなげることが可能である。このように、提案手法はモデルへの入力とするグラフを分析の目的によって変更させることができ、非常に拡張性の高い実用的なモデルであるといえる。

6.2 知識グラフの埋め込み方法

提案手法では TransR に基づいて、知識グラフのエンティティおよびリレーションのベクトル表現の学習を行った。しかし、TransR は知識グラフ上で定義されたトリプルを埋め込み空間でいかに厳密に表現するかに焦点を当てて提案されている手法であり、分析モデルを構築するにあたっては、目的に応じて他の手法を用いることも考えられる。提案モデルでは、リレーションごとに異なる埋め込み空間を定義しており、各 Relation space において、その関係性を表現している。そのため、リレーションが異なるエンティティどうしを同一の意味空間上で比較することができない。TransR ではなく、より単純なモデルである TransE を用いることで、事実関係性の厳密性は欠くものの、すべてのトリプルを同一の埋め込み空間上で定義することができる。これによって、属性とアイテムなど異なる性質を持ったエンティティどうしが同一の意味空間で表現され、ある「属性」と類似度の高い「アイテム」などの議論が可能になる。これによって、ある属性を代表するようなアイテムの抽出などの分析が可能になると考えられる。このように、分析のしやすさの観点から、モデルを再構築することも可能である。

6.3 提案モデルを用いた分析の安定性

ECサイトなどのデータセットにおいては、ユーザの購買件数やアイテムの被購買件数にばらつきがあることが一般的である。よって、このようなデータセットを対象とする場合、FMなどの次元削減をベースとする手法では、データ間の既知の関係性を考慮することができないため、データ数が十分に存在しないユーザやアイテムはモデルの効率的な学習の妨げとなる可能性が高い。しかし、提案手法においては、データ数が少ない場合に与えられた知識をAttention機構を用いて重要視することによって、適切なベクトル表現の学習を行うことが可能である。そのため、ECサイトのように膨大なアイテムが存在するデータセットにおいても、適切な学習を行うことが可能であり、安定した分析が行えるといえる。

6.4 マーケティングへの応用

マーケティング施策を実施する際、企業は顧客全体を一定の基準に基づきユーザを分割して、分割後の各セグメントに対して施策を検討することが多い。この分割の基準としては、年代や性別といったデモグラフィック属性のほかにも、居住地などの地理的や特徴や、アンケートに基づく心理的な分類であるサイコグラフィック属性などが用いられることがある。そして実務において、セグメントの分割基準の決定は経験に基づいた判断によって行われることが多い。本提案手法においては、顧客の属性情報を自由に与えることが可能であり、式(16)によって属性ごとの嗜好の共通度合いを定量化することができる(図5)。ここで、 $\pi(u, r, a)$ の値が大きく嗜好が共通化しているユーザ群に対しては、セグメントマーケティングが効果的であると考えられる。そのため、セグメント分割の候補となる各属性 a を与えた場合のモデルの精度とユーザと属性の間のエッジの重みの平均値 $\mu(u, r, a)$ から、企業が有する属性情報のうち、いずれがセグメントマーケティングを行う際の分割基準として有効であるかを定量的に評価することができる。具体的には、モデルの予測精度が高く、 $\mu(u, r, a)$ が大きい属性 a を有するユーザ群にはセグメントマーケティングが有効であると判断することができる。これにより、提案手法を用いることにより、企業が有するユーザの属性情報の中から、効果的なセグメント分割基準を発見することが可能となると考えられる。

7. 結論と今後の課題

本研究では、様々なデータを統合的に扱い、多角的な分析を可能とすることを目的とし、KGATをユーザの属性情報を考慮したモデルへと拡張することによって、ユーザの嗜好を属性情報と購買アイテムおよびその補助情報の観点から分析可能なモデルを提案した。提案モデルを実データに適用し、予測精度の評価とユーザの購買行動に関する分

析を行った。予測精度の比較においては、従来手法と比較して評価指標の観点からも優れていることを示した。ユーザの嗜好に関する分析においては、ユーザの嗜好をモデルの学習により得られた知識グラフ上のエッジの重みを用いて解釈することが可能であることを示した。また、分析結果に基づいてマーケティング施策を提案できることから、提案手法は実応用可能性の高い手法であるといえる。加えて、アイテムと店舗においても同様の分析を行い、アイテムと店舗の関係性の強さから店舗の性質について考察を行い、グラフのエッジの重みを用いた分析の有用性を示した。以上により、提案モデルでは、グラフ上のエッジの重みを用いて様々な観点から分析を行うことが可能となり、顧客の嗜好を理解するために有用な分析モデルであるといえる。提案モデルへの入力とするグラフは分析の目的によって自由に構築することが可能であるため、様々な分野への応用が期待される。

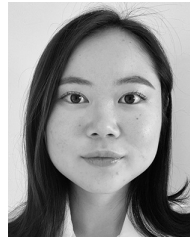
今後の課題として、購買順序などのユーザの行動系列の時系列性を考慮することがあげられる。これにより、ユーザの嗜好の変化やトレンドなどについても分析が可能になると考えられる。

謝辞 本研究では、国立情報学研究所のIDRデータセット提供サービスにより楽天株式会社から提供を受けた「楽天データセット」を使用いたしました。貴重なデータの提供に深く感謝いたします。また、本研究の一部はJSPS科研費21H04600の助成を受けたものです。

参考文献

- [1] Blei, D.M., Ng, A.Y. and Jordan, M.I.: Latent Dirichlet Allocation, *the Journal of machine Learning research*, Vol.3, pp.993–1022 (2003).
- [2] Barkan, O. and Koenigstein, N.: Item2vec: Neural Item Embedding for Collaborative Filtering, *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp.1–6, IEEE (2016).
- [3] 保坂大樹, 河部瞭太, 山下 遥, 後藤正幸ほか: 意味空間上の分布表現に基づく Web サイトと閲覧ユーザの統合分析モデル, *情報処理学会論文誌*, Vol.60, No.8, pp.1390–1402 (2019).
- [4] Wang, X., He, X., Cao, Y., Liu, M. and Chua, T.-S.: KGAT: Knowledge Graph Attention Network for Recommendation, *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp.950–958 (2019).
- [5] 坂元哲平, 山下 遥, 荻原大陸, 後藤正幸ほか: 就職ポータルサイトにおける企業のアピールポイントと学生の志望理由のマッチング分析モデルに関する一考察, *情報処理学会論文誌*, Vol.58, No.9, pp.1535–1548 (2017).
- [6] Shimizu, R., Sakamoto, T., Yamashita, H. and Goto, M.: Proposal of a Purchase Behavior Analysis Model on an Electronic Commerce Site using Questionnaire Data, *Total Quality Science*, Vol.4, No.1, pp.1–12 (2018).
- [7] Kobayashi, M., Mikawa, K., Goto, M., Matsushima, T. and Hirasawa, S.: Collaborative Filtering Based on the Latent Class Model for Attributes, *2017 16th IEEE International Conference on Machine Learning and Ap-*

- plications (ICMLA)*, pp.893–896, IEEE (2017).
- [8] Mikolov, T., Chen, K., Corrado, G. and Dean, J.: Efficient Estimation of Word Representations in Vector Space, *arXiv preprint arXiv:1301.3781* (2013).
- [9] Rendle, S.: Factorization Machines, *2010 IEEE International Conference on Data Mining*, pp.995–1000, IEEE (2010).
- [10] He, X. and Chua, T.-S.: Neural Factorization Machines for Sparse Predictive Analytics, *Proc. 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, pp.355–364 (2017).
- [11] Lian, J., Zhou, X., Zhang, F., Chen, Z., Xie, X. and Sun, G.: XdeepFM: Combining Explicit and Implicit Feature Interactions for Recommender Systems, *Proc. 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp.1754–1763 (2018).
- [12] Wang, X., Wang, D., Xu, C., He, X., Cao, Y. and Chua, T.-S.: Explainable Reasoning over Knowledge Graphs for Recommendation, *Proc. AAAI Conference on Artificial Intelligence*, Vol.33, No.1, pp.5329–5336 (2019).
- [13] Huang, X., Fang, Q., Qian, S., Sang, J., Li, Y. and Xu, C.: Explainable Interaction-driven User modeling over Knowledge Graph for Sequential Recommendation, *Proc. 27th ACM International Conference on Multimedia*, pp.548–556 (2019).
- [14] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I.: Attention is all you need, *arXiv preprint arXiv:1706.03762* (2017).
- [15] Hu, B., Shi, C., Zhao, W.X. and Yu, P.S.: Leveraging Meta-path based Context for Top-N Recommendation with a Neural co-Attention Model, *Proc. 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp.1531–1540 (2018).
- [16] Seo, S., Huang, J., Yang, H. and Liu, Y.: Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction, *Proc. 11th ACM conference on recommender systems*, pp.297–305 (2017).
- [17] Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J. and Yakhnenko, O.: Translating Embeddings for modeling Multi-Relational data, *Neural Information Processing Systems (NIPS)*, pp.1–9 (2013).
- [18] Lin, Y., Liu, Z., Sun, M., Liu, Y. and Zhu, X.: Learning Entity and Relation Embeddings for Knowledge Graph Completion, *Proc. AAAI Conference on Artificial Intelligence*, Vol.29, No.1 (2015).
- [19] Kingma, D.P. and Ba, J.: Adam: A Method for Stochastic Optimization, *arXiv preprint arXiv:1412.6980* (2014).
- [20] Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X. and Guo, M.: Ripplenet: Propagating User Preferences on the Knowledge Graph for Recommender Systems, *Proc. 27th ACM International Conference on Information and Knowledge Management*, pp.417–426 (2018).
- [21] 楽天株式会社：楽天市場データ，国立情報学研究所情報学研究データリポジトリ（データセット），入手先（<https://doi.org/10.32130/idr.2.1>）（2014）。



伊藤 史世

1997年生。2020年早稲田大学創造理工学部経営システム工学科卒業。現在、同大学大学院創造理工学研究科経営システム工学専攻修士課程在学中。機械学習を用いたデータ分析に関する研究とその実応用に興味を持つ。



張 志穎

1996年生。2018年山東大学統計学部応用数学科卒業。2021年早稲田大学大学院創造理工学研究科経営システム専攻修士課程修了。在学時 Long Short-Term Memory のモデル構造に関する研究に従事。



雲居 玄道

2008年早稲田大学理工学部経営システム工学科卒業，2008年同大学理工学術院総合研究所嘱託研究員。2015年浄土真宗本願寺派総合研究所研究助手。2017年早稲田大学大学院創造理工学研究科博士後期課程入学。2019年早稲田大学創造理工学部経営システム工学科助手，現在に至る。情報数理応用・テキストマイニングの研究に従事。経営情報学会，日本気象学会各会員。



後藤 正幸（正会員）

1969年生。1994年武蔵工業大学大学院修士課程修了。2000年早稲田大学大学院理工学研究科博士課程修了。博士（工学）。1997年同大学理工学部助手。2000年東京大学大学院工学研究科助手。2002年武蔵工業大学環境情報学部助教授。2008年早稲田大学創造理工学部経営システム工学科准教授。2011年同大学教授。情報数理応用とデータサイエンス，ならびにビジネスアナリティクスの研究に従事。著書に、『入門パターン認識と機械学習』コロナ社（2014），『ビジネス統計-統計基礎とエクセル分析』オデッセイコミュニケーションズ（2015）等。IEEE, INFORMS, 電子情報通信学会, 人工知能学会, 日本経営工学会, 経営情報学会等各会員。