

2024年度修士論文

アスペクト情報を活用した
知識グラフ上のパス推論による
説明可能なホテル推薦システム

東京科学大学 工学院 経営工学系

指導教員 市瀬 龍太郎

学籍番号 23M14834

氏名 安田 大輝

English Abstract

Explainable Hotel Recommendation System through Path Reasoning on Knowledge Graphs Utilizing Aspect Information

Hiroki Yasuda

The hotel recommendation system automatically recommends the best choice for users in the hotel reservation service. By using this system, users can efficiently select hotels from a large number of choices, and service providers can improve customer satisfaction and increase sales. In this study, we aim to develop a recommendation system that provides users with more understandable reasons for recommendation and helps them make a convincing choice in hotel reservation, which is a more important decision than video viewing or product purchase. To achieve this, we propose a new method for constructing a knowledge graph that utilizes aspect information such as location, meals, and baths included in user reviews of hotels. Then, by using a recommendation algorithm based on this knowledge graph, we determine the path between the user entity and the hotel entity to be recommended, which is called the explanation path, and provide a more understandable explanation of the reason for recommendation based on the aspect. Additionally, we also devise a aspect-oriented post-processing method for models to increase the number of more understandable explanation paths. Furthermore, we propose a new quantitative matrix, AO-Score, to evaluate the quality of recommendation explanation paths. Through these proposals, we can ensure the scalability and versatility of the system. In our experiments, we used user reviews of actual hotel reservation services and found that the proposed method with high explainability performs as well as the conventional matrix factorization-based methods that focus on accuracy improvement. We also confirmed that the type of explanation paths can effectively be changed through post-processing of the proposed model and that the quality of the paths can be quantitatively evaluated by AO-Score. Finally, through case studies, we qualitatively examine how explanation paths and sentences improve the ease of understanding of the reasons for the recommendation.

論文要旨

ホテル推薦システムは、ホテル予約サービスにおいてユーザに最適な選択を自動的に推薦する。このシステムを利用することで、ユーザは膨大な選択肢から効率的にホテルを選択し、サービス提供者は顧客満足度の向上を図ることができる。本研究では、動画閲覧や商品購入よりも重要な意思決定であるホテル予約において、ユーザにとってより理解しやすい推薦理由を提供し、納得感のある選択を支援する推薦システムの開発を目指す。これを実現するために、ホテルのユーザレビューに含まれる立地、食事、風呂などのアスペクト情報を活用した新たな知識グラフの構築方法を提案する。そして、この知識グラフに基づく推薦アルゴリズムを利用することで、ユーザとホテルの間のパスを決定し、アスペクトによる理解のしやすい推薦理由の説明を提供する。さらに、アスペクトを重視したモデルの後処理方法と、推薦の説明パスを評価する新たな定量指標を考案し、システムの拡張性や汎用性を確保する。推薦精度をはじめとした定量的な検証と、提供された説明文による定性的なケーススタディにより、提案システムの効果が示された。

目次

English Abstract	iii
論文要旨	v
表目次	ix
図目次	xi
1 序論	1
1.1 研究背景	1
1.1.1 ホテル推薦システム	1
1.1.2 知識グラフを用いた推薦理由の提供	2
1.1.3 既存システムの問題点	4
1.2 研究目的	4
1.2.1 説明の理解のしやすさ	4
1.2.2 ユーザレビューにおけるアスペクト情報	5
1.3 本論文の構成	6
2 関連研究	7
2.1 一般的な推薦アルゴリズム	7
2.1.1 内容ベースフィルタリング	7
2.1.2 協調フィルタリング	8
2.2 知識グラフを用いた説明可能な推薦アルゴリズム	9
2.2.1 埋め込み表現に基づく手法	9
2.2.2 パス推論に基づく手法	9
2.3 知識グラフを用いた推薦の説明に着目した研究	10
2.4 PGPR (Policy-Guided Path Reasoning)	11
2.4.1 MDP 定式化	12
2.4.2 強化学習とパス推論	13

3	提案手法	15
3.1	AO-PGPR (Aspect-Oriented PGPR)	15
3.2	知識グラフの構築	16
3.2.1	エンティティとリレーションの定義	16
3.2.2	トリプルの構築	17
3.2.3	説明パスの種類	18
3.3	アスペクトを重視したモデルの後処理	19
3.3.1	パス選択	19
3.3.2	リランキング	20
3.4	説明パスの定量評価	20
4	実験	23
4.1	実験設定	23
4.1.1	データセット	23
4.1.2	評価指標	24
4.1.3	ベンチマーク	26
4.2	実験結果	26
4.2.1	推薦精度	26
4.2.2	パスごとのカバー率の変化	28
4.2.3	AO-Score の比較	29
4.2.4	リランキングの推薦精度への影響	30
4.2.5	ケーススタディ	31
5	考察	33
6	結論	35
	謝辞	37
	参考文献	39
A	付録 A アスペクトベース感情分析	43

表目次

1.1 ユーザレビューの例	5
4.1 フィルタリング後のインタラクションデータの集計	24
4.2 知識グラフのエンティティ, リレーション, パスの集計	25
4.3 リレーションの組み合わせが推薦精度に与える影響	26
4.4 パラメータ θ_{r_B} が推薦精度に与える影響	27
4.5 パラメータ θ_{r_C} が推薦精度に与える影響	27
4.6 推薦精度の比較	28
4.7 説明文のテンプレート	31

図目次

1.1	推薦システムに利用される知識グラフの例 (音楽領域)	3
3.1	提案システムの概観図	16
3.2	知識グラフの構築方法	17
4.1	モデルの後処理におけるパスごとのカバー率の変化	29
4.2	AO-Score の比較	30
4.3	リランキングによる推薦精度の各評価指標の変化	31
4.4	説明文によるケーススタディ	32

1 序論

本章では, 研究背景として, ホテル推薦システムと知識グラフによる推薦理由の提供について解説する. 次に, 指摘した既存システムの問題点に対処するための新しいアプローチを研究目的として提示する. 最後に, 本論文の構成について説明する.

1.1 研究背景

本節では, ホテル推薦システムの概要と, 説明可能な推薦が必要とされる理由を説明する. さらに, 説明可能性を実現するための知識グラフを用いた手法とその限界についても検討する.

1.1.1 ホテル推薦システム

推薦システムは, インターネットサービスにおいて重要な役割を担い, 機械学習をはじめとした人工知能 (Artificial Intelligence, AI) 技術などを用いて, ユーザの過去の行動や好みを分析し, 個々に最適化されたアイテムを自動的に提案する. これらは映画や音楽のストリーミングサービス, オンラインショッピング, ソーシャルメディアでの情報発見など, 日常生活の様々な場面で活用されている. これにより, ユーザはインターネット上の膨大な選択肢の中から, 自分の興味や好みに合ったアイテムを容易に発見することができる. さらに, サービス提供者はこれを利用して顧客満足度を高め, 売上向上を図ることが可能である. 映画ストリーミングサービスの Netflix では, 推薦システムによる影響が特に顕著であり, 2015 年時点での視聴活動の約 80 % が推薦システムに支えられていたと報告されている [1]. また, オンライン

ンショッピングサービスの Amazon では、2018 年に推薦システム経由での売上が全体の約 35 % を占めるほどであった [2]。このように、推薦システムはサービスの成長と顧客体験の向上に寄与し、今後その影響はさらに増すことが期待される。

ホテル推薦システムは、旅行や出張時にユーザが宿泊するホテル選択を支援することで、ホテル予約サービスにおける重要な役割を果たす。このシステムは、動画閲覧や商品購入に用いられる他の推薦システムと技術的に共通点がありながらも、意思決定の支援という点では根本的な違いを持つ。楽天インサイト株式会社が 2022 年の夏期に実施した調査 [3] によると、宿泊を伴う日本国内旅行の平均支出額の中で宿泊費は 26,933 円に達し、総支出の約 30% を占めている。また、旅行計画の開始は出発の 1 ヶ月以上前に行われるケースが 70% 以上に上り、計画期間の長さが重要な意思決定であることを示している。これらの事実を、即座に利用可能で低コストな動画サービスや、短期間での衝動的な購入が可能であるオンラインショッピングと比較すると、ホテル選択の重要性が一層明確となる。したがって、ホテル推薦システムは、高額な出費と長期的な検討を必要とするホテル選択をサポートし、ユーザの旅行体験の満足度を高めるという極めて重要な役割を担っているのである。

この文脈において、説明可能な AI (Explainable AI, XAI) の必要性が浮き彫りとなる。現代の AI システム、特に深層学習に基づくものは、その出力の決定プロセスが極めて複雑であり、しばしば人間には理解しがたい「ブラックボックス問題」を引き起こす [4]。説明可能な AI は、この問題に対応するために開発された技術で、AI がどのようにして特定の出力を導出したのかを人間が理解できる形で明確に説明する。EU 一般データ保護規則 (GDPR) などの規則は、「説明を受ける権利」を要求しており [5]、AI の出力結果の根拠をシステムのユーザに説明することが、現代のシステムにとって不可欠な性質となっている。そして、説明可能な AI の概念を推薦システムに適用したものが、説明可能な推薦システム (Explainable Recommendation System) である。このシステムは、ユーザの好みや以前の選択に基づいてアイテムを推薦するだけでなく、その推薦理由もユーザに提供する。説明可能な推薦は、推薦システムの説得力や満足度の向上に寄与する [6]。このことから、重要な意思決定プロセスに関わっているホテル推薦システムにおいて、ユーザに推薦理由を説明することは、納得感のある選択を支援するために、必須の技術であるといえる。

1.1.2 知識グラフを用いた推薦理由の提供

説明可能な推薦システムを実現する技術の 1 つに、知識グラフを用いた手法が存在する。知識グラフは、エンティティとリレーションを基本要素として使用し、様々なデータや知識の関連性をグラフ形式で表現するネットワークである。エンティティはある物事を表すノードであり、リレーションはエンティティ間の関係を表すエッジである。エンティティとリレーションを組み合わせたものをトリプルと呼び、形式は (ヘッドエンティティ, リレーション, テイルエンティティ) となる。この形式は、自然言語の文法における主語、述語、目的語に類似しており、知識の関係性を明確に表現する。これらのトリプルによって構成される知識グラフを用いることで、複雑なデータの関係が明確となり、高度な検索や推論を容易に行うことが可能となる。

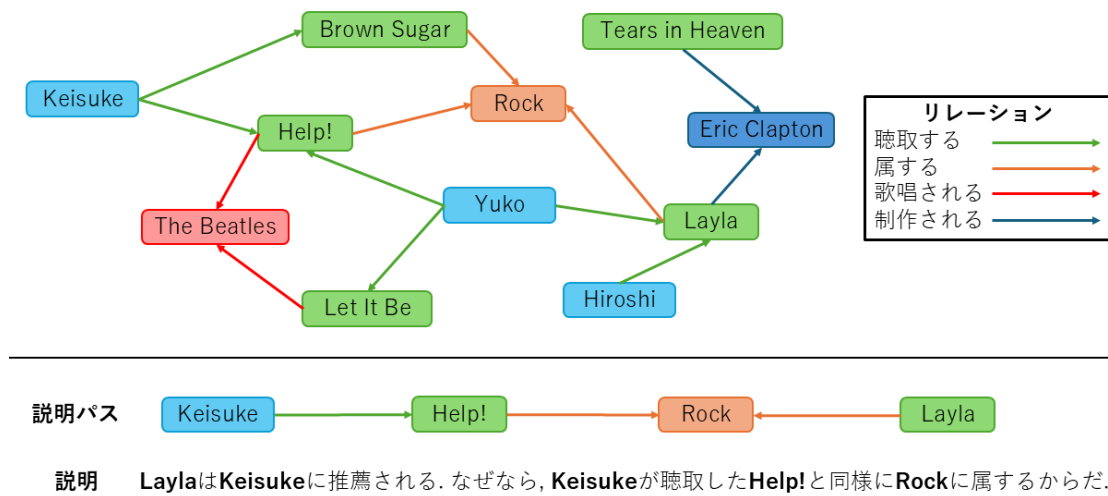


図 1.1: 推薦システムに利用される知識グラフの例 (音楽領域). エンティティは、水色がユーザ、緑がアイテム (楽曲), オレンジがカテゴリ, 赤がアーティスト, 青が制作者を表す. パス推論で決定した説明パスによって, 推薦理由の説明が提供される.

推薦システム用のデータセットとして, 映画に関連する MovieLens 1M (ML1M) [7] や音楽に関連する Last.FM 1M (LFM1M) [8] があり, これらには, 知識グラフが構築されている. LFM1M において構築される知識グラフの例を図1.1に示す. エンティティには「ユーザ」, 「アイテム (楽曲)」, 「アーティスト」など, リレーションには「聴取する」や「制作される」などが該当する. また, (Keisuke, 聴取する, Help!) や (Tears in Heaven, 制作される, Eric Clapton) といったトリプルが存在する.

知識グラフによる推薦システムは, 豊富な属性情報を活用し, 推薦の精度を向上させる. さらに, パス推論 (Path Reasoning) という特有の技術を用いることで, 具体的な推薦理由をユーザに提供することが可能である. パス推論は, 知識グラフ上で定義されたユーザエンティティとアイテムエンティティ間のパスを機械学習を用いて特定する. ここでいうパスとは, エンティティ間の関係性を表すリレーションが形成する連なりで, この連続をたどることにより, ユーザの好みやアイテムの特性が明らかになる. 推薦対象のユーザがシステムに入力されたとき, 最も関連性が高いと推定されるパスが決定され, これを「説明パス」と呼ぶ.

説明パスは, 推薦するアイテムに至るまでの論理的な経路を提供し, ユーザには最終的に推薦アイテムを選ぶ直接的な理由として提示される. 推薦対象ユーザ「Keisuke」に対して, 図1.1の例に示された説明パスが決定された場合, 推薦アイテムとして「Layla」(楽曲) が選ばれる. このシナリオでは, 説明パスは「Keisuke が聴取した『Help!』(楽曲) と同様に『Rock』(カテゴリ) に属するからだ」という形で推薦理由が提供される. このように, 知識グラフ上のパス推論によって, アイテムの推薦と同時に, その推薦理由を直観的にユーザに説明することが可能となる.

1.1.3 既存システムの問題点

知識グラフを用いた説明可能な推薦システムに関する研究については、第2章で詳細に解説する。これらのシステムには、主に以下の2つの重要な問題点が存在する。

第一に、多くのシステムは単に推薦理由をユーザに提供することに留まり、推薦の背後にある論理の理解のしやすさを向上させるための具体的な知識グラフ構築方法については十分に議論されていない。また、これらのシステムは、ML1M や LFM1M のような豊富な情報を含む知識グラフの存在を前提とするが、その構築方法についての詳細はしばしば省略される。その結果、ユーザにとって納得できる意思決定を支援しない推薦理由が提供されてしまう可能性がある。第二に、ほとんどの既存システムでは説明パスがケーススタディとして例示されるに留まり、説明パスの質そのものを定量的に評価する指標が欠如している。このため、ユーザにとって実際に有益な説明が提供されているかを評価することが困難である。次節ではこれらの問題を解決するためのアプローチを述べる。

1.2 研究目的

前節では、既存の知識グラフを用いた推薦システムに存在する主要な問題点を2点指摘した。第一に、説明の理解のしやすさを高めるための知識グラフ構築方法に関する議論が不十分であることを挙げたが、これに対しユーザレビューからアспект情報を抽出し利用する新たな構築方法を提案する。詳細は本節にて展開する。第二に、説明パスの定量評価指標の欠如にも対処し、ユーザレビューのアспект別評価に基づく新しい評価指標（AO-Score）を導入することで、説明パスに含まれるアспектに対するユーザの実際の満足度が測定可能とする。

これらの解決策を通じて、ユーザにとってより理解のしやすい推薦理由を提供し、納得感のある選択を促すホテル推薦システムを開発することを本研究の目的とする。本研究は、ユーザの選択支援だけでなく、サービス開発者がビジネス企画者へ提案する際の説得力を向上させることや、サービス提供者がユーザに対して説明責任を果たしやすくする仕組みを整備することなどにも寄与する。

1.2.1 説明の理解のしやすさ

本研究では、ユーザが提示された推薦アイテムの背後にある論理を直観的に把握できるか、つまり説明の理解のしやすさに着目する。パス推論によって提供される説明がユーザにとって分かりやすいものであれば、システムへの信頼感やアイテムの受容性が向上し、より納得のいく意思決定が促されることが考えられる。

この考えに基づき、本研究では従来の単なる推薦理由の説明を超え、ユーザが直観的に理解できる推薦理由の提供方法を探求する。具体的には、後述するアспект情報に基づいた知識グラフの構築手法を通じて、推薦アイテムとユーザとの関連性を明確に示すアプローチを提案する。

表 1.1: ユーザレビューの例.

項目	サンプル
ユーザ ID	user_1
投稿日時	2019-01-01 12:34:56
アイテム ID	121212
立地	4
部屋	5
食事	5
風呂	3
サービス	4
設備	3
コメント	初めて利用しました。朝食が充実していてよかったです。 空気清浄機があればもっとよかったです。

実験で利用した楽天トラベルデータ [9] の紹介ページに記載の例を基に作成.

1.2.2 ユーザレビューにおけるアスペクト情報

本研究では, 説明の理解のしやすさを高めるために, 一般的なユーザレビューから得られるデータを基にした知識グラフの構築方法を提案する. ユーザレビューは, ユーザが経験したアイテムに関する詳細なフィードバックを含み, ユーザ ID, 投稿日時, アイテム ID, 各アスペクトに対する評価であるアスペクト別評価と具体的なコメントから成り立っている.

表 1.1 は, 具体例としてホテルのユーザレビューの形式を示している. アスペクトとは, 表における「立地」, 「部屋」, 「食事」, 「風呂」, 「サービス」, 「設備」といった観点のことを意味し, 各アイテムの特性や利点を明確に示すものである. アスペクトに関する情報は, アスペクト別評価の点数だけでなく, コメント内にも含まれる. 例えば, 表内のコメントにおいて, 「朝食」という単語が示す「食事」アスペクトに対して, ユーザはポジティブな印象を示していることが分かる. 同様に, 「空気清浄機」について言及されている箇所では, 「設備」アスペクトに対するネガティブな感想が示されている. これらのコメントを通じて, ユーザが「食事」と「設備」のアスペクトに特に関心を持っていることが示唆される. このように, ユーザレビュー内のアスペクト別評価とコメントは, ユーザの重視するアスペクトや, アイテムの得意領域に関するアスペクトを特定できる可能性を持っている. このため, 本研究ではアスペクト別評価とコメントを合わせて, アスペクト情報と定義する.

本研究で提案するアスペクト情報を活用した推薦システムは, ホテル領域に留まらず, 同様の形式のユーザレビューのデータが存在するあらゆる領域に適用可能である. 領域の例としては, レストラン, 美容院, 不動産などが想定できる.

1.3 本論文の構成

本論文は, 以下に示す流れで論じられる. 第2章では, 推薦システムに関する既存研究についての具体的な説明を行い, 本研究との関連を明らかにする. 第3章では, 提案手法について述べる. 提案する知識グラフ構築方法, モデルの後処理, パスの定量評価指標の3点を詳細に解説する. 第4章では, 実際のホテル予約サービスのユーザーレビューを用いて, 提案手法の効果を検証する. 従来の推薦システムで用いられる指標などによる定量的な評価と, 実際のシステムによって出力された推薦理由の説明文による定性的なケーススタディで構成される. 第5章では, 実験結果を踏まえて明らかになった本研究における課題を中心に考察する. 第6章では, 本研究の結論を述べ, 本論文のむすびとする.

2

関連研究

本章では, 本研究の関連研究を紹介する. 一般的な推薦アルゴリズムについて説明した後, 知識グラフを用いた説明可能な推薦システムに関する既存研究に焦点を当てて解説を行う. 推薦アルゴリズムとその応用研究について, それぞれ技術的進展と知識グラフの構築方法や説明の質を向上させる手法を概説する. これらの調査を通じて, 本研究の位置づけや意義を明確にすることが, 本章の目的である. また, 最後に本研究で用いるアルゴリズムについて, 提案手法の準備として説明する.

2.1 一般的な推薦アルゴリズム

本節では, 一般的な推薦アルゴリズムとして, 特に内容ベースフィルタリングと協調フィルタリングについて紹介する. それぞれに後述する利点や欠点が存在するため, これらのアプローチを組み合わせたハイブリッドな手法も提案されている [10]. 本研究における知識グラフを用いた手法もその 1 つである.

2.1.1 内容ベースフィルタリング

内容ベースフィルタリングは, テキスト (コメント) 内の記述, 画像の特徴, 音楽のジャンルなどのアイテムの特徴に基づいてユーザに推薦を行う手法である. 基本的なアプローチとして, ユーザが過去に興味を示したアイテムの特徴を分析し, ユーザプロファイルを作成後, それに類似度が高い新たなアイテムを推薦する.

近年では, テキスト情報を扱う際に LDA (Latent Dirichlet Allocation) [11] などのトピックモデルや, word2vec [12], doc2vec [13] といった機械学習モデルを利用し,

アイテムの内容をより深く解析する方法が導入されている。LDA を用いた手法ではテキストの集合から潜在的なトピックを抽出し、それらに基づいてアイテムの類似性を評価する。また、word2vec や doc2vec を用いて単語や文書をベクトル空間に埋め込むことで、意味的な類似性を捉えた推薦を可能にする。

Marko らによると、内容ベースフィルタリングには以下の利点と欠点が存在する [14]。利点としては、アイテムの特徴やユーザプロファイルに基づいて推薦が行われるため、新規アイテムや新規ユーザであっても推薦が可能であることが挙げられる。一方、アイテムの特徴を抽出することが困難な場合の推薦の質の低下や、推薦がユーザの既知の好みに基づいていることによる新しいカテゴリや未知のアイテムを発見する機会の減少などの欠点が指摘されている。

2.1.2 協調フィルタリング

協調フィルタリングは、ユーザとアイテムの過去のインタラクション（閲覧、購入、クリックなど）に基づいて、ユーザが興味を持つ可能性のあるアイテムを推薦するアプローチである。1つの分類方法として、過去のインタラクションが似ているアイテムに基づいて推薦を行うアイテムベース協調フィルタリングと、過去の行動が似ているユーザに基づいて推薦を行うユーザベース協調フィルタリングがある。

アイテムベース協調フィルタリングの例として、アソシエーションルール [15] を紹介する。このアプローチは、アイテム間の共起関係を分析し、同様の傾向があるアイテム群を特定する。インタラクションデータを基に、例えば「ユーザが牛乳を購入した場合、パンも購入する可能性が高い」といった関連性のルールを発見する。アソシエーションルールは、シンプルながらも強力な方法であり、ユーザの明確なインタラクションのパターンを基に推薦を行うことができる。

他の協調フィルタリングのアプローチとして行列分解がある。行列分解はユーザを行方向、アイテムを列方向に配置した行列を基に動作する。この行列は、ユーザとアイテム間のインタラクションが存在すればそのセルに1を、存在しなければ0を配置して構成される。行列分解の目的は、この疎な行列から、ユーザとアイテムの潜在的な特徴を捉えることで、インタラクションのないユーザとアイテムの組み合わせに対する好みを予測することである。代表的なアルゴリズムとして IMF (Implicit Matrix Factorization) [16] と BPR (Bayesian Personalized Ranking) [17] を紹介する。IMF は、観測されたユーザとアイテムのインタラクションを、ユーザとアイテムの潜在的な特徴ベクトル間の内積で近似することで、ユーザが各アイテムをインタラクションする確率を予測する。BPR は、ユーザが実際にインタラクションしたアイテムは他のアイテムよりも好ましいという前提で、モデルのパラメータを最適化する手法で、ランキングの精度を高めるように設計されている。

Marko らによると、協調フィルタリングには以下の利点と欠点が存在する [14]。協調フィルタリングでは、アイテムの特徴を利用しないため、あらゆる種類のドメインに対応でき、ユーザにとって未知のアイテムを推薦することもできる。しかし、新規アイテムや評価データの少ない状況では推薦が困難になるという欠点がある。

2.2 知識グラフを用いた説明可能な推薦アルゴリズム

本節では、知識グラフを用いた説明可能な推薦システムにおける代表的なアルゴリズムに焦点を当てて解説する。具体的には、知識グラフの埋め込み表現に基づく手法と、本研究で用いるパス推論に基づく手法の2つに分けて紹介する。

2.2.1 埋め込み表現に基づく手法

知識グラフの埋め込み表現に基づく手法は、エンティティやリレーションを低次元のベクトル空間に埋め込み、その構造を維持することで推薦精度と説明可能性の向上を目指した手法である。このアプローチでは、ユーザ、アイテム、属性などをベクトルで表現し、ベクトル間の類似度に基づいて推薦が行われる。代表的な手法として、異種情報を統合する HeteroEmbed (Heterogeneous Knowledge Base Embeddings) [18] と、Graph Attention Networks (GAT) [19] を知識グラフに応用した KGAT (Knowledge Graph Attention Networks) [20] が存在する。

HeteroEmbed は、ユーザの行動履歴、アイテムの属性、ユーザレビューのコメントなど、異なる種類の情報を1つに統合した知識グラフを利用する。この手法では、異種情報を低次元のベクトル空間に埋め込むことで、エンティティ間の複雑な関係をベクトルの位置関係として表現する。推薦理由の説明は、知識グラフ上でユーザとアイテム間のパスをソフトマッチングすることで提供される。このソフトマッチングは、ベクトル間の距離や類似度を計算することで、どのアイテムがユーザの好みに最も適合するかを判断するプロセスである。HeteroEmbed の重要な特徴は、異種情報を統合的に扱いながら、推薦の説明を直観的に提供できる点にある。

KGAT は、ユーザとアイテムのインタラクションと、アイテムの属性やその他の関連知識を統合した協調知識グラフ上で動作する。この手法は、GAT の Attention メカニズムを使用して、知識グラフ内の各エンティティが他のエンティティとどれだけ関連しているかを評価し、推薦を行う。Attention メカニズムは、高次の関係性を持つノード（エンティティ）からの情報をより重視するように、重みを調整する。KGAT の推薦理由は、Attention の重みを通じて提供され、これらはどのエンティティが推薦の決定に大きく寄与したのかを示す。KGAT は高次の関係性を直接的にモデル化することで、推薦の説明可能性と精度を同時に追求している。

これらの手法は、知識グラフを用いた説明可能な推薦システムの発展に寄与しているが、提供される説明はポストホック（事後的な）アプローチに依存しており、生成される説明が必ずしも推薦プロセスの直接的な洞察を提供するわけではない。ポストホックな説明は、モデルが既に出した推薦に対して、後から理由をつける形で説明を生成するため、根本的な説明可能性には限界が存在する。

2.2.2 パス推論に基づく手法

パス推論に基づく手法は、知識グラフ上のパス探索によって、推薦理由を直接的に提供することに焦点を当てている。このアプローチでは、埋め込み表現を補助的に

利用しながらも、ユーザとアイテム間の具体的な関係性をパスとして示すことで、より直接的な説明を実現する。ここでは、強化学習を用いてパス探索を行う PGPR (Policy-Guided Path Reasoning) [21], ユーザプロファイルに基づく段階的な推論を行う CAFE (Coarse-to-Fine Neural Symbolic Reasoning) [22], パスを言語モデルで学習する PLM-Rec (Path Language Modeling) [23] の3つの手法を説明する。

PGPR は、強化学習を用いて知識グラフ上でのパスを探索する手法であり、ユーザからアイテムへの最適なパスを発見する。このプロセスでは、強化学習エージェントがユーザの現在の状況を基に次に進むべきリレーションを選択し、関連性の高いアイテムに到達することで報酬を獲得する。パス探索では、ユーザからアイテムへの複数の可能性を検討し、各パスの確率を算出する。最終的に、最も確率の高いパスを選択し、そのパスを通じてアイテムが推薦された理由をユーザに提示する。本研究で利用する推薦アルゴリズムであるため、詳細は2.4節で述べる。

CAFE は、ユーザの行動履歴を基にして初期の粗いユーザプロファイルを生成し、このプロファイルをガイドとして知識グラフ上で詳細なパスを探索する。この初期プロファイルは、ユーザの過去の行動から派生したパターンを集約したものであり、関連アイテムにつながる可能性の高いパスを特定するのに役立つ。そして、このプロファイルを活用して、効率的かつ効果的にパスを探索し、最終的に推薦アイテムとその説明パスを提供する。提案されたニューラルシンボリック推論を活用してパスを推論することにより、ユーザプロファイルに最も適合するパスを識別する。この段階的な推論プロセスにより、より洗練された推薦理由が提供可能となる。

PLM-Rec は、知識グラフ上のユーザとアイテム間のパスを言語的なシーケンスとしてモデル化し、言語モデルを使用して新しいパスを生成する手法である。このアプローチでは、Transformer [24] のデコーダー層を利用して、パスの生成に確率的な手法を適用する。このプロセスにより、既存の知識グラフの制約を超えて新たなパスを発見することが可能であり、推薦システムの探索範囲と潜在的な推薦の多様性を大幅に拡張する。一方、PLM-Rec が生成するパスは、時にユーザにとって未知のものであるため、提供される推薦理由としての正確性に課題が残る可能性がある。

これらのパス推論に基づく手法は、推薦理由を説明パスとして明示的に提示することによって、説明の直接性を強化している。埋め込み表現に基づく手法と比較して、より具体的かつ直観的に理解できる説明をユーザに提供することが実現可能である。

2.3 知識グラフを用いた推薦の説明に着目した研究

知識グラフを活用した推薦システムにおける説明の質に焦点を当てた研究はまだ発展途上であり、本節ではこの分野について2つの研究を取り上げる。既存の研究が限られている中で、これらの事例は説明可能な推薦システムの設計と評価における重要な洞察を提供する。

Iván らによる研究 [25] は、ユーザレビューのコメントから抽出された評価とアスペクトベースの意見を組み込んだ知識グラフの構築方法を提案している。このア

アプローチでは、ユーザ、アイテム、アスペクト間の関係をグラフ構造で表現し、グラフ埋め込み技術を用いてこれらを同一の潜在空間にマッピングすることにより、従来の推薦システムよりも詳細な情報を提供し、推薦されたアイテムに対するユーザの意見や感情（例えば「好き」「嫌い」「無関心」など特定のアスペクトに対する感情）を基にした説明が可能となる。

しかし、この手法にはいくつかの制限も存在する。具体的には、単一のコメントの言及のみでリレーションを構築するため、ユーザの長期的な傾向を考慮していないことや、アスペクト別評価を組み込んでいないことが挙げられる。これにより、確実な情報を捉えることが困難となる。また、知識グラフの構築方法がポストホックな説明に限定され、直観的な説明を提供するパス推論のアプローチには不向きであるという課題も存在する。したがって、推薦アイテムの背後にある論理の理解のしやすさを根本的に向上させるための手法としては、一定の限界があるといえるだろう。

次に、Giacomo らの研究 [26] は、知識グラフを活用した推薦システムの説明において、推薦結果を後処理することで説明の質を最適化する手法を提案している。具体的には、推薦アイテムとその説明パスを、最新性、人気度、多様性の観点から再評価しリランキングを行う。最新性では、ユーザの最近の行動に基づいて関連性の高い説明を提供することを重視する。人気度では、説明パス内で頻繁に利用されるエンティティを優先する。多様性では、推薦アイテムリストの中で異なるタイプの説明パスを提示することを重視し、説明の多様性を高めることを目指す。この手法は、推薦精度を維持しながら、ユーザにとってより質の高い説明を提供することを目的としている。また、異なるデモグラフィックグループ間で説明の質が公平になるように配慮している。

この研究は、本研究の3.3節で提案するアスペクトを重視したモデルの後処理の考え方に影響を与えた。しかしながら、推薦の背景にある論理をより分かりやすい形でユーザに提示するという観点でみると、最新性、人気度、多様性という評価指標は不十分であるといえる。説明の質をさらに深掘りし、ユーザにとって理解のしやすい説明を追求することが、今後の課題として残る。

2.4 PGPR (Policy-Guided Path Reasoning)

本節では、本研究の提案システムで利用する推薦アルゴリズムである PGPR (Policy-Guided Path Reasoning) について詳説する。PGPR は、エージェントが知識グラフ上でパス推論を行うことで、ユーザに適したアイテムを発見することを目指す。具体的には、ユーザとアイテムの間の関係性を知識グラフ上のパスとして定義し、そのパスを探索する問題をマルコフ決定過程 (MDP) として定式化する。この際、強化学習のフレームワークを利用することで、複雑な知識グラフ上での効率的なパス推論を実現し、具体的な推薦理由を提示することが可能になる。以下では、PGPR のアルゴリズムの概要について、定式化および学習とパス推論の観点から説明する。

2.4.1 MDP 定式化

知識グラフ G_R は, エンティティ集合 E とリレーション集合 R によって定義され, $G_R = \{(e, r, e') | e, e' \in E, r \in R\}$ と表される. ここで, エンティティ集合には, ユーザエンティティ $u \in U$ とアイテムエンティティ $i \in I$ が含まれる. U はユーザ集合, I はアイテム集合である.

パスは, エンティティとリレーションの列として定義され, 具体的には, エンティティ e_0 からエンティティ e_k への k -hop パス $p_k(e_0, e_k)$ は, 式2.1のように表される.

$$p_k(e_0, e_k) = \{e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} e_k\} \quad (2.1)$$

ここで, $e_{i-1} \xrightarrow{r_i} e_i$ は, $(e_{i-1}, r_i, e_i) \in G_R$ または $(e_i, r_i, e_{i-1}) \in G_R$ のいずれかを表す. つまり, 一方向のリレーションだけでなく, 逆方向の場合も考慮に入れる.

PGPR のような知識グラフのパスに基づく説明可能な推薦システムが解く問題は, 次のように定義される. 知識グラフ G_R , 推薦対象ユーザ $u \in U$, 整数 K と N が与えられたとき, 各ペア (u, i_n) が1つの推論された説明パス $p_k(u, i_n)$ ($2 \leq k \leq K$) に関連付けられる長さ N の推薦アイテムのリスト $\{i_n\}_{n \in [N]} \subseteq I$ を特定する.

PGPR では, この問題を以下の MDP としてモデル化する.

- 状態 (s_t): ステップ t における状態は, ユーザエンティティ u , 現在のエンティティ e_t , 過去のステップの履歴 h_t のタプルとして定義される. T をある固定期間とすると, 初期状態は $s_0 = (u, u, \emptyset)$, 終端状態は $s_T = (u, e_T, h_T)$ である.
- 行動 (a_t): 状態 s_t における行動空間 A_t は, エンティティ e_t から出る全てのエッジであり, 過去に訪問したエンティティとリレーションは除外される.
- 報酬 (R_t): ユーザに対して, どのアイテムがターゲットであるか事前に分からないため, 終端状態 s_T において, アイテムに到達した場合にのみ報酬を与える. 終端報酬 R_T は, 以下の式2.2で定義される.

$$R_T = \begin{cases} \max(0, \frac{f(u, e_T)}{\max_{i \in I} f(u, i)}) & \text{if } e_T \in I \\ 0 & \text{otherwise} \end{cases} \quad (2.2)$$

ここで, $f(u, i)$ はユーザ u とアイテム i の間のスコアリング関数を表す. この関数は u と i の関連度を意味し, 複数のエンティティとリレーションを統合し, それらの埋め込み表現とパスパターンに基づいて計算される.

- 遷移: 状態 $s_t = (u, e_t, h_t)$ において行動 $a_t = (r_{t+1}, e_{t+1})$ が選択されたとき, 次の状態は $s_{t+1} = (u, e_{t+1}, h_{t+1})$ と一意に決定される.

2.4.2 強化学習とパス推論

PGPR の目的は, 任意の初期ユーザ u に対して, 式2.3で表される期待累積報酬を最大化するための確率の方策 π を学習することである.

$$J(\theta) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{T-1} \gamma^t R_{t+1} | s_0 = (u, u, \emptyset) \right] \quad (2.3)$$

ここで, $\gamma \in [0, 1]$ は割引率である. この問題の解法に関しては, 原論文 [21] で述べられている強化学習アルゴリズムに基づく. パス推論の段階では, 学習された方策を利用して, 推薦対象ユーザからアイテムへのパスを効率的に探索する. 具体的には, ビームサーチを導入し, 行動の確率と報酬によって導かれるパスのサンプリングを行う.

最終的に, ユーザとアイテムのペアに対して, 最も高い生成確率を持つパスが, 推薦理由の説明パスとして選択される. そして, これらの説明パスは, スコアリング関数に基づいてランキングが行われ, 対応するアイテムがユーザに推薦される.

3

提案手法

本章では, 説明可能性の高い推薦システムを実現するための3つの提案手法を述べる. 具体的には, アスペクト情報を活用した知識グラフの構築, アスペクトを重視したモデルの後処理, 説明パスの定量評価指標である AO-Score によって構成される.

3.1 AO-PGPR (Aspect-Oriented PGPR)

本節では, はじめに提案手法およびシステムの全体像を説明する. 次頁の図3.1は, 本研究における提案システムの概観図である. アスペクト情報を重視したシステムであること, 知識グラフのパス推論に基づく説明可能な推薦アルゴリズムとして PGPR [21] を採用したことを踏まえて, 提案システムを AO-PGPR (Aspect-Oriented PGPR) と命名した. パス推論を用いた代表的なアルゴリズムであること, 解釈がより直観的であること, 関連する研究が比較的多いことを総合的に鑑みた結果, PGPR を本研究で用いるに至った.

AO-PGPR では, まずユーザレビューを用いて3.2節で提案する手法によって知識グラフを構築する. 次に, 知識グラフと PGPR を用いて推薦モデルの学習を行う. そして, 新たに推薦対象のユーザ ID と, 旅行の目的地の地域を表す地域 ID の組を入力として, PGPR によって学習済みのモデルを適用したパス推論を行うことで, 推薦アイテムであるホテルのリストとそれに対応する説明パスを出力する. 提案するモデルの後処理を行う場合には, それによって新たな推薦アイテムのリストと説明パスを出力する. モデルの後処理を行う必要性や手法については, 3.3節で詳細に述べる. 最後に, 出力された推薦アイテムやパスを, 従来の推薦精度の指標と3.4節で提案する AO-Score (Aspect-Oriented Path Score) によって, 定量的に評価する.

AO-PGPR

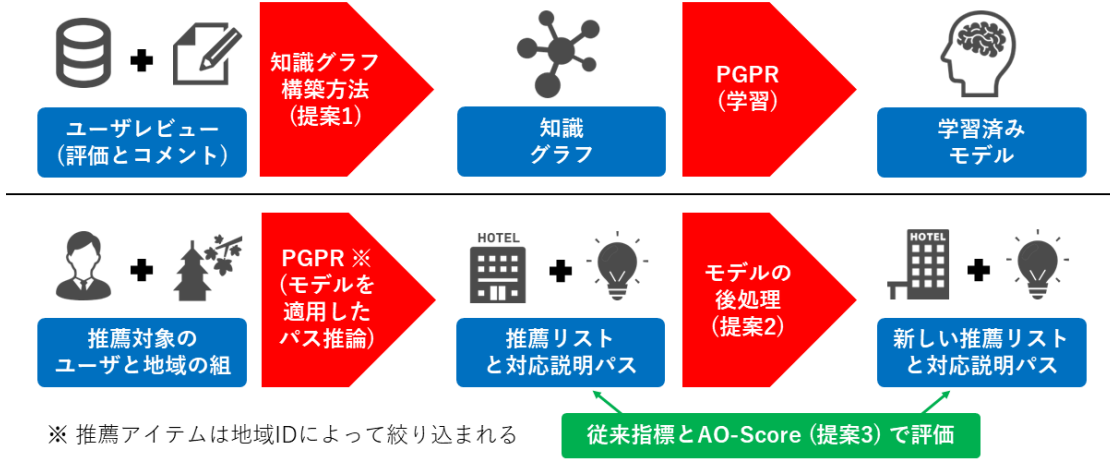


図 3.1: 提案システムの概観図. 青が入力と出力, 赤が処理, 緑が評価を表す.

パス推論の処理において, ユーザ ID に加え地域 ID も入力とする工夫を行っている. ユーザ ID を入力としてパス推論によって出力された推薦アイテムは, 該当の地域 ID に一致しているもののみに絞り込まれる. これにより, ユーザが目的地を決めた上でホテルを選択するという, 現実のシナリオを考慮したシステムを実現する.

3.2 知識グラフの構築

本節では, 提案するアスペクト情報を活用した知識グラフの構築方法を述べる. 次頁の図3.2に概要を示した. 図の知識グラフの例において, 推薦対象ユーザ u_A に対して, パス推論によってアイテム i_C が推薦された場合, 出力される 3-hop 以下の説明パスは 4 種類となる. 詳細は本節で解説する.

3.2.1 エンティティとリレーションの定義

はじめに, 知識グラフを構成するエンティティとリレーションについて定義する. 以下の 3 種類のエンティティを定義する.

- ユーザエンティティ u : 各ユーザを一意に識別するユーザ ID に対応.
- アイテムエンティティ i : 各アイテムを一意に識別するアイテム ID に対応.
- アスペクトエンティティ a : 具体的なアスペクト (立地, 食事, 風呂など) に対応.

また, 以下の 3 種類のリレーションを定義する.

- (過去に) 利用した r_A : ユーザ u からアイテム i へのリレーションに対応.
- 頻繁にコメントする r_B : ユーザ u からアスペクト a へのリレーションに対応.
- 高評価を受けている r_C : アイテム i からアスペクト a へのリレーションに対応.

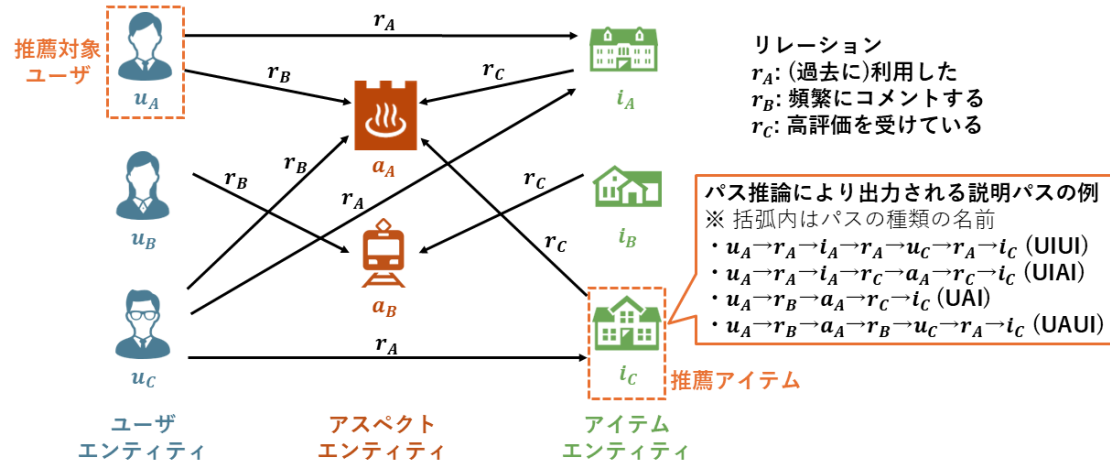


図 3.2: 知識グラフの構築方法. 知識グラフは, 図中の3種類のエンティティと, 3種類のリレーションから構成される. 説明パスは, 図の都合上, 経由するエンティティとリレーションを順番に列挙することで表現.

3.2.2 トリプルの構築

定義したエンティティとリレーションを組み合わせることで, アスペクト情報を含んだユーザレビューからトリプルを構築する. 以下に, 具体的な方法を述べる.

1. ユーザとアイテムの関連トリプル (u, r_A, i):
 ユーザ u とアイテム i 間の過去のインタラクションが存在する場合にトリプルを構築する. つまり, ユーザレビューの行数がトリプルの数に対応する.
2. ユーザとアスペクトの関連トリプル (u, r_B, a):
 ユーザ u がアスペクト a について, コメント内で頻繁に言及している場合にトリプルを形成する. ここで, ユーザ u がアスペクト a に言及した回数をそのユーザの全レビュー数で割った値を「言及割合」と定義する. 各アスペクトに対して, 全ユーザの言及割合のヒストグラムを作成し, 特定の上側パーセント点 θ_{r_B} を超える割合を持つユーザとアスペクトのペアに対してトリプルを構築する. なお, アスペクトの言及は大規模言語モデルを使用して特定される. この過程で, アスペクトに対するユーザの持続的な関心が明らかとなる.
3. アイテムとアスペクトの関連トリプル (i, r_C, a):
 アイテム i がアスペクト a に関して, アスペクト別評価の中で特に高評価を受けている場合にトリプルを構築する. 具体的には, 各アスペクトに対して, 全アイテムの平均評価点数のヒストグラムを作成し, 特定の上側パーセント点 θ_{r_C} を超える平均評価点数を持つアイテムとアスペクトのペアに対してトリプルを構築する. このプロセスを通じて, アイテムが各アスペクトにおいてどれだけ総合的に優れているかが確認することが可能である.

このように3種類のトリプルを構築することで, 長期的なユーザの趣向や総合的なアイテムの特徴を踏まえた知識グラフの構築が実現する. この工夫の結果, 最終的にユーザにとってより理解のしやすい説明の提供が可能となる.

3.2.3 説明パスの種類

推薦対象のユーザ $u \in U$ に対して, 長さ N の推薦アイテムのリスト $\{i_n\}_{n \in [N]} \subseteq I$ と共に, ユーザと各アイテムをそれぞれ結ぶ k -hop の説明パス $p_k(u, i_n)$ ($2 \leq k \leq K$) が, PGPR のパス推論によって出力される. 本研究では, この説明パスの長さの上限 K を3とする制約を加えた. これは, 4-hop 以上のパスの場合, かえって説明が冗長になってしまうことを考慮したためである.

上記の制約下において, ユーザエンティティ u とアイテムエンティティ i の間の説明パスは図3.2の4種類が存在する. これらのパスは, 経由するエンティティの種類の頭文字を取ることで命名した.

1. UIUI パス: 推薦対象ユーザが以前に利用したアイテムと同様のアイテムを利用した他のユーザが, 過去に利用した別のアイテムを推薦するパターンを表す. 従来のユーザベース協調フィルタリングと類似したパスである.
2. UIAI パス: 推薦対象ユーザが過去に利用したアイテムと同じく, 特定のアスペクトにおいて高評価を得ているアイテムを推薦するパターンを表す. アイテムの優位な点をアスペクトで強調したパスである.
3. UAI パス: 推薦対象ユーザが頻繁にコメントするアスペクトにおいて評価の高いアイテムを推薦するパターンを表す. ユーザの好みとアイテムの優位な点をアスペクトを通じて直接結び付けた 2-hop のパスである.
4. UAUI パス: 推薦対象ユーザと同じアスペクトについて頻繁にコメントする他のユーザが過去に利用したアイテムを推薦するパターンを表す. ユーザの趣向を反映した他ユーザの利用履歴に基づくパスである.

上記の4種類の説明パスのうち, アスペクトエンティティ a を経由する UIAI パス, UAI パス, UAUI パスのグループと, ユーザエンティティ u とアイテムエンティティ i のみで構成される UIUI パスに分類することができる. 前者はユーザやアイテムをアスペクトによって特徴づけることが可能であり, 提供する説明の理解のしやすさが比較的高いと考えられる. 一方, 後者は従来のユーザベース協調フィルタリングと同等の説明の提供に留まるため, ユーザにとって十分に好ましい説明を提供可能であるとはいえない. この分類方法は, 次節以降の提案手法および実験において, 重要な意味を持つ.

3.3 アスペクトを重視したモデルの後処理

本節では, PGPR のパス推論によって推薦対象ユーザに推薦アイテムとそれに対応する説明パスが出力された後の, モデルの後処理方法を提案する. 3.2節において, 3-hop 以内の説明パスには4種類があり, 説明が比較的理解しやすいUIAIパス, UAIパス, UAUIパスと, 理解しにくいUIUIパスに分類することができると述べた. このことを踏まえて, ユーザにとってより理解がしやすい説明を提供するシステムを構築するために, 出力される全説明パスのうち, アスペクトエンティティを含む3種類のパスの占める割合の増加およびUIUIパスの割合の減少を目指す. 本研究では, パス選択とリランキングの2種類のモデルの後処理を行うことによって, これを実現することを提案する.

3.3.1 パス選択

PGPR は, 強化学習によって, 知識グラフ上でユーザの趣向やアイテムの特徴に関するパターンを学習する. 提案するパス選択のプロセスでは, 学習の際に最適化される方策に対応する, 各パスの選択確率を利用する. ユーザエンティティとアイテムエンティティの間のパスが複数存在する場合, それらが選択される確率がモデルによって計算される. 例えば, 推薦対象のユーザ u とアイテム i の間の m 番目のパス $p_k(u, i)^m$ の選択確率 $\pi_m(u, i)$ は, 式3.1で表される.

$$\pi_m(u, i) = \Pr(p_k(u, i)^m \mid u, i) \quad (3.1)$$

パス選択は, 式3.2に従って行われる. ここで, $A_{u,i}$ は, ユーザ u とアイテム i の間のパスのうち, アスペクトエンティティを含むパス (UIAI, UAI, UAUI) の集合を表す. $p_{k,s}(u, i)$ は, 最終的に決定されるパスを表す.

$$p_{k,s}(u, i) = \begin{cases} \arg \max_{p_k(u, i)^m \in A_{u,i}} \pi_m(u, i) & \text{if } A_{u,i} \neq \emptyset \\ \arg \max_{p_k(u, i)^m} \pi_m(u, i) & \text{if } A_{u,i} = \emptyset \end{cases} \quad (3.2)$$

アスペクトエンティティを含むパスが存在する場合, これらのパスの中から最も高い確率を持つパスが選択される. 一方, アスペクトエンティティが含むパスが存在しない場合は, 全ての選択されうるパスの中から最も高い確率を持つパスが選択される. これは, モデルの後処理が行われない通常の PGPR においてパスが決定されるプロセスと同様である.

この提案するパス選択によって, アスペクトエンティティを含むパスが優先的に推薦対象ユーザと推薦アイテムの間の説明パスとして決定される. その結果, アスペクトエンティティを含む説明パスが全ての説明パスに占める割合が増加し, より理解しやすい説明を提供しやすくなることが容易に期待できる.

3.3.2 リランキング

提案されたパス選択だけでは、元々UIUIパスのみによって接続されているユーザエンティティ u とアイテムエンティティ i に対しては、アスペクト情報を活用した説明を出力することができない。そのため、パス選択によって特定のユーザとアイテム間のパスが1つに決定された後に、アスペクトエンティティを経由するパスによって決定されたアイテムを優先的に推薦を行う手法である、リランキングを提案する。これにより、理解しやすい説明のさらなる増加を目指す。

PGPRにおいて、ユーザエンティティ u とアイテムエンティティ i の関連度を示すスコアリング関数 $f(u, i)$ が、MDPにおける報酬の計算と推薦アイテムのランキングに用いられることは、2.4節で説明した。 $f(u, i)$ は0以上1以下の値を取り、値が大きいほどユーザ u とアイテム i の関連性が高いとされるため、推薦アイテムのランキングはこの関数の値が高い順に行われる。本研究では、PGPRのスコアリング関数を拡張し、式3.3で定義されるアスペクトを重視したスコアリング関数 $g(u, i)$ によって、アイテムのリランキングを行う。

$$g(u, i) = (1 - \alpha) f(u, i) + \alpha I_{\text{aspect}}(u, i) \quad (3.3)$$

ここで、 $f(u, i)$ はPGPRで使用されるスコアリング関数を示し、 $I_{\text{aspect}}(u, i)$ は新たに定義したアスペクト指示関数である。アスペクト指示関数は、パス選択によって決定された $p_{k,s}(u, i)$ が、アスペクトエンティティを含むパス（UIAI, UAI, UAUI）である場合は1を、UIUIパスである場合は0を出力する。また、パラメータ α はリランキングを制御する重みであり、 $\alpha = 0$ の場合は元のPGPRのスコアリングに従い、 $\alpha = 1$ の場合はアスペクトエンティティの存在を最優先してアイテムのリランキングを行う。

アスペクト情報の重要性を加味した推薦アイテムの選択を行うリランキングにより、アスペクトエンティティを含む説明パスの割合の増加が実現する。その結果、ユーザにとってより理解しやすい推薦理由の提供が可能となる。

3.4 説明パスの定量評価

本節では、説明パスの定量的な評価指標である AO-Score（Aspect-Oriented Path Score）を提案する。この評価指標により、説明パスが、推薦対象ユーザによる推薦アイテムのアスペクト別評価に対して、どの程度適合しているかを測定することが可能となる。説明パスに対する AO-Score の算出は、次の2条件を満たすケースで行われる。

1. 説明パスがアスペクトエンティティを含むパス（UIAI, UAI, UAUI）である。
2. 説明パスによって推薦されたアイテムが、実際にユーザがインタラクションを行ったアイテム（正解アイテム）である。

$AO - Score(p_k(u, i), a)$ は, ユーザ u がアイテム i に対してインタラクションを行った際に選択される正解の説明パス $p_k(u, i)$ において, アスペクト a を経由する場合の AO-Score を示す. AO-Score の計算には, 式3.4の標準化関数 S が使用される.

$$S(x, \mu, \sigma) = \frac{x - \mu}{\sigma} \quad (3.4)$$

ここで x は評価点数, μ は評価点数の平均, σ は標準偏差である. ユーザ u が過去にインタラクションを行った全アイテムにおけるアスペクト a の評価をこの関数によって平均が 0, 標準偏差が 1 となるように標準化する. 式3.5によって, 対象の説明パス $p_k(u, i)$ の AO-Score は, アイテム i のアスペクト a に対する評価点 $x_{u,i,a}$ の標準化スコアとして計算される. 過去のユーザ u のアスペクト a に対する評価点数において, $\mu_{u,a}$ はその平均, $\sigma_{u,a}$ は標準偏差を表す.

$$AO - Score(p_k(u, i), a) = S(x_{u,i,a}, \mu_{u,a}, \sigma_{u,a}) \quad (3.5)$$

このスコアが高いほど, ユーザ u はアイテム i のアスペクト a に対して, 実際により高い満足度を示しているといえる. つまり, AO-Score が高い場合, 説明パスがユーザの好みやアイテムの特性を正確に捉えていることが考えられる.

4 実験

本章では, 実際のホテル予約サービスにおけるユーザレビューを用いて, 提案システムの効果を検証する. 具体的には, ベンチマークとなる従来の行列分解ベースの手法と, 伝統的な評価指標を用いて推薦精度を比較する. また, アスペクトを重視したモデルの後処理の影響を, パスごとのカバー率や AO-Score, 推薦精度の変化から検証する. 最後に, 実際に提供された説明文によるケーススタディを行い, 提案する知識グラフの構築方法やモデルの後処理の効果を定性的に確認する.

4.1 実験設定

本節では, 実験に使用するデータセット, 評価指標, そしてベンチマークとなる他の推薦システムについて述べる.

4.1.1 データセット

本研究では, ホテルのユーザレビューとして, 楽天グループ株式会社の楽天トラベルデータ [9] を使用する. ユーザレビューは, ユーザ ID, 投稿日時, アイテム ID, 6 種類のアスペクト別評価, 日本語のコメントで構成されている. アスペクトは, それぞれ立地, 部屋, 食事, 風呂, サービス, 設備であり, 評価は最高が 5, 最低が 1 の整数による 5 段階である. また, アイテム ID とホテル名は対応づけが可能であり, 提供された説明文によるケーススタディの際に用いる.

データの前処理として, はじめに以下の 3 ステップでフィルタリングを行う. まず, ユーザレビューの数に対応するインタラクションの数が 20 以上のユーザおよび

表 4.1: フィルタリング後のインタラクションデータの集計.

集計項目	集計値
総インタラクション数	993,423
ユニークユーザ数	34,609
ユニークアイテム数	9,218
最大インタラクション数 (ユーザごと)	375
最大インタラクション数 (アイテムごと)	1,804
平均インタラクション数 (ユーザごと)	28.7
平均インタラクション数 (アイテムごと)	107.8

アイテムのユーザレビューのみを抽出し、インタラクションの少ないユーザやアイテムを除外する。その後、その中からコメントの文字数が10以上のユーザレビューを抽出する。最後に、同一ユーザが同一ホテルに対して複数回インタラクションを持つ場合、ランダムに1件のユーザレビューを選択し、重複を避ける。このようにして行われたフィルタリング後のデータを表4.1に集計した。

次に、フィルタリング後のユーザレビューを時系列順に整理し、全体を訓練データ、検証データ、テストデータの3部に分割する。具体的には、全ユーザレビューの8割を訓練データとして使用し、残りの2割を等分して検証データとテストデータに割り当てる。この分割により、モデルの学習と評価が実際の推薦システムにおけるシナリオを反映させて、適切に行われることを保証する。

そして、3.2節の提案手法に基づき、訓練データから知識グラフを構築する。コメント内で言及されたアスペクトの特定には、日本語大規模言語モデルの「tokyotech-llm/Swallow-13b-instruct-hf」[27]を使用する。詳細は付録Aに記載した。また、トリプル構築の際の θ_{r_B} は10、 θ_{r_C} は20とする。構築後の知識グラフに関する集計値は、次頁の表4.2にまとめた。パスの集計値の括弧内は、全パスに占める割合を示す。

推薦対象のアイテムを地域で絞り込む工程で使用する地域IDは、楽天トラベル¹の実際の地域分割方法に準拠し、Webスクレイピングを行うことで決定する。地域IDの数は323、1地域あたりの平均アイテム数は28.5となった。また、PGPRで学習とパス推論の際に使用する知識グラフのエンティティとリレーションの埋め込み表現の取得には、代表的なアルゴリズムであるTransE[28]を利用する。

4.1.2 評価指標

推薦精度の評価には、Normalized Discounted Cumulative Gain (NDCG), Hit Ratio (HR), Precision, および Recall の4指標を使用し、推薦リストの大きさ N を5に設定して、リストの上位5アイテムに対する評価を行う。いずれの指標も0以上1以下の値を出力し、値が大きいほど性能が良いことを示す。これらの伝統的な4指標を組み合わせることで、システム全体の推薦精度を総合的に評価することができる。

¹<https://travel.rakuten.co.jp/>

表 4.2: 知識グラフのエンティティ, リレーション, パスの集計.

カテゴリ	集計項目	集計値
エンティティ	ユーザ (u)	34,609
	アイテム (i)	9,218
	アスペクト (a)	6
リレーション	r_A	794,737
	r_B	20,392
	r_C	10,427
パス	UIUI	6,513,024,620 (68.9%)
	UIAI	1,452,500,280 (15.4%)
	UAI	35,439,776 (0.4%)
	UAUI	1,453,004,038 (15.4%)

NDCG@5 はランキングの品質を評価するために使用され, 推薦リストの上位アイテムがユーザにとってどれだけ価値があるかを計測する. 具体的には, 各アイテムに対してゲイン (正解アイテムはゲイン 1, 不正解アイテムはゲイン 0) を割り当て, それをランクに基づいた割引率 $\frac{1}{\log_2(\text{rank}+1)}$ で割り引いた値の累積和である Discounted Cumulative Gain (DCG) を計算する. そして, DCG を理想的なゲイン順序で得られる最大の DCG である IDCG で正規化する. HR@5 は推薦リスト内に少なくとも 1 つの正解アイテムが含まれるケースの割合を示す. リストに 1 つでも正解アイテムが存在すればカウントされ, 全ユーザ (本研究では全ユーザと地域の組み合わせ) に対する割合として表される. Precision@5 は, 推薦リストの 5 アイテム中, 正解アイテムの数の割合を示し, 推薦されたアイテムのうちどれだけが実際にユーザにとって価値あるかを測る. Recall@5 は, 全ての正解アイテムにおいて, 推薦リストにどれだけの割合で現れるかを示し, 実際の正解のアイテムのうち推薦されたアイテムがどれだけカバーをしているかを測る.

また, 3.3 節で提案したアスペクトを重視したモデルの後処理の影響を測定するために, 従来の評価指標に加えて, パスごとのカバー率と AO-Score を計算する. パスには 3.2 節で述べたように, 3-hop 以下の制約では, UIUI, UIAI, UAI, そして UAUI の 4 種類が存在する. パスごとのカバー率は, 推薦の際に説明パスとして出力された全パスのうち, 各パスが占める割合を示す. モデルの後処理の目的は, アスペクトエンティティを含むパスの割合を増加させ, より理解しやすい説明を提供することである. このことを踏まえると, UIUI のカバー率がより減少し, それ以外のパスのカバー率がより増加することが望ましい. さらに, 3.4 節で提案した AO-Score を計算することで, 正解アイテムにおけるアスペクトエンティティを含む説明パスの質を定量的に評価する. AO-Score の値が大きいほど, 実際にユーザが正解アイテムにインタラクションを行った際に, 説明パス上のアスペクトに対してより満足していることを示す.

表 4.3: リレーシンの組み合わせが推薦精度に与える影響.

リレーション	NDCG@5	HR@5	Prec.@5	Recall@5	(ユーザ, 地域)
r_A	0.1849	0.2899	0.0611	0.2564	71,869
r_A, r_B	0.1985	0.3072	0.0650	0.2734	71,969
r_A, r_C	<u>0.1877</u>	<u>0.2947</u>	<u>0.0621</u>	<u>0.2611</u>	71,955
r_A, r_B, r_C	0.1830	0.2876	0.0605	0.2546	72,006

各評価指標において、最も優れた結果は太字で、次点の結果は下線で表示.

4.1.3 ベンチマーク

推薦精度のベンチマークとして、2.1節で説明した従来の行列分解ベースの推薦手法である IMF [16] と BPR [17] を採用する. これらは、説明可能性よりも推薦精度の向上を主目的とした手法で、幅広い推薦タスクにおいて広く利用されている. 本研究では、これらの既存手法と、説明可能性を重視した提案システムである AO-PGPR との比較を通じて、推薦精度の面での性能評価を行う.

4.2 実験結果

本節では、実験結果を述べる. 具体的には、提案システムの推薦精度に関して 2 種類の予備実験とベンチマークとの比較、複数の定量的な評価指標によるアスペクトを重視したモデルの後処理の影響の検証、そして提供された説明文による定性的なケーススタディで構成される.

4.2.1 推薦精度

まず、予備実験として AO-PGPR においてリレーシンの組み合わせが推薦精度に与える影響を検証する. r_A を除いた場合、知識グラフがスパースとなり推薦精度が期待できないため、 r_A は必ずリレーションとして採用する. また、知識グラフ構築の際のパラメータである θ_{r_B} と θ_{r_C} は初期的にいずれも 20 を採用する. リレーションに関する予備実験の結果を表 4.3 に示した. 表における、(ユーザ, 地域) 列の数値は、推薦対象のユーザと地域の組み合わせの数を指し、次の予備実験の表においても同様である. 少なくとも 5 個のアイテムが推薦されない場合は、各評価指標の計算が行われないため、モデルに入力する知識グラフによって (ユーザ, 地域) 列の数値が変動する. 全推薦対象のユーザと地域の組み合わせの数は 72,786 であり、これは行列分解ベースの手法における数値と一致する. リレーションに関する予備実験の結果より、 r_A と r_B を用いたケースが全体として最も高い推薦精度を記録したことが分かる. 一方、全てのリレーションを用いた場合は最も推薦精度が低くなったが、4 種類の説明パスを提供できるため、説明可能性の高さを鑑みると、今後の実験も同様のケースを採用する.

表 4.4: パラメータ θ_{r_B} が推薦精度に与える影響.

θ_{r_B}	NDCG@5	HR@5	Prec.@5	Recall@5	(ユーザ, 地域)
10	0.1992	0.3080	0.0650	0.2736	71,951
20	<u>0.1985</u>	<u>0.3072</u>	0.0650	<u>0.2734</u>	71,969
30	0.1963	0.3044	0.0643	0.2705	71,996

各評価指標において, 最も優れた結果は太字で, 次点の結果は下線で表示.

表 4.5: パラメータ θ_{r_C} が推薦精度に与える影響.

θ_{r_C}	NDCG@5	HR@5	Prec.@5	Recall@5	(ユーザ, 地域)
10	<u>0.1846</u>	<u>0.2880</u>	<u>0.0609</u>	<u>0.2557</u>	71,939
20	0.1877	0.2947	0.0621	0.2611	71,955
30	0.1825	0.2857	0.0602	0.2531	71,940

各評価指標において, 最も優れた結果は太字で, 次点の結果は下線で表示.

AO-PGPR において, 3.2 節で説明したトリプル構築の際のパラメータである θ_{r_B} , θ_{r_C} の推薦精度への影響を表 4.4 と表 4.5 にそれぞれ示した. パラメータには, それぞれ 10, 20, 30 を採用した. パラメータの変化による影響を直に図るため, 全てにおいて他のリレーションは r_A のみを採用している. パラメータの値が大きくなるほど, 知識グラフ内の該当のリレーションの数は多くなる. 予備実験の結果から, どちらのパラメータの値も推薦精度に対して, 大きく影響を与えていないことが分かる. また, $\theta_{r_B} = 10$ と $\theta_{r_C} = 20$ のときにそれぞれ最高の精度を記録した. 以上を踏まえて, 今後の本研究の実験において, θ_{r_B} は 10, θ_{r_C} は 20 を知識グラフ構築の際のパラメータとして採用する.

提案した AO-PGPR とベンチマークの手法の推薦精度を比較した結果を次頁の表 4.6 に示した. AO-PGPR の結果はリランキングを行う前のものであり, リランキングの影響については後述する. また, 表のスペースの都合により記載できなかったが, AO-PGPR におけるユーザと地域の組み合わせの数は 71,995 であり, 行列分解を用いた手法における組み合わせの数である 72,786 に対して, 約 1,000 ケースほど不足する結果となった. 各評価指標において, ある 1 指標の値が高い場合, 他の指標の値も同様に高い傾向が見られ, 指標間での結果の入れ替わりは観察されなかった. AO-PGPR は, 行列分解ベースの手法である BPR を上回るパフォーマンスを示した. しかし, IMF と比較すると, AO-PGPR の推薦精度は若干下回った. 全体的にみると, 提案システムはより理解しやすい推薦理由を提供でき (本節のケーススタディを参照), 説明可能性に優れたシステムでありながら, 推薦精度においては従来のベンチマークと同等またはそれ以上の結果を得ることができたといえる. モデルの精度と説明可能性が通常トレードオフの関係にあることを踏まえると, これは十分妥当な結果であるといえる.

表 4.6: 推薦精度の比較.

手法	タイプ	説明可能性	NDCG@5	HR@5	Prec.@5	Recall@5
IMF	行列分解	低い	0.2055	0.3133	0.0664	0.2785
BPR	行列分解	低い	0.1734	0.2758	0.0580	0.2455
AO-PGPR	提案手法	高い	<u>0.1831</u>	<u>0.2878</u>	<u>0.0606</u>	<u>0.2549</u>

各評価指標において、最も優れた結果は太字で、次点の結果は下線で表示.

AO-PGPR が IMF の推薦精度に若干及ばなかった理由について分析する. IMF は行列分解を用いることで、価格帯や利用目的などの潜在的な要因を効果的に考慮することが可能である. これに対し、提案された知識グラフではこれらの要因を十分に取り込んでいないため、情報が不足していることが考えられる. さらに、推薦アルゴリズムとして利用した PGPR では、ユーザとアイテム間のパスが推薦の根拠となるが、3-hop 以内にパスが見つからないアイテムは推薦候補から外れる. このため、地域内で利用可能な全アイテムを網羅的に推薦することが難しいという、行列分解手法との明確な違いがある.

また、計算時間については、AO-PGPR は行列分解ベースの手法と比較して大幅に増加する. 行列分解では解析的に解を求めることができ、短時間での最適化が可能であるが、PGPR では知識グラフ上の膨大なパスを探索する必要があるため、学習や推論に長時間を要する. これは実運用における効率性の課題であるといえる.

4.2.2 パスごとのカバース率の変化

3.3節で提案した、アスペクトを重視したモデルの後処理の効果を検証する. 具体的には、アスペクトエンティティを含むパスを優先的に説明パスとして決定するパス選択と、同様のパスによって決定されたアイテムを優先的に推薦リストの上位アイテムとするリランキングの2種類である. これらの後処理を行うことで、アスペクトエンティティを含むパスのカバース率を増加させ、結果的にアスペクトを含んだより理解しやすい説明の増加を目指す.

次頁の図4.1に、推薦において出力されたパスごとのカバース率の変化を示した. 本実験において、リランキングは、パス選択を行った後に行われた. また、リランキングにおいて、 $\alpha = 0.5$ 以降はカバース率の変化が小さかったため、 $\alpha = 0.5$ の実験結果を採用した. アスペクトエンティティを含む3種類の説明パス (UIAI, UAI, UAUI) のカバース率の合計は、元のモデルの 4.3% から、パス選択後には 51.7% に、さらにリランキング後には 64.2% にまで増加した. この結果から、提案したモデルの後処理は、アスペクトを活用したより理解しやすい説明の提供に対して有効であるといえる. 一方、UAI パスのカバース率に関しては、全てのケースにおいて、5% 未満となっている. これは、表4.2において、UAI パスの元の知識グラフにおけるカバース率が 0.4% で非常に小さい値であることが原因だと考えられる.

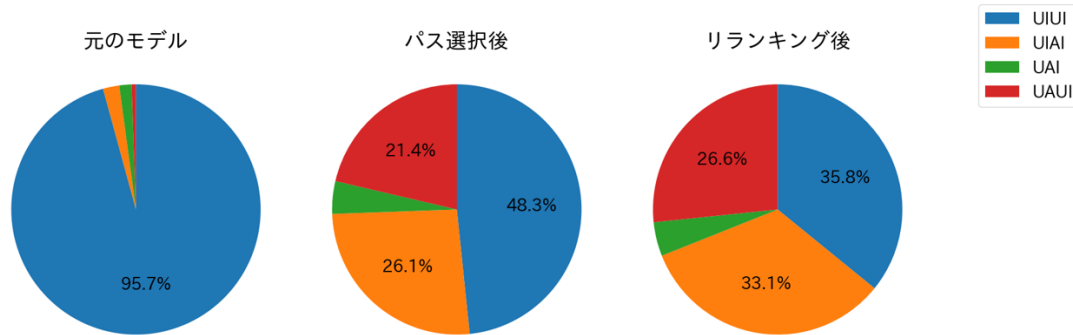


図 4.1: モデルの後処理におけるパスごとのカバー率の変化. リランキング後は, $\alpha = 0.5$ の場合. 5%未満の割合の表示は, スペースの都合上省略.

4.2.3 AO-Score の比較

3.4節で提案した, 説明パスの定量的な評価指標である AO-Score のリランキングにおける比較, およびパスごとの比較を次頁の図4.2に示した. AO-Score はアスペクトエンティティを含むパスのみに対して算出されるため, UIUI パスの結果は存在しない. また, パス選択前の元のモデルにおいて, UIUI パスが出力される説明パスの 9 割以上を占めることにより, 算出の対象とするケースが極端に少なくなるため, 実験結果から除外した. 全体 (Overall) およびパスごとの結果を見ると, α の値の変動による AO-Score の顕著な変化は見られなかった. これは, リランキングが説明パスの品質に大きな影響を与えていないことを示す. しかし, パスによっては AO-Score に差が存在することが確認される. 特に, サンプル数が少ないといえど, UAI パスが 0 に近い値を示しているのは問題である. これは, ユーザの評価が平均付近のアスペクトのエンティティを含むパスが提供されてしまっていることを意味する.

パスによる AO-Score の差異について, 特にユーザとアスペクトの関連トリプル (u, r_B, a) がユーザの趣向を十分に反映していない可能性が原因として考えられる. 3.2節で説明した通り, このトリプルはユーザのコメントから大規模言語モデルを用いて特定されたアスペクトを基に構築される. しかし, 本研究ではこのアスペクト特定タスクの精度向上には焦点を当てておらず, 結果として (u, r_B, a) の精度も低下している可能性がある. これが, (u, r_B, a) を含む UAU パスおよび UAI パスの AO-Score が相対的に低くなっている一因であると考えられる.

以上より, AO-Score による評価がもたらす洞察の価値は明らかである. この指標を用いることで, 説明パスがユーザのアスペクトに対する適切性を定量的に捉えることが可能となる. これにより, システムが提供する説明パスがユーザにとって本当に有益であるかを確認し, 必要に応じて改善を行うことができる.

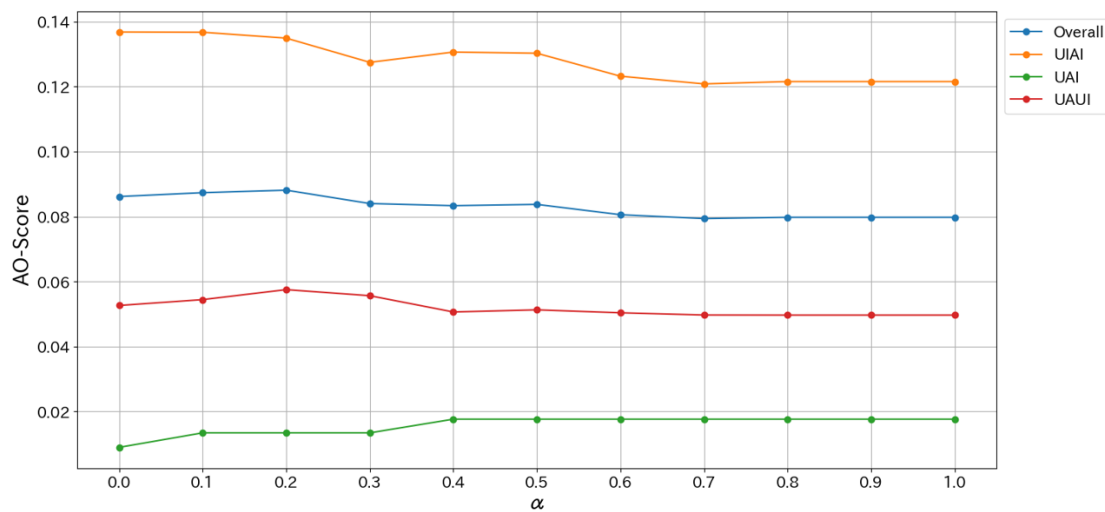


図 4.2: AO-Score の比較. パスごとの結果の重みづけ平均が Overall である.

4.2.4 リランキングの推薦精度への影響

アスペクトを重視したモデルの後処理のうち, リランキングはアスペクトを利用した説明の是非によって推薦リスト内のアイテムを並び替える. これによって, アスペクトエンティティを含むパスが増加し説明の理解のしやすさが向上したとしても, 肝心の推薦精度が大幅に悪化してしまう場合, 本末転倒である.

以上を踏まえて, リランキングを制御するパラメータ α による評価指標ごとの推薦精度の変化を, 次頁の図4.3に示した. 一般的に, リランキングにおいて α の値が増加すると, 式3.3によりユーザとアイテムの関連度よりパス内のアスペクトエンティティの有無が重視されるため, 推薦精度は下がるのが想定できる. 結果として全ての指標において, 想定通りとなった. これにより, PGPR が適切にユーザとアイテムの特徴を捉えていることを改めて確認することができる. しかし, 図をより詳細に観察すると興味深いことに, NDCG@5 と Precision@5 では $\alpha = 0.2$ のとき, HR@5 と Recall@5 では $\alpha = 0.3$ のときに最高の推薦精度を記録したことが分かる.

$\alpha = 0$ ではないときに, それぞれの指標において最高精度を記録した要因として, アスペクトがユーザの趣向やアイテムの特徴を適切に捉えた可能性が考えられる. このことから, リランキングはアスペクトエンティティを含むパスを増加させるだけでなく, 推薦精度自体を向上させる可能性を持っているという示唆が得られた. この結果を踏まえて, システム提供者は, アスペクトエンティティを含むパスの増加と推薦精度の維持のバランスを考えて, リランキングを制御するパラメータ α を決定する必要があるといえる.

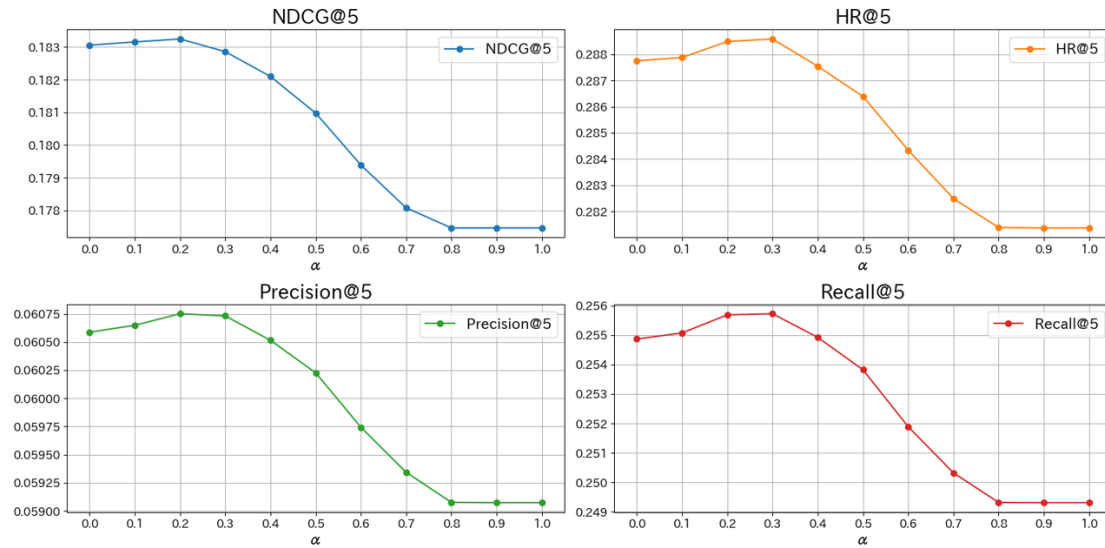


図 4.3: リランキングによる推薦精度の各評価指標の変化.

表 4.7: 説明文のテンプレート.

説明パス	説明文
UIUI	あなたと同様に $\langle i_0 \rangle$ を利用したユーザが利用した
UIAI	あなたが利用した $\langle i_0 \rangle$ と同様に $\langle a_0 \rangle$ について高評価を受けている
UAI	あなたが頻繁にコメントする $\langle a_0 \rangle$ について高評価を受けている
UAUI	あなたと同様に $\langle a_0 \rangle$ について頻繁にコメントするユーザが利用した

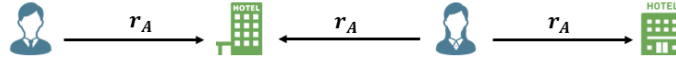
4.2.5 ケーススタディ

最後に, AO-PGPR によって提供された推薦理由の説明文のケーススタディを行う. 説明パスを説明文へと変換する際に, 表4.7に示された4種類のパスそれぞれに対応するテンプレートを使用する. テンプレート内のプレースホルダである i_0 や a_0 は, 説明パスの中ではじめに登場するアイテムやアスペクトのエンティティを表し, これらに具体的なエンティティの名前を挿入することによって, 推薦理由の説明文が生成される. この方法を用いることで, ユーザに対して直観的かつ具体的な推薦理由を提供することができる.

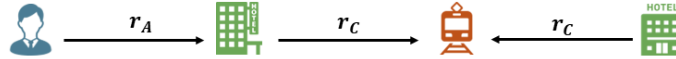
2種類のケーススタディの結果を図4.4に示した. フィルタリング後のインタラクション数が最多である「品川プリンスホテル」と, 3番目に多い「天然温泉 加賀の湧泉 ドーミーイン金沢」が, それぞれ別のあるユーザに推薦されたケースを対象とする. 楽天トラベルの対象ホテルのページによると, 前者は東京都品川駅前の便利な場所にある高級ホテルで, 後者は石川県金沢駅前唯一の天然温泉大浴場について評判が良いホテルである. 本研究では, 説明文の提示に加え, パス選択前後における比較も行う.

ケース1: 品川プリンスホテルの推薦理由

<パス選択前> あなたと同様にホテルJALシティ 関内 横浜を利用したユーザが利用したホテル

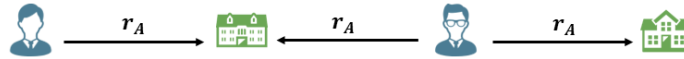


<パス選択後> あなたが利用した京急EXイン 京急川崎駅前と同様に立地について高評価を受けているホテル



ケース2: 天然温泉 加賀の湧泉 ドーミーイン金沢の推薦理由

<パス選択前> あなたと同様にホテルアクティブ！博多を利用したユーザが利用したホテル



<パス選択後> あなたと同様に風呂について頻繁にコメントするユーザが利用したホテル

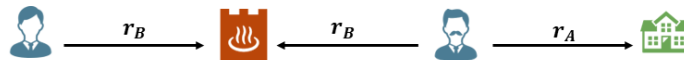


図 4.4: 説明文によるケーススタディ. タイプの異なる2種類の推薦ホテルを選出. 出力された説明文および対応する説明パスをパス選択前後で比較.

ケース1では, パス選択前は図中のUIUIパスによって, 「あなたと同様にホテルJAL シティ 関内 横浜を利用したユーザが利用したホテル」という説明が提供された. この説明では, 「ホテルJAL シティ 関内 横浜」と「品川プリンスホテル」の関係が曖昧であり, ユーザにとって理解しにくい. 一方, パス選択後は図中のUIAIパスによって, 「あなたが利用した京急EX イン 京急川崎駅前と同様に立地について高評価を受けているホテル」という説明に変更された. この場合, アスペクト「立地」が, 2つのホテルを結びつけ, 理解しやすい説明となっている. ケース2においても, ケース1と同様に, パス選択前はUIUIパスを通じて分かりづらい説明が提供されていたが, パス選択後にはUAUIパスを用いることで, 「風呂」というアスペクトを活用し, より理解しやすい説明を実現したことが確認できる.

実際のケースの観察によって, アスペクトが推薦の基準となり, ユーザにとってより直観的に理解しやすい説明を出力可能なことが明らかとなった. ユーザが重視するアスペクトと推薦ホテルの関連の理解が容易となっていることが, 説明全体の分かりやすさに繋がっていると考えられる. 本研究では, アスペクトに注目した知識グラフの構築とモデルの後処理を提案した. 双方の提案手法が説明の理解のしやすさを高めるための重要な手段として有効であることが, ケーススタディによって定性的に示されたといえる.

5

考察

本研究の課題について、技術的及び有益性の観点から考察を行う。

本システムは、従来の行列分解ベースの手法である IMF に対して若干の推薦精度の劣勢が見られた。推薦精度の向上のためには、アイテムの価格帯やユーザの利用目的などの情報をより効果的に知識グラフに組み込むことが必要であると考えられる。また、知識グラフ構築に利用する大規模言語モデルによるコメント内のアスペクトを特定するタスクの精度を追求することで、推薦精度の向上が期待できる。さらに、異なる知識グラフ構築手法や関連研究で述べた別のアルゴリズムを使用して得られる結果との比較を行い、システムの精度を評価することも重要である。地域によるアイテムの絞り込み処理については、現在のシステムでは計算効率が低く、処理時間が長いという課題がある。地域絞り込みを推薦システムのより前段階で実施することにより、全体の計算速度を向上させることが可能だと考えられる。

システムの有益性の観点では、出力された説明文がユーザにとって実際に理解しやすいかどうかの検証が不足しており、アンケート調査や A/B テストによるオンライン評価を通じて、定量的に評価されることが必要である。また、複数の説明パスを提供し、言語モデルを用いてこれらを効果的に集約するアプローチも検討する必要性もある。1つの推薦アイテムに対して単に1つのパスを提供するよりも複数のパスの集約結果を提供する方が、ユーザがより豊かで理解しやすい情報を得られる可能性があると考えられるためである。

6

結論

本研究では, 説明可能なホテル推薦システムとして, 知識グラフ上のパス推論を利用した AO-PGPR を開発した. このシステムにおいて, 推薦理由の理解のしやすさの向上のためにアスペクト情報を活用する新しい知識グラフの構築方法を提案した. さらに, アスペクトエンティティに着目し, モデルの後処理を通じてユーザにとって理解しやすい説明パスを増やす工夫を行った. また, 説明パスの質を定量的に評価する新しい指標である AO-Score の提案も行った.

実験では, ホテル予約サービスのユーザレビューを使用し, 説明可能性に優れた AO-PGPR が推薦精度においても, 精度向上を重視した従来の行列分解ベースの手法と同等のパフォーマンスを記録した. そして, モデルの後処理を通じて効果的に出力される説明パスの種類を変更できること, AO-Score によってパスの品質を定量評価できること, リランキングが推薦精度に影響を与えることを, それぞれ確認した. 最後に, ケーススタディを通じて, 説明パスと説明文がどのように直観的な理解のしやすさを向上させるのかを, パス選択の前後で比較し検証した.

本研究の貢献は, ユーザレビューのアスペクト情報に注目することによって, 理解のしやすい説明を提供するホテル推薦システムを開発した点にある. このアプローチにより, データ活用の新たな可能性を広げ, ユーザの納得感を高める意思決定を支援するための基盤を築いたといえる.

最後に, 今後の展望を述べる. 本研究では, アスペクトに関する粒度の適切性が問題として残っており, 例えばアスペクト「食事」においても, 「朝食のバイキング」と「夕食の懐石料理」を重視するユーザを区別することが困難である. そこで, より詳細なアスペクト情報を組み込むことで, 個別化された推薦と説明の提供が実現する可能性がある. ただし, これにより知識グラフの構造の変化が考えられ, 推薦精度にどのような影響を及ぼすのかを慎重に検討する必要があるだろう.

謝辞

この場を借りて、私の研究を通じて絶えずサポートしてくれた家族に心から感謝いたします。また、専門的な指導と貴重な助言を惜しみなく提供してくださった市瀬龍太郎教授に心からの感謝の意を表します。市瀬教授は、私の些細な研究の相談にも真摯に向き合ってくださいました。本研究への多大な洞察だけでなく、研究の進め方や発表方法など、多岐にわたる学びを提供してくださいました。心より感謝申し上げます。さらに、多様なバックグラウンドを持つ研究室のメンバーや研究員の皆様にも感謝の意を表します。彼らとの刺激的な議論は、私の研究アプローチを形作る上で不可欠でした。最後に、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、楽天グループ株式会社の楽天トラベルデータを利用させていただきました。ここに深く感謝の意を表します。

参考文献

- [1] Carlos A. Gomez-Uribe and Neil Hunt. The Netflix Recommender System: Algorithms, Business Value, and Innovation. ACM Transactions on Management Information Systems, Vol. 6, No. 4, pp. 1–19, 2016.
- [2] Blake Morgan. How Amazon Has ReOrganized Around Artificial Intelligence And Machine Learning. <https://www.forbes.com/sites/blakemorgan/2018/07/16/>, 2018. Accessed 2025-01-15.
- [3] 楽天インサイト株式会社. 2022 年の夏は約 4 人に 1 人が「宿泊を伴う国内旅行」へ。コロナ前と比べて旅行計画時に重視するようになったのは「密を避けられる場所であること」がトップで約 7 割. <https://insight.rakuten.co.jp/report/20220914>, 2022. Accessed 2025-01-15.
- [4] Bas H. M van der Velden. Explainable AI: current status and future potential. European Radiology, Vol. 34, pp. 1187–1189, 2024.
- [5] Bryce Goodman and Seth Flaxman. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". AI Magazine, Vol. 38, No. 3, pp. 50–57, 2017.
- [6] Yongfeng Zhang and Xu Chen. Explainable Recommendation: A Survey and New Perspectives. Foundations and Trends® in Information Retrieval, Vol. 14, No. 1, pp. 1–101, 2018.
- [7] F. Maxwell Harper and Joseph A. Konstan. The MovieLens Datasets: History and Context. ACM Transactions on Interactive Intelligent Systems, Vol. 5, No. 4, pp. 1–19, 2015.
- [8] Markus Schedl. The LFM-1b Dataset for Music Retrieval and Recommendation. In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, pp. 103–110, 2016.
- [9] 楽天グループ株式会社. 楽天トラベルデータ. 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). <https://doi.org/10.32130/idr.2.2>, 2020. Accessed 2025-01-15.

- [10] Robin Burke. Hybrid Recommender Systems: Survey and Experiments. User Modeling and User-Adapted Interaction, Vol. 12, No. 4, p. 331–370, 2002.
- [11] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet Allocation. The Journal of Machine Learning Research, Vol. 3, p. 993–1022, 2003.
- [12] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed Representations of Words and Phrases and their Compositionality. In Advances in Neural Information Processing Systems, pp. 3111–3119, 2013.
- [13] Quoc Le and Tomas Mikolov. Distributed Representations of Sentences and Documents. In Proceedings of the 31st International Conference on Machine Learning, pp. 1188–1196, 2014.
- [14] Marko Balabanović and Yoav Shoham. Fab: content-based, collaborative recommendation. Communications of the ACM, Vol. 40, No. 3, p. 66–72, 1997.
- [15] Rakesh Agrawal and Ramakrishnan Srikant. Fast Algorithms for Mining Association Rules in Large Databases. In Proceedings of the 20th International Conference on Very Large Data Bases, pp. 487–499, 1994.
- [16] Yifan Hu, Yehuda Koren, and Chris Volinsky. Collaborative Filtering for Implicit Feedback Datasets. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, pp. 263–272, 2008.
- [17] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. BPR: Bayesian Personalized Ranking from Implicit Feedback. In Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, pp. 452–461, 2009.
- [18] Qingyao Ai, Vahid Azizi, Xu Chen, and Yongfeng Zhang. Learning Heterogeneous Knowledge Base Embeddings for Explainable Recommendation. Algorithms, Vol. 11, No. 9, 2018.
- [19] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph Attention Networks. In Proceedings of the 6th International Conference on Learning Representations, 2017.
- [20] Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. KGAT: Knowledge Graph Attention Network for Recommendation. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 950–958, 2019.

- [21] Yikun Xian, Zuohui Fu, S. Muthukrishnan, Gerard de Melo, and Yongfeng Zhang. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 285–294, 2019.
- [22] Yikun Xian, Zuohui Fu, Handong Zhao, Yingqiang Ge, Xu Chen, Qiaoying Huang, Shijie Geng, Zhou Qin, Gerard de Melo, S. Muthukrishnan, and Yongfeng Zhang. CAFE: Coarse-to-Fine Neural Symbolic Reasoning for Explainable Recommendation. In Proceedings of the 29th ACM International Conference on Information Knowledge Management, pp. 1645–1654, 2020.
- [23] Shijie Geng, Zuohui Fu, Juntao Tan, Yingqiang Ge, Gerard de Melo, and Yongfeng Zhang. Path Language Modeling over Knowledge Graphs for Explainable Recommendation. In Proceedings of the ACM Web Conference 2022, pp. 946–955, 2022.
- [24] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. In Advances in Neural Information Processing Systems, pp. 6000–6010, 2017.
- [25] Iván Cantador, Andrés Carvallo, and Fernando Diez. Rating and aspect-based opinion graph embeddings for explainable recommendations. [arXiv:2107.03385](https://arxiv.org/abs/2107.03385), 2021.
- [26] Giacomo Balloccu, Ludovico Boratto, Gianni Fenu, and Mirko Marras. Post Processing Recommender Systems with Knowledge Graphs for Recency, Popularity, and Diversity of Explanations. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 646–656, 2022.
- [27] Kazuki Fujii, Taishi Nakamura, Mengsay Loem, Hiroki Iida, Masanari Ohi, Kakeru Hattori, Hirai Shota, Sakae Mizuki, Rio Yokota, and Naoaki Okazaki. Continual Pre-Training for Cross-Lingual LLM Adaptation: Enhancing Japanese Language Capabilities. In Proceedings of the First Conference on Language Modeling, 2024.
- [28] Antoine Bordes, Nicolas Usunier, Alberto Garcia Duran, Jason Weston, and Oksana Yakhnenko. Translating Embeddings for Modeling Multi-relational Data. In Advances in Neural Information Processing Systems, pp. 2787–2795, 2013.



付録A アスペクトベース感情分析

大規模言語モデルを利用したアスペクトベース感情分析について説明する. 本研究では, ユーザレビューの各コメントに対して, この自然言語処理タスクを適用することで, 3.2節で提案した知識グラフの構築に利用する, コメント内で言及されているアスペクトの特定を実現した. このタスクは, 文中の特定のアスペクトに対する感情を識別することを目的としており, 以下の2つのステップで行われる.

1. 文中で, 特定のアスペクトに言及している部分を特定.
2. 言及された部分の感情 (ポジティブ, ネガティブ) を特定.

感情分析の結果を知識グラフの構築に利用するため, ポジティブまたはネガティブと明確に識別された感情を持つアスペクトを特に重視している. このアプローチは, ユーザが感情を強く示すアスペクトはそのユーザにとって特に重要であるという仮説に基づいている. 中立的な感情や評価の言及は, ユーザの関心が比較的低い可能性を示すため, この分析では主に極端な感情の表現に焦点を当てている.

大規模言語モデルにプロンプト (指示) を与えることで, アスペクトベース感情分析を実施した. 本研究で利用したユーザレビューのコメントは日本語によるものであったため, 日本語に対応した大規模言語モデルを選定し, 東京工業大学 (現東京科学大学) の「Swallow-13b-instruct-hf」[27]を使用した. また, この自然言語処理タスクは本研究の主要な焦点ではなかったため, 精度を向上させるためのモデルのファインチューニングやプロンプトの詳細な調整を実施していない. プロンプトは, タスクの定義と説明, 具体的な例示, 具体的な入力と期待される出力の形式を含む構成となっている. 実際に利用したプロンプトを次頁に示す.

以下の感情分析タスクを行ってください。与えられた文から、すべての [アスペクト, 感情] ペアを特定してください。アスペクトは以下のカテゴリのいずれかに該当します: ['食事', '部屋', '風呂', '立地', '設備', 'サービス']。

具体的には、「食事」はホテルの夕食や朝食、「部屋」は部屋の内装や部屋からの景色、「風呂」は大浴場や露天風呂などのお風呂、「立地」は駅からの近さや観光地へのアクセス、「設備」はネットワークやアメニティ、「サービス」は店員の接客やエンターテインメントを指します。アスペクトは ['食事', '部屋', '風呂', '立地', '設備', 'サービス'] のいずれか、感情は ['ネガティブ', 'ポジティブ'] のいずれかで表されます。アスペクト-感情ペアが文中に存在しない場合は、空のリストを返してください。それ以外の場合は各ペアを二つの文字列としてシングルクォートで囲み、Python リスト形式で返してください。結果は Python のリストの形式のみで返してください。以下に例を示します。

文: 朝食のビュッフェは地元の新鮮な食材をふんだんに使用しており、特にフレッシュなフルーツと手作りのパンが印象的でした。客室の清潔さと快適なベッドも満足できるレベルで、良質な睡眠を提供してくれました。フロントデスクのスタッフは非常に親切で、迅速なチェックインが印象的でした。ホテルの立地は市街地に近く、観光地へのアクセスが非常に便利でした。

ラベル: ['食事', 'ポジティブ'], ['部屋', 'ポジティブ'], ['サービス', 'ポジティブ']

文: 露天風呂からの山並みの景色が息をのむ美しさで、温かいお湯が旅の疲れを癒してくれました。しかし、部屋の壁が薄く、隣室の話し声が聞こえて少し気になりました。

ラベル: ['風呂', 'ポジティブ'], ['部屋', 'ネガティブ']

文: フィットネスセンターの設備が充実しており、最新のトレーニング機器で効果的なワークアウトができました。夕食ではシェフの特製コースが提供され、地元の新鮮な食材を活かした創造的な料理が絶品でしたが、サービスの遅さが少し残念でした。

ラベル: ['設備', 'ポジティブ'], ['食事', 'ポジティブ'], ['サービス', 'ネガティブ']

文: 2 回目の大浴場の利用では、夕暮れ時の景色が特に印象的で、リラックス効果が抜群でした。部屋に備え付けられた高速 Wi-Fi がビジネス作業に大変役立ちました。しかし、ホテルは最寄り駅からかなり離れた場所にあり、公共交通機関を利用するアクセスが不便でした。

ラベル: ['風呂', 'ポジティブ'], ['設備', 'ポジティブ'], ['立地', 'ネガティブ']

文: 今回の滞在で特に不満はありませんでしたが、ホテル全体としては平均的な印象で、特に際立った点は見当たりませんでした。清潔で快適ではありましたが、他と差別化できる特別なサービスや施設はありませんでした。

ラベル: []

上記のガイドラインに従って、以下の文のラベルを特定してください。説明は不要です。