

指導教員(主査)：牛尼 剛聡 准教授

副査：上岡 玲子 准教授

副査：森本 有紀 助教

修 士 論 文

レビューへのフィードバックに基づいたアイテムの絞り込みと品定
めの支援に関する研究

A Study for Supporting Item Selection Using Feedback on Reviews

九州大学大学院芸術工学府 芸術工学専攻 コンテンツクリエー
ティブデザイン

平成28年4月入学 学籍番号 2DS16081W

南 大智

2018年1月提出

要旨

近年、インターネットの普及に伴い、利用者はアイテムの購入や調査をインターネットを介して行うことが一般的になった。しかし、インターネット上には莫大な数のアイテムが存在しており、ユーザは自分の好みに合ったアイテムを発見することが困難になっている。こうした中で、アイテムに対するレビューを閲覧可能な投稿型レビューサイトの役割が重要になっている。しかし、中には、不誠実な業者や悪意を持ったユーザによる信頼性の低いレビューも存在し、投稿されているレビューの信頼性が問題になっている。また、ユーザの嗜好の違いによって、悪意はないがユーザにとって好ましくないレビューが投稿されている場合も存在する。このような、信頼に値しないレビューワーカーの存在が、アイテム選別の妨げになる場合が考えられる。また、ユーザがオンラインレビューサイト上でアイテム選別を行う際に、見知らぬレビューワーカーが信頼できるかどうかを判断することは困難である。本研究では、対象ユーザのレビューに対するフィードバックに基づいて、未知のレビューワーカーに対する信頼を予測する手法を提案する。提案手法に基づき、投稿型レビューサイトにおけるレビューワーカーへの信頼の予測を行い、得られた信頼情報を用いて、アイテム推薦の最適化及びレビュー推薦を行う。私たちは、アイテム推薦の最適化による「莫大な数のアイテムの中からの絞り込み」と、レビューランキングの個人化による「候補アイテムに対する価値判断（品定め）」のそれぞれの観点から、アイテム選別の支援を行う。プロトタイプを用いた被験者実験において、それぞれの手法の有効性の検証を行なった。

目次

要旨	i
第1章 はじめに	1
1.1 研究の背景	1
1.2 研究の目的	2
1.3 論文の構成	5
第2章 関連研究	7
2.1 情報推薦におけるフィードバックの活用	7
2.2 レビューの品質推定	8
2.3 レビューを利用した情報推薦	8
2.4 協調フィルタリング拡張手法	9
第3章 レビューへのフィードバックによる嗜好情報の取得	10
第4章 レビューに基づくアイテムの絞り込み支援	12
4.1 観点の違いを考慮したレビューの重み付け	12
4.2 レビューの確信度を考慮したレビューの重み付け	17
4.3 レビューの重みを利用した協調フィルタリングの拡張	21
4.4 評価実験	21
4.5 議論	24
第5章 レビューの信頼度に基づくアイテムの品定め支援	28
5.1 アイテムの品定めにおけるレビュー推薦の必要性	28
5.2 レビューの信頼度に基づくレビュー推薦手法	28
5.3 著者推定モデルによるレビューの確信度計算	30
5.4 評価実験	31

目次	iii
5.5 議論	34
第6章 まとめ	36
謝辞	38
参考文献	39

第1章

はじめに

1.1 研究の背景

近年，インターネットの普及に伴い，ECサイトやブログ，SNSなどの様々なWebサービスが登場し，インターネットを介して商品の購入や調査などを行うことが一般的になった．また，インターネット上に存在するアイテムの数も増加している．例えば，2017年3月の時点で，日本における代表的なECサイト「楽天市場」[1]には194,023,903点の商品が存在する [2]．また，特定のカテゴリのアイテムも莫大な数が存在し，Google Booksでは，2010年8月5日時点での書籍総数は1億2986万4880冊としている [3]．このような「情報過多」は近年の重要な課題となっており [4]，ユーザが自分の好みに合うアイテムを発見することは困難である．

一方で，特定の分野における商品のレビューを集めたSNS形式のWebサイト（オンラインレビューサイト）が多くのユーザを獲得している．日本におけるオンラインレビューサイトの代表例として，化粧品に対するレビューを集めた「@cosme」（アットコスメ）[5]，書籍を対象とした「読書メーター」[6]，映画を対象とした「Yahoo!映画」[7]などがあげられる．レビューはアイテムの選別を行う上で非常に有用であり，株式会社情報通信総合研究所の調査 [8]では，インターネットショッピングサービスを利用する際に，全体の74.8%の人がレビューを参考にすると回答している．投稿されているレビューには，利用者の実際の経験に基づく感想や説明が書かれており，ユーザはアイテムが自分にとって価値があるかどうかについて，購入する前に判断することができる．

また，オンラインレビューサイトでは，ユーザが指定したアイテムのレビューを表示するだけでなく，レビューを利用したアイテム推薦機能も提供されている場合が多い．情報推薦の方式には，一般的に，コンテンツに基づくフィルタリング(content-based filtering)と，協調フィルタリング(collaborative filtering)の2種類がある [9]．例えば，「読書メーター」においては，「この本を読んだ人はこんな本も読んでいます」という協調フィルタリングを用いた書籍推薦機能を提供し

ている。「読書メーター」では、ユーザが読んだ書籍を登録できるため、登録データの分析によって求められるユーザ同士の類似性を基にして書籍の類似性を推定し、ユーザが興味を持った書籍に基づいて、そのユーザが興味を持つ可能性が高いと考えられる書籍を推薦する。

一般的に、オンラインレビューサイトにおいて、全てのレビューの価値は等価に扱われている。しかし、投稿されているレビューは対象ユーザにとって役に立つレビューばかりではなく、中には対象ユーザにとって役に立たないレビューも存在する場合がある。例えば、図1.1に示すような、悪意のあるユーザや業者によるスパムレビューがECサイトやオンラインレビューサイト上には存在し、ユーザがアイテムの価値を判断する上での妨げになることが考えられる。また、ユーザの感性の違いによって、物事をどう捉えるかは異なり、図1.2に示すように、同一アイテムに対するレビューであっても、評価値や評価の視点が異なる場合がある。そのため、悪意はないが、ユーザと感性が合わず、ユーザの共感できないレビューも存在する。しかし、未知のアイテムに対するレビューの中で、ユーザはどのレビュワーを信頼するべきか、判断することは困難である。

★★★★★ これは多分完璧にiphone 7 Plusにぴったり合うと思っております。
投稿者 Amazon カスタマー 投稿日 2016/9/8
色名: グレー | サイズ名: iPhone 7 Plus | Amazonで購入
但し、皆様に、申し訳ありません。
こちらは既にiphone 7 Plusというスマホを入手しておりました。
ここで皆様にこのケースの効果を示しております。
まず、このケースは耐衝撃性があります。なぜなら、不意にケースが3mのところから落ちてしまったのに、無事でした。それで、この効果にびっくりしました。
それに、四隅共に耐衝撃のスペーサーがありますので、上から落下するリスクも削減できます。
また、写真の通りに示しております。。iphone 7 Plusの全ての機能は正常として使えます。イヤホンはairpodのものであろうと、充電ポートのイヤホンであらう、いずれも正常として使えます。しかも、カメラ、Homeボタンの利用も問題がありません。
スクリーンのフィルムも問題がないと思っております。
ここで私も
Naliom iPhone 7 Plus 用全画面保護フィルム強化液晶保護ガラスフィルム を購入しました。
これは多分完璧にiphne 7 Plusにぴったり合うと思っております。
信号への干渉は殆どありません。

図1.1 スパムレビューの例

1.2 研究の目的

本研究では、レビューの価値はユーザによって異なるという仮定に基づいて、ユーザのレビューに対するフィードバックを用いて、レビュワーの価値を推定することで、より効率的なアイテム



図1.2 レビューによる評価値と評価観点の違いの例

の選別を支援する。具体的には、以下の2つのアプローチでアイテム選別の支援を行う。

1.2.1 協調フィルタリングによるアイテムの絞り込み

図1.3は、ユーザのアイテム選別のプロセスを表している。例えば、ユーザがアイテムを選別する際には、まず、購入候補アイテムを絞り込むことが考えられる。莫大な数のアイテムの中から購入候補を絞り込む上で、推薦システムは有用である。代表的な推薦アルゴリズムの一つである、協調フィルタリングは、ユーザの過去の履歴を見て、対象ユーザと同じアイテムを多く選択しているユーザは嗜好が類似していると判断し、類似ユーザが選択・購入したアイテムを推薦することで、高い精度をもった推薦が行える。例えば、図1.4では、対象ユーザがアイテムBを選択したことがあり、それ以外のアイテムは未選択である状況だとする。対象ユーザ以外で、アイテムBを評価した二人のユーザは、それぞれアイテムAとアイテムCを選択しているので、対象ユーザにはアイテムAとアイテムCを推薦するという仕組みである。

上記のアイテム推薦機能で用いられている協調フィルタリングは一般的に、「同一のアイテムを選択したユーザ同士はアイテム選択における嗜好が似ている」という仮定に基づいて推薦を行っている。しかし、物事を見たり考えたりするとき、人はそれぞれの立場から評価や考察を行う。書籍を対象としたオンラインレビューサイトにおいては、同一書籍に対するレビューであっても、

各レビューのレビューにおける評価の視点に違いが見られる場合が多い。例えば、漫画書籍に対するレビューの場合、絵についての評価をするレビューーもいれば、ストーリーについての評価をするレビューーもいる。さまざまな内容のレビューの中で、どのレビューに対して共感できるかどうかは、ユーザによって違いがあることが考えられ、同一アイテムを選択したユーザ同士であっても、嗜好の類似性が低い場合がある。そこで、書籍を対象としたオンラインレビューサイトにおける協調フィルタリングを利用した書籍情報の推薦の際に、対象ユーザのレビューに対する好みに基づいて、レビューの重みを変更して推薦書籍の計算を行うことで、推薦精度を向上させる。



図1.3 ユーザのアイテム選別のプロセス

1.2.2 レビュー推薦によるアイテムの品定め

次に、アイテム推薦により、選択候補のアイテムが絞り込まれた後、ユーザはアイテムが本当に自分の好みに合うかどうかについて、価値を判断する（品定め）行為を行うことが考えられる。ユーザが未知のアイテムの品定めを行う際に、アイテムに対する感想や説明が書かれている口コミ情報、すなわちレビューは有用な情報である。しかし、オンラインレビューサイトに投稿されているレビューの数は莫大であるため、効率的にレビューを閲覧することは難しいという問題がある。実際、Amazonなどのオンラインレビューサイトで商品を購入する際に、8割のユーザは

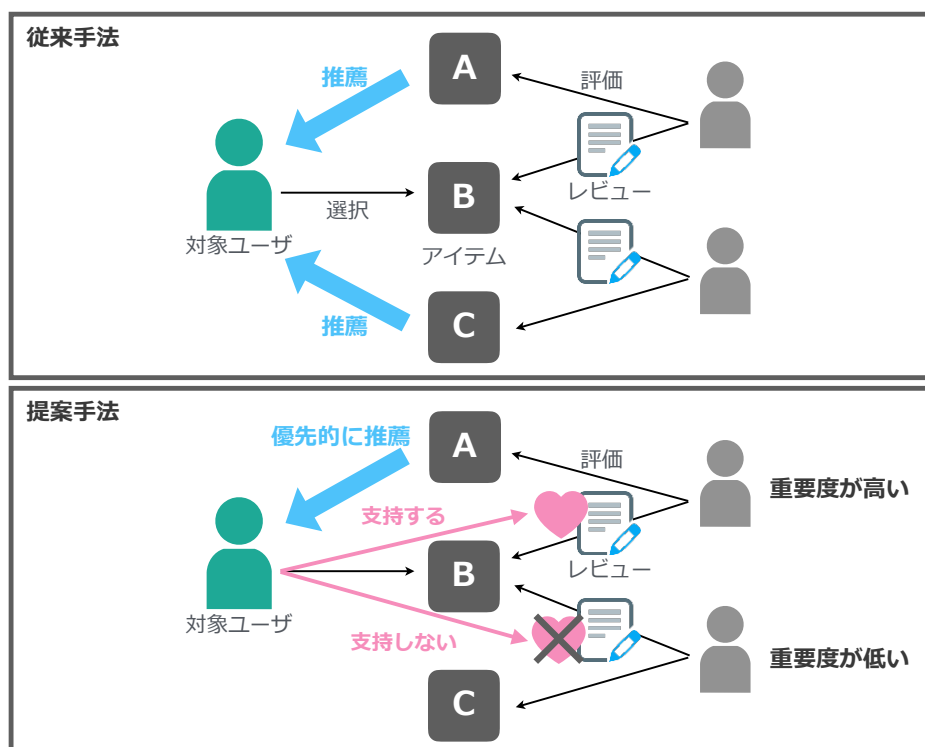


図1.4 従来型の協調フィルリングと提案手法の比較

レビューを最大でも10件程度しか読まない [10]。また、投稿されているレビューは対象ユーザにとって役に立つレビューばかりではなく、中には対象ユーザにとって役に立たないレビューも存在する。そこで、オンラインレビューサイトにおける協調フィルタリングによる推薦アイテムを品定めする際、図1.5のように、ユーザにとって役に立つ順番にレビューを並び替えることによって、効率的な品定めの支援を行う。

1.3 論文の構成

本研究では、レビューに対するフィードバックを基に、推定した対象ユーザのレビュワーに対する価値を用いて、アイテム選別の支援を行う。具体的には、アイテムの「絞り込み」と「品定め」の2つの行為に着目し、それぞれにおける評価実験を行う。本論文の構成を以下に示す。第2章では、本研究に関連する研究を紹介し、本研究の位置付けについて述べる。第3章では、レビューに対するフィードバックを用いた嗜好情報の取得について述べる。第4章では、アイテムの絞り込みのために、第3章で得られたレビューの評価情報を基に、レビュワーの信頼を推定する手法を述べる。また、推定されたレビュワーの信頼を利用した協調フィルタリング手法について説明し、評価実験の結果を示し、有効性を検証する。第5章では、アイテムの品定め支援のために、信頼情報を用いたレビュー推薦手法について説明し、評価実験の結果を示し、有効性を検

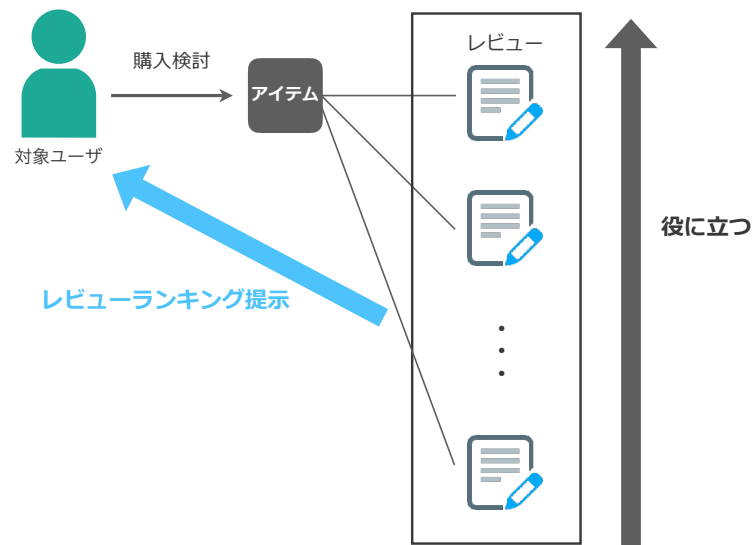


図1.5 レビューランキングの提示

証する。第6章では、今後の課題と展望を示し、本論文の総括を行う。

第2章

関連研究

2.1 情報推薦におけるフィードバックの活用

情報検索において、ユーザがクエリを正しく指定することができない場合に要求する情報を取得できないことは、重要な問題として認識されてきた。これを解決する代表的な手法として、適合フィードバックが知られている [11]。適合フィードバックは、情報検索において検索結果として出力された文章の内容に対するユーザの反応に基づいて、検索質問や検索戦略、検索式を修正する手法である [12]。このとき、クエリを修正する基となるのは、文書自体に対してのフィードバックである。代表的なクエリの修正は、Rocchioのアルゴリズムが用いられることが多く、次のように定義される。

$$\mathbf{q}_m = \alpha \mathbf{q}_0 + \frac{\beta}{|D_R|} \sum_{\mathbf{d}_i \in D_R} \mathbf{d}_i - \frac{\gamma}{|D_N|} \sum_{\mathbf{d}_i \in D_N} \mathbf{d}_i \quad (2.1)$$

ここで、 D_R と D_N は、それぞれユーザに与えられた検索結果のうち、興味のあるとした文書、興味がないとした文書である。また、 α 、 β 、 γ はそれぞれ元のクエリベクトル、正の影響ベクトル、負の影響ベクトルにどれくらいの重みを与えるかを定めるパラメータである。推薦するコンテンツの種類がテキストである場合は、ベクトルはキーワードの出現頻度で表され、tf・idf法などを用いて、キーワードに重みを付ける方法が適用される。

このように、推薦コンテンツ自体に対するフィードバックを利用する研究として、例えば、顔らは、スマートフォンでの効率的な商品選別を行うために、ユーザの振る舞いを基にした適合フィードバックを用いて推薦の最適化を行う手法を提案している [13]。

一般的に、適合フィードバックは、対象とするアイテム自体にフィードバックを与え、結果を最適化する手法である。それに対して、本研究では対象となるアイテムである書籍自体ではなく、書籍に関するレビューに対するフィードバックを与えるという点で独自性がある。

2.2 レビューの品質推定

Amazonでは、レビューが役に立ったかについての投票ボタンを設置しており、投票に基づいてレビューの品質を推定するための情報として利用している。しかし、Amazonのレビュー全体のうち約10%程度しか評価されていない [14]。また、古いレビューは多くの投票を集め、上位に表示されるが、まだ票を持っていない新しいレビューは役に立たないと考えられている [15]。また、多くのポジティブな投票を得たレビューは、より多くのポジティブな投票を集めやすい。これらの問題を解決するために、レビューの品質スコアを推定する研究が行われている [16]。また、レビューなどのテキスト情報の信頼性に関して、さまざまな研究が行われてきた [17–20]。これらの研究では、スパムなど、不当なレビューの投稿を行う悪意を持った投稿者を問題視しており、「スパムであるか否か」「虚偽の投稿であるか否か」ということに関する議論を展開している。従来のレビューのランキングやフィルタリングに関する研究は、レビューの客観的な価値に関するものが多く、「対象ユーザにとって有益であるか」や「対象ユーザが好感を抱くか」など、レビューに対する個人的な嗜好についてはあまり議論されていない。本研究においては、個人的な観点からのレビューの価値について着目する。

2.3 レビューを利用した情報推薦

近年、レビューを利用した情報推薦に関して多くの研究が行われている [21–32]。レビューを利用した推薦の方向性として主に以下の3点がある [33]。

- ユーザの好みに関する情報を抽出することでデータの希薄さ（データスパースネス）の問題に対応する。
- コールドスタート問題に対応する。
- ユーザのコンテキストや潜在的な嗜好因子など評価値以外の推薦に有用な情報をレビューから導出する。

レビューを利用してユーザの嗜好を抽出し情報推薦を行う研究として、林ら [34]は、ユーザが書いたレビューからユーザの嗜好を示す興味語とその極性を抽出し、興味語が含まれ極性が一致するレビューをもつ映画を推薦することで、ユーザの好みに合った推薦が可能であることを示した。

一般的に、上記のような情報推薦に関する研究は、アイテムの効率的な選別を支援するためにアイテムを推薦し、推薦アイテムがどれだけ対象ユーザの好みに合っているかという点を評価しているが、アイテムを推薦した後の価値判断の支援は行っていない。本研究では、推薦アイテム

が提示された後に、ユーザが行う品定め行為を支援することで、アイテムの効率的な選別を支援する。また、ユーザが過去に選択したアイテム全てに対してレビューを入力させることでユーザの嗜好を獲得するアプローチは、ユーザへの負担が大きいという点で現実的ではない。そこで、本研究ではアイテムのレビューに対するフィードバックを用いて、間接的に、ユーザの嗜好を推定するアプローチをとる。

2.4 協調フィルタリング拡張手法

協調フィルタリングを拡張する手法に関して、今までにさまざまな提案がされてきた。協調フィルタリングでは、ユーザのアイテムに対する評価値データを基にして推薦を行うが、評価値の欠損により推薦が上手く行えないというスパース性の問題に対して、少ない評価値情報から効率的に未知の評価値を推定する手法として、Matrix Factorization (MF) [35]が有効であることが確認されている。MFは、次元削減を行うことで、ユーザの潜在的特徴を利用し、より疎な部分が少ない評価値行列を作成することで、精度の良い推薦を行うことができる。また、ユーザやアイテムの潜在的特徴に関して、評価値以外の情報を利用する手法についても研究されている。協調フィルタリングとコンテンツに基づくフィルタリングを混合させたモデルであるコンテンツベース協調フィルタリング [36]においては、ナイーブベイズ法を用いて、ユーザの属性（年齢や性別など）からユーザアイテムマトリックスの欠損値を補完して、擬似マトリックスを生成する。また近年では、潜在的特徴の解析手法として近年注目されているトピックモデルを協調フィルタリングに用いた研究も行われており、代表的な手法としてLDAの著者であるBleiらが提案したCTRが知られている [37]。CTRは、アイテムに関するテキスト情報からトピックを抽出して、ユーザとアイテムの特徴ベクトルに反映させることで、評価値が存在しないアイテムであってもトピック情報を基に評価値を推定する手法である。これらの手法では、いずれもユーザの潜在的特徴を用いて評価値の推定を行っているが、評価における観点は考慮されていない。

本研究における評価観点を推定するためには、ユーザの書籍に対する印象を知る必要がある。GroupLens [38]においてはユーザのアイテムに対する評価を明示的に入力させるが、ユーザが過去に選択したアイテム全てに対してレビューを入力させることでユーザの嗜好を獲得するアプローチは、ユーザへの負担が大きいという点で現実的ではない。そこで、本研究ではアイテムのレビューに対するフィードバックを用いて、間接的に、ユーザの嗜好を推定するアプローチを採用する。

第3章

レビューへのフィードバックによる嗜好情報の取得

本論文では、対象ユーザの既知のレビューに対する評価に基づいて、未知のレビューに対する信頼を予測するモデルを提案する。提案モデルでは、対象ユーザは既知のアイテムに投稿されているレビューに対する評価を入力する。評価されたユーザの特徴と未知のユーザの特徴との間の類似度を求めることで未知のユーザに対する信頼度を推定する。図3.1は、オンラインレビューサイトにおけるレビューに対する評価の入力インターフェースの例を表す。対象ユーザはこのインターフェースを用いて、既知のユーザに対する評価を入力する。この図においては、レビュー文の右下に表示されている二つのアイコンが評価の入力ボタンであり、支持と書いてあるアイコンはポジティブな評価、不支持と書いてあるアイコンはネガティブな評価を入力するために利用する。また、レビューに対する評価がどちらとも言えない場合は、どちらのアイコンもクリックしないことで、中立の表現とする。システムはユーザから入力された各レビューへの評価を、ユーザプロフィール情報として登録する。



タッチ (1) (小学館文庫)

タッチ 1巻の感想・レビュー



双子なのに、兄はグウタラ、弟はスーパーマン。隣の南ちゃんとはとにかく
カワイイ!! 大ブームを巻き起こした青春まんが!!

★ナイス! ★32 - コメント(0) - 2015年8月20日

コメントする

👍 支持

👎 不支持



比喩表現が特徴的な作品。見開き1ページつかった迫力さを出したりもせず、コマも小さく淡々とした構成だからこそ生きてくる表現なんだろうなあ。

★ナイス! ★13 - コメント(0) - 2014年9月27日

コメントする

👍 支持

👎 不支持



先週のアメトーークを見ていたらまた読みたくなりました。やっぱり南ちゃん、可愛い! 昔はタッチちゃんのほうが好きだったけど、今はカッチちゃんのほうが好きかもしれません。

★ナイス! ★12 - コメント(0) - 2014年9月19日

コメントする

👍 支持

👎 不支持

図3.1 インターフェース上での評価の入力例

第4章

レビュー者に基づくアイテムの絞り込み支援

本論文では、レビュー者の重み付けに関して、以下の2つの手法を提案し、それぞれの手法の有効性について、検証する。

4.1 観点の違いを考慮したレビュー者の重み付け

4.1.1 アイテム推薦における評価観点情報の活用

協調フィルタリングは一般的に、「同一のアイテムを読んだユーザ同士はアイテム選択における嗜好が似ている」という仮定に基づいて推薦を行っている。しかし、物事を見たり考えたりするとき、人はそれぞれの立場から評価や考察を行う。例えば、漫画書籍の場合、絵が美しい漫画を好む人は、「絵が美しいかどうか」が漫画を評価する上での重要な要素だと考え、ストーリーが面白い漫画を好む人は、「ストーリーが面白いかどうか」が漫画を評価する上での重要な要素だと考える場合がある。協調フィルタリングにおいて、このようなユーザの評価における観点を考慮しないことによって、ユーザの嗜好に合わないアイテムが推薦される可能性がある。例えば、アイテム i に対して、 A という観点からの特徴を好んで選択したユーザ集合 U_A と、アイテム i に対して、 B という観点からの特徴を好んで選択したユーザ集合 U_B が存在するとする。対象ユーザがアイテム i を選択したとき、一般的な協調フィルタリングでは U_A と U_B は対象ユーザと同じアイテムを選択しているため、嗜好が類似していると判断される。しかし、対象ユーザがアイテム i を観点 A からの特徴に基づいて選択していた場合、 U_A は対象ユーザとの嗜好の類似性が高いと言えるが、 U_B は U_A と比べると、対象ユーザとの嗜好の類似性は低いと考えられる。つまり、このような場合であれば、 U_B に基づいて推薦されるアイテムは、対象ユーザの嗜好に合わない可能性があるという問題点がある。

本節では、オンラインレビューサイトにおける協調フィルタリングを利用したアイテム推薦の際に、対象ユーザの評価に対する嗜好（評価嗜好）を考慮し、対象ユーザが支持する評価観点とレビューの評価観点の類似性に基づいて書籍推薦を行うことで、推薦精度を向上させる手法を提案する。書籍を対象としたオンラインレビューサイトにおいては、同一書籍に対するレビューであっても、各レビューのレビューにおける評価の観点に違いが見られる場合が多い。例えば、漫画書籍に対するレビューの場合、絵についての評価をするレビューもいれば、ストーリーについての評価をするレビューもいる。このようなレビューの観点の違いを解析することによって、評価観点を推定できると考えられる。レビューの評価観点の計算には、Latent Dirichlet Allocation (LDA) [39]を用いる。具体的には、書籍を評価するための代表的なトピックを学習し、学習したトピックモデルを用いて、対象とするレビューのレビューはどのトピックが生成しやすいかについて推論計算することで求められる多項分布をレビューの評価観点とする。

4.1.2 システムの概要

本論文では、書籍を対象としたオンラインレビューサイトにおいて、ユーザが書評レビューに対する評価をフィードバックとして与えることで、推薦の内容を最適化してユーザに再提示するシステムを提案する。図4.1に本システムの概要を示す。本システムは、対象ユーザが選択した書籍に関するページにおいて、その書籍に対するレビュー、推薦書籍を表示する。初期の推薦書籍は各レビューの重要度は均一として一般的な協調フィルタリング手法に基づいて求める。対象ユーザはレビューを読み、個々のレビューに対する評価を支持・不支持・中立のいずれかとして入力する。評価されたレビュー集合を解析することで、対象ユーザの評価嗜好を求める。レビューの評価観点は、あらかじめLDAを用いて学習した、書籍を評価するためのトピックに対する分布を推論計算することで求める。そして、対象ユーザの評価嗜好と各レビューの評価観点との間のコサイン相関値を求め、その値をレビューに対する重みとして、推薦書籍の再計算および再提示を行う。

4.1.3 LDAを用いた評価観点の解析

Latent Dirichlet Allocation (LDA)

LDA [39]とは、Bleiらが提案したトピックモデルである。トピックモデルとは、1つの文書が1つまたは複数のトピックをもつと仮定し、文書ごとのトピック分布 $\theta_d = (\theta_{d1}, \dots, \theta_{dK})$ と、それらのトピックごとの単語分布 $\phi_z = (\phi_{z1}, \dots, \phi_{zN})$ に基づいて確率的に文書を生成するモデルである。ここで、 $\theta_{dz} = p(z | \theta_d)$ は文書 d にトピック z が割り当てられる確率で、 $\phi_{zw} = p(w | \phi_z)$ はトピック z で単語 w が生成される確率である。具体的な文書の生成は次のように行われる。

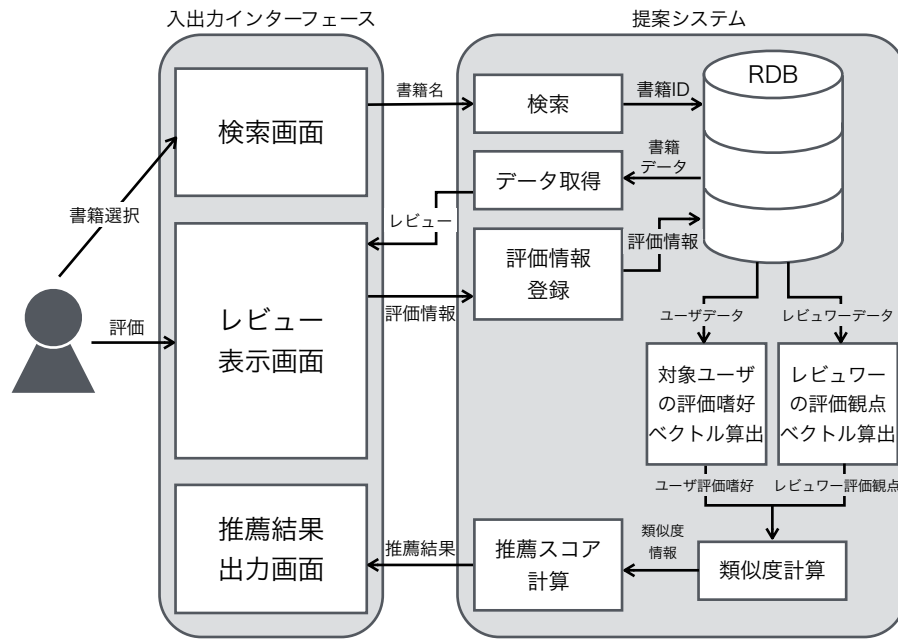


図4.1 システムの概要

1. For トピック $z = 1, \dots, K$
 - (a) 単語分布を生成 $\phi_z \sim \text{ディリクレ分布}(\beta)$
2. For 文書 $d = 1, \dots, D$
 - (a) トピック分布を生成 $\theta_d \sim \text{ディリクレ分布}(\alpha)$
 - (b) For 単語 $w = 1, \dots, N$
 - i. トピックを生成 $z_{dw} \sim \text{多項分布}(\theta_d)$
 - ii. 単語を生成 $w_{dw} \sim \text{多項分布}(\phi_{z_{dw}})$

この生成過程を表すグラフィカルモデルは図4.2のように表される。トピック分布 θ と各トピックにおける単語分布 ϕ の推定には、GriffithsらのGibbs Samplingによる推定手法を用いた [40]。

LDAによる評価観点の学習

本節では、全ての書籍レビューに対してLDAを用いて、書籍を評価する際の多様な観点をトピックとして抽出する。このとき、書籍自体に関するトピックが抽出されるのは避けることが望ましい。そこで、書籍の登場人物名や作者名など、特定の書籍のみに出現するような単語を除外するために、以下の前処理を行う。

全書籍のレビューを Mecab [43]を用いて形態素解析し、名詞・形容詞・形容動詞を抽出する。そして、それぞれの書籍ごとに、その書籍に対する全レビュー中の単語 w を含むレビューの割合を

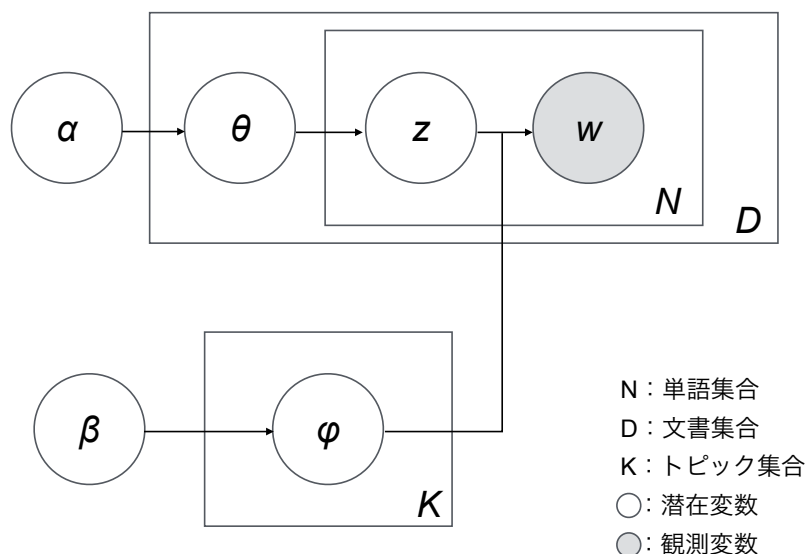


図4.2 LDAのグラフィカルモデル

求める．形式的には，書籍 b における単語 w を含むレビューの出現率 $r(b, w)$ を以下の式で表す．

$$r(b, w) = \frac{|\{r \mid r \in \text{reviews}(b), w \in \text{words}(r)\}|}{|\text{reviews}(b)|} \quad (4.1)$$

ここで， $\text{reviews}(b)$ は書籍 b を含むレビュー集合であり， $\text{words}(r)$ はレビュー r を含む全単語からなる集合を表す．書籍集合 B における，それぞれの書籍に対する単語 w の平均出現率が低い場合は，その単語は特定の書籍固有の単語であると考えられるため，評価観点要素の抽出には用いないこととする．書籍集合 B における単語の w の一般度 $\text{major}(B, w)$ を以下の式で定義する．

$$\text{major}(B, w) = \frac{\sum_{b \in B} r(b, w)}{|B|} \quad (4.2)$$

上記の式によって，求められた一般度 $\text{major}(B, w)$ が設定した閾値より低い単語をトピック抽出の対象から除外する．一般度の低い単語が除外された，全レビューの全レビューに対してトピック抽出を行い，得られたトピックの集合を，書籍評価における観点トピックの集合とする．抽出された書籍評価の観点トピックの例を表4.1に示す．ここで対象としたのは「読書メーター」に含まれる書籍レビューである．対象としたレビュー数は9万5,719人であり，これらのレビューの投稿した全てのレビュー，計1233万1749件を用いた．トピック数は100とし，各トピックあたりに含まれる単語数は10と設定した．

表4.1 観点トピックの例

topic1	topic2	topic3	topic4
キャラ	自分	漫画	最後
主人公	子供	絵	登場人物
感じ	本	作品	描写
これ	それ	購入	印象
ヒロイン	家族	表紙	もの
かわいい	ため	先生	ラスト
期待	心	絵柄	物語
ラブコメ	親	作者	伏線
設定	時代	楽しみ	人間
漫画	気持ち	連載	文章

LDAによる評価観点の推論

レビューーの全レビューを連結したものを1つの新規文書として、LDAの推論計算により、観点トピックに対して、新規文書はどのトピックが生成しやすいかの多項分布 $\theta_{d_{new}}$ を以下の式で推定することができる。

$$\theta_{d_{new}} = \sum_{z=1}^K p(z)p(\mathbf{d}_{new} | z) \quad (4.3)$$

$p(z)$ は混合比と呼ばれ、トピック z が選ばれる確率を表し、4.1.3で学習したモデルでの全体における θ_z に対応している。 $p(\mathbf{d}_{new} | z)$ はトピック z が選ばれたときに文書 \mathbf{d}_{new} が生成される確率を表し、これに $p(z)$ をかけることで、ベイズ推定より、文書 d_{new} においてトピック z が選ばれる確率である $\theta_{d_{new},z} = p(z | \mathbf{d}_{new})$ を求めている。これをレビューーの評価観点ベクトルとして扱う。

4.1.4 対象ユーザの評価嗜好の推定

3で得られた、評価情報のうち、ポジティブなフィードバックを受けたレビューーの評価観点ベクトルを正の影響ベクトル、ネガティブなフィードバックを受けたレビューーの評価観点ベクトルを負の影響ベクトルとして、対象ユーザの評価に対する嗜好を推定する手法を述べる。フィードバックによって得られるのは、ユーザの評価観点そのものではなく、評価観点に対する嗜好である。もし、ユーザがレビューーに対して正の評価も負の評価も行わなかった場合には、ユーザは評価に対する嗜好が存在しないため、全てのレビューーを等価に扱う。2.1で記述したRocchioのアルゴリズムと同様なアプローチで、対象ユーザ u のフィードバックによる評価嗜好ベクトル

\mathbf{p}_u を以下の式で定義する.

$$\mathbf{p}_u = \frac{\beta}{|R_{pos}|} \sum_{r \in R_{pos}} \mathbf{v}_r - \frac{\gamma}{|R_{neg}|} \sum_{r \in R_{neg}} \mathbf{v}_r \quad (4.4)$$

R_{pos} は, 対象ユーザ u がポジティブなフィードバックを返したレビュー集合である. 一方, R_{neg} は, 対象ユーザ u がネガティブなフィードバックを返したレビュー集合であり, \mathbf{v}_r はレビュー r の評価観点ベクトルである.

4.1.5 レビューに対する共感度の推定

本手法では, ユーザごとにレビューに対する共感度が異なるを考える. 対象ユーザにとって共感度が高いレビューは協調フィルタリングの際に重要視され, 共感度が低いレビューは協調フィルタリングの際に重要視されないものとする. ここでは, 対象ユーザの評価嗜好に基づいて, レビューの評価観点に対する共感度を推定する手法について述べる. 本手法では, 対象ユーザの評価観点に対する嗜好を適用したレビューの評価観点ベクトルと, 嗜好を適用しないレビューの評価観点ベクトルとの間のコサイン相関値を求め, 得られた値を対象ユーザのレビューに対する共感度とする. 対象ユーザ u のレビュー r に対する共感度 $\text{emp}(u, r)$ は, 以下の式で定義する.

$$\text{emp}(u, r) = \cos(\mathbf{v}_r, \alpha \mathbf{v}_r + \mathbf{p}_u) \quad (4.5)$$

ここで, $\alpha \mathbf{v}_r + \mathbf{p}_u$ は, 対象としているレビュー r の評価観点に対して, ユーザの評価嗜好を反映した評価観点ベクトルを表している. この式では, 対象とするレビューの評価観点ベクトルと, ユーザの評価嗜好を反映させた評価観点ベクトルとのコサイン相関値をとることにより, 対象ユーザのレビューに対する共感度を推定している.

4.2 レビューの確信度を考慮したレビューの重み付け

4.2.1 評価されたレビューに対する価値推定

3章で得られた評価情報に基づいて, 対象ユーザが評価したユーザに対する価値を推定する. 対象ユーザはレビューが投稿したレビューに対する評価を行う. しかし, 一件のレビューからレビューを正しく評価できるとは限らないという問題がある. 一般的に, レビューは1つのアイテムに対してだけではなく, 様々なアイテムに対してレビューを投稿しており, それぞれのレビューには, レビューごとの個性や癖が反映されと考えられる. しかし, 様々なアイテムに対するレビューを投稿する中で, レビューが普段は投稿しないような内容のレビューを投稿することも考えられる. 例えば, 図4.3のように, 負の評価を入力されたレビューが普段のレビュー

があまり投稿しないような内容のレビューだった場合、そのレビューの他のレビュー全てが対象ユーザにとって価値が低いとは言えない。本節では、レビューの価値を計算する際に、評価された文書がどれだけそのユーザらしい発言であるかという確信度（Confidnece）を考慮したレビューの信頼推定手法を提案する。レビューの確信度は、任意のレビューがレビューらしい文書である時に高くなり、レビューらしくない場合には低くなる値であり、レビューに対する信頼値を計算する際に、レビュー自体の評価をレビューの信頼値にどれくらい反映させるかを決定するために用いられる。図4.4は、直接的にレビューを評価されたレビューに対する価値を推定する概念図である。対象ユーザ u が評価したレビューを投稿したレビュー r に対する信頼度 $\text{trust}(u, r)$ は、評価に基づき、以下の式で定義する。

$$\text{trust}(u, r) = \frac{1}{\|D_r(\text{pos})\|} \sum_{d \in D_r(\text{pos})} \text{conf}(d, r) - \frac{1}{\|D_r(\text{neg})\|} \sum_{d \in D_r(\text{neg})} \text{conf}(d, r) \quad (4.6)$$

この式では、対象ユーザからポジティブな評価を受けた文書集合の典型性の総和から、ネガティブな評価を受けた文書集合の典型性の総和を引くことにより、対象ユーザが既知であるユーザに対する信頼度を推定している。ここで、 $D_r(\text{pos})$ はレビュー r の投稿した文書集合のうち、対象ユーザからポジティブな評価を受けた文書集合を表し、 $D_r(\text{neg})$ はレビュー r の投稿した文書集合のうち、対象ユーザからネガティブな評価を受けた文書集合を表し、 d は1つの文書を表している。 $\text{conf}(d, r)$ は文書 d がどれだけレビュー r らしい文書であるかについての確信度を表し、本研究では、 $\text{conf}(d, r)$ をレビュー r における文書 d の確信度と定義する。確信度の詳細に関しては??で後述する。



図4.3 評価されたレビューが他の投稿の内容とは異なる例

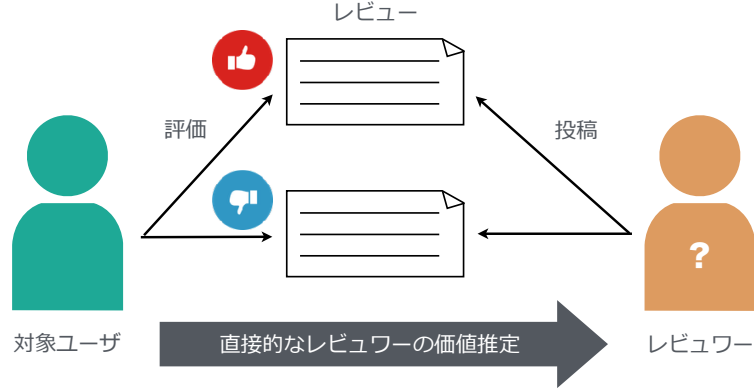


図4.4 評価されたレビューに対する価値推定

4.2.2 未評価のレビュワーに対する価値推定

ここでは、4.2.1で得られた既知のユーザに対する信頼情報に基づいて、対象ユーザがまだ評価していないユーザに対する信頼度を推定する手法について述べる。本手法では、対象ユーザに評価されたユーザの特徴ベクトルと、まだ評価されていないユーザの特徴ベクトルとの間のコサイン類似度を求め、得られた値が、まだ実際には評価されていないレビュワーに対する信頼度に影響を与える。信頼度の高いレビュワーと似ているレビュワーの価値は高く、反対に信頼度の低いレビュワーと似ているレビュワーの価値は低いという考えに基づいて、対象ユーザ u に対する任意のレビュワー r の信頼スコア $\text{trust}(u, r)$ は以下の式で計算される。

$$\text{trust}(u, r) = \left(\sum_{d \in D(\text{pos})} \text{sim}(d, d_r) - \sum_{d \in D(\text{neg})} \text{sim}(d, d_r) \right) \times \text{conf}(d_r, r) \quad (4.7)$$

この式では、対象ユーザからポジティブな評価を受けたレビュー集合と信頼を推定したい未知のユーザのレビューとの類似度の総和からネガティブな評価を受けたレビュー集合と信頼を推定したい未知のユーザのレビューとの類似度の総和を引き、得られた値にレビューの典型性を掛け合わせることで、対象ユーザが未知であるレビュワーに対する信頼度を推定している。ここで、 \mathbf{D}_{pos} は対象ユーザからポジティブな評価を受けたレビュー集合、 \mathbf{D}_{neg} は対象ユーザからネガティブな評価を受けたレビュー集合を表している。 $\text{sim}(d, d_r)$ は、レビュー d と未評価レビュワー r のレビュー d_r との間の類似度を表し、Doc2Vecによりベクトル化された \mathbf{d} と \mathbf{d}_r の間のコサイン類似度を計算することで求める。

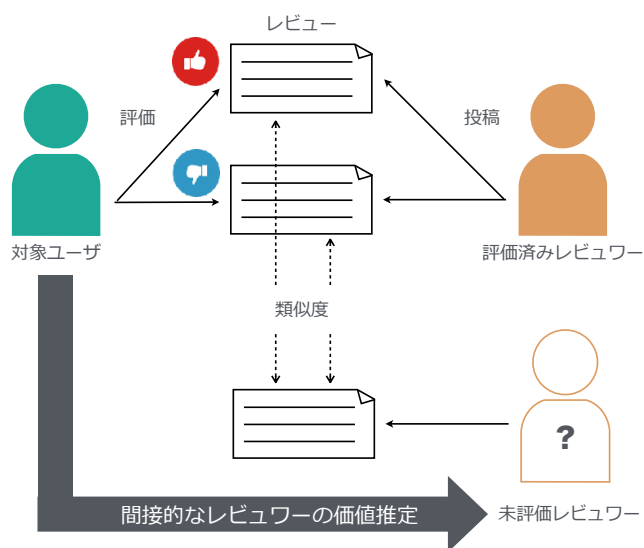


図4.5 未評価のレビュワーに対する価値推定

4.2.3 普段の発言との比較に基づくレビューの確信度計算

本節では，評価されたレビューがどれだけそのユーザらしい発言であるかという確信度を信頼推定モデルの中に組み込むことで，計算される信頼度の確信性を向上させる．レビューの確信度は，ユーザが今まで投稿した全てのレビューをひとまとめにした文書のベクトルと，対象となるレビューのベクトルとの間のコサイン類似度を計算することで求める．レビュワー r における文書 d の確信度 $\text{conf}(d, r)$ は以下の式で定義される．

$$\text{conf}(d, r) = \cos(\mathbf{d}, \mathbf{D}_r) \quad (4.8)$$

\mathbf{d} は典型性を求めたいレビューのベクトルであり， \mathbf{D}_r はユーザ u が投稿したレビュー全体のベクトルである．

4.2.4 Doc2Vecによる文書のベクトル化

Doc2Vecは文書をベクトル化するための手法の一つである [41]．この手法は，語順を基にして単語の次の単語を予測できるような表現ベクトルを獲得するWord2Vec [42]を文章に拡張したものである．この手法により，Bag-of-Wordsなどの従来手法では難しいとされる，文脈も含めたベクトル化を行うことが期待できる．Doc2vecによるベクトル化を行うためには，まず膨大な量の文書データをコーパスとして学習することが必要である．今回は，書籍レビューサイト「読書メーター」から取得したレビューを学習データとして用いたDoc2vecモデルを用いてレビュー文書の

ベクトル化を行う．4.1.3と同様の前処理を行い，平均出現率の低い単語が除外された，全レビューの全レビューを用いてDoc2Vecの学習を行う．ここで対象としたのは「読書メーター」に含まれる書籍レビューである．対象としたレビュー数は9万5,719人であり，これらのレビューの投稿した全てのレビュー，計1233万1749件を用いた．圧縮後の次元数は300とし，順番を考慮する単語幅は8と設定した．

4.3 レビューの重みを利用した協調フィルタリングの拡張

提案モデルに基づき，オンラインレビューサイトにおけるレビューに対する価値を推定し，その値をレビューに対する重みとして，推薦書籍の計算および提示を行う．書籍 b_1 を選択した時の，ユーザ u に対する書籍 b_2 の予測評価値は，以下の式で求める．

$$\text{Pred}(u, b_1, b_2) = \sum_{r \in \mathbf{R}(b_1)} (\text{trust}(u, r) \times \text{read}(r, b_2)) \quad (4.9)$$

ここで， $\mathbf{R}(b)$ は書籍 b に対してレビューを投稿したレビュー集合を表し， $\text{trust}(u, r)$ は対象ユーザ u にとってのレビュー r の信頼度を表す．また， $\text{read}(r, b)$ は，レビュー r が書籍 b を読んだかどうかによって値が決まる．

$$\text{read}(r, b) = \begin{cases} 1 & (\text{読んだ}) \\ 0 & (\text{読んでいない}) \end{cases} \quad (4.10)$$

書籍 b_1 を読んだレビュー集合のうち，書籍 b_2 を読んだレビューの信頼度の合計値がユーザ u に対する書籍 b_2 の予測評価値となり，これにより，ユーザの信頼度を考慮した協調フィルタリングによる推薦を行う．

4.4 評価実験

プロトタイプを用いた被験者実験を行い，提案手法の有効性を評価した．

4.4.1 実験環境

オンライン書籍レビューサイト「読書メーター」[6]に登録しているレビューのうち計9万5,719人のレビュー，全1233万1749件を対象として，Doc2Vecの学習を行った．日本語解析にはMeCabを用いて，辞書はmecab-ipadic-neologd [44]を使用した．ユーザに提示するレビューは，読書メーターに投稿されているレビューを用いた．また，4.1におけるパラメータ検証のため，表4.2に示すパラメータの組み合わせによって予測評価値の計算を行った．

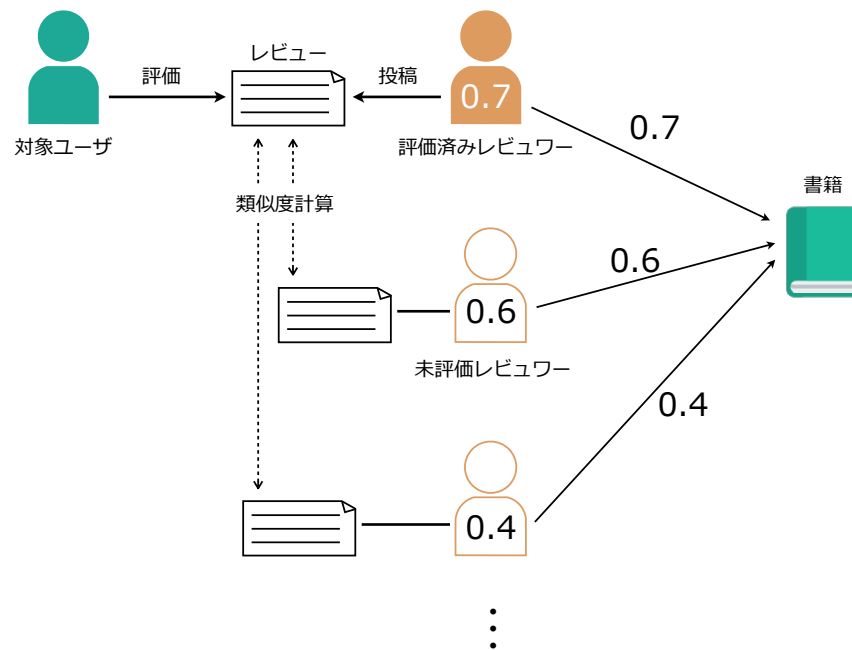


図4.6 レビューの重みを利用した推薦システム

4.4.2 評価手法

4.1による実験の対象とした被験者は20代の男女12人であった。4.2による実験の対象とした被験者は20代の男女8人であった。被験者には、最近読んだ書籍の中から、自由に書籍を選んでもらい、その書籍のレビューに対して、評価のフィードバックを入力してもらった。評価については「このレビューを支持するか」という基準で「支持する」、「わからない」、「支持しない」の3段階の評価とし、「支持する」「支持しない」がそれぞれ10件以上入力されるまでを入力してもらった。プロトタイプシステムを通して得られた推薦結果上位20件のうち、未読書籍について「あらすじ」と「レビュー」を読んでもらい、「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の5段階でアンケート調査を行った。書籍の「1巻」「2巻」などは、まとめてシリーズとしての評価を行った。ベースラインは、レビューに対する重みを均一に設定した協調フィルタリングによる推薦とした。また、文書の典型性を考慮することの有効性を評価するために、典型性を考慮しない場合の提案手法による推薦も行った。

表4.2 実験で使ったパラメータの組み合わせ

	α	β	γ
set1	0.0	0.5	0.5
set2	0.1	0.4	0.5
set3	0.2	0.3	0.5
set4	0.3	0.2	0.5
set5	0.4	0.1	0.5
set6	0.1	0.5	0.4
set7	0.2	0.5	0.3
set8	0.3	0.5	0.2
set9	0.4	0.5	0.1
set10	0.1	0.3	0.6
set11	0.2	0.2	0.6
set12	0.2	0.1	0.7

4.4.3 実験結果

観点の違いを考慮したレビュー者の重み付けによる結果

「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の評価を、それぞれ5点、4点、3点、2点、1点として、手法ごとに、推薦書籍への評価の平均値を求めた。提案手法とベースラインのスコアを図4.7に示す。縦軸は推薦書籍への評価の平均値であり、横軸は左からベースライン、提案手法である。このときの提案手法のパラメータは、 $\alpha = 0.1$, $\beta = 0.3$, $\gamma = 0.6$ であった。

ベースラインと提案手法において、推薦への評価に有意な差があるかどうかを、t検定を用いて調べた結果、有意水準1%において有意な差が見られた。

また、図4.8は γ を最大（0.5）に固定し、 $\alpha + \beta = 0.5$ になるような条件で α を遷移させたときの、推薦結果への評価を表す。図4.9は β を最大（0.5）に固定し、 $\alpha + \gamma = 0.5$ になるような条件で α を遷移させたときの、推薦結果への評価を表す。

確信度に基づくレビュー者の重み付けによる結果

「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の評価を、それぞれ5点、4点、3点、2点、1点として、手法ごとに、推薦書籍への評価の平均値を求めた。提案手法とベースラインのスコアを図4.10に示す。縦軸は推薦書籍への評価の平均値であり、横軸は左からベースライン、提

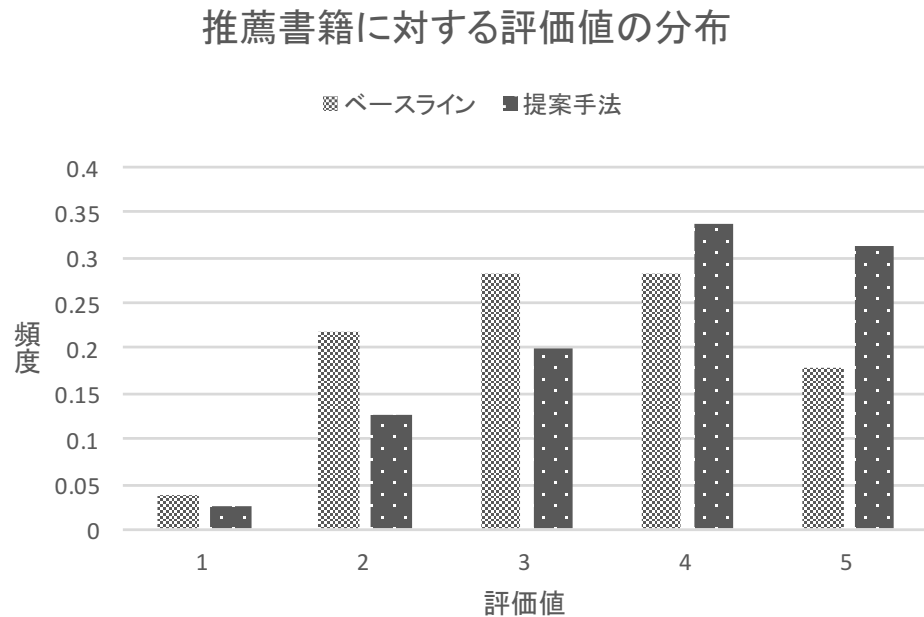


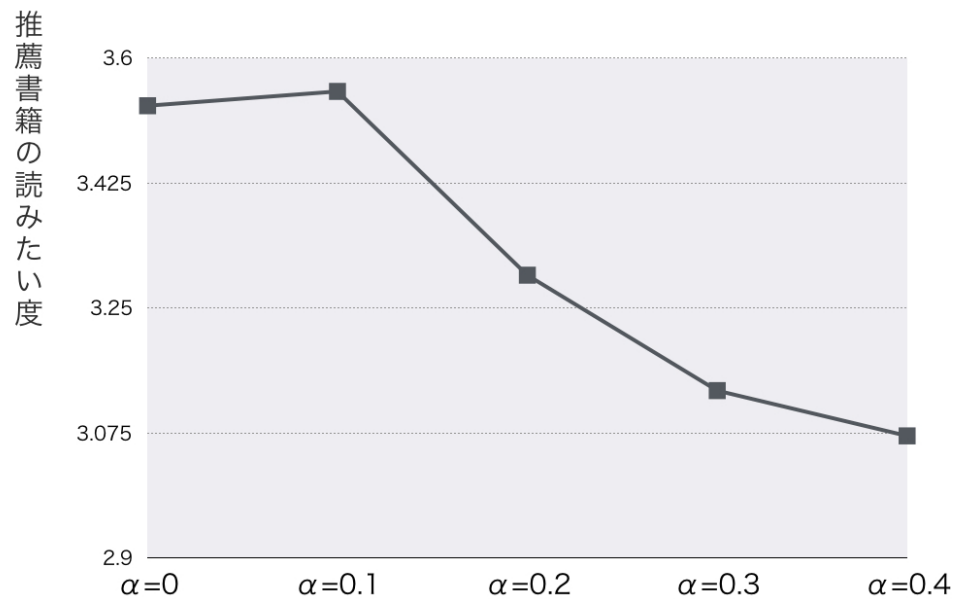
図4.7 提案手法とベースラインの結果

案手法（確信度あり）、提案手法（確信度なし）である．また、評価値ごとの頻度分布を図4.11に示す．

4.5 議論

4.5.1 観点の違いを考慮したレビュー者の重み付け

提案手法とベースラインとの間で、有意水準1%において有意な差が見られたことにより、提案手法は書籍推薦の精度を上げるうえで有効であることが分かった．また、適切なパラメータに関する幾つかの知見が得られた．まず、レビュー者の評価観点ベクトルの影響の大きさを表す α の値に関しては、フィードバックによって計算される対象ユーザの評価嗜好の影響の大きさを表す β と γ に対して、同等かそれ以上の値になると、ほとんどのレビュー者の共感度が1に近い値となり、レビュー者ごとに差が見られなかったことから、ベースラインとほとんど変わらない推薦内容となった．そのため、提案手法においては、 α は、 β と γ に対して、比較的小さな値が適している．また、ポジティブなフィードバックの影響の大きさを表す β とネガティブなフィードバックの影響の大きさを表す γ との間においては、 $\beta < \gamma$ の方が推薦への評価の平均値が高い傾向にあり、 $\alpha = 0.1$ のとき、 $\gamma < \beta$ と $\beta < \gamma$ との間で t 検定を行った結果では、有意水準5%において、有意な差が見られた．このことから、ユーザの好みではない観点をもつレビュー者の重要度を低く

図4.8 γ を最大に固定したときのパラメータ推移の結果

することが、ユーザにとって精度の良い推薦結果につながると考えられる。

4.5.2 確信度に基づくレビューの重み付け

提案手法とベースラインとの間で、有意水準5%において有意な差が見られ、また、提案手法（確信度あり）はベースラインと比較して、評価値1～3の頻度が低く、反対に評価値4～5の頻度が高いことから、提案手法は書籍推薦の精度を上げるうえで有効であることが分かった。また、提案手法の中でも、レビューの確信度を考慮した手法としなかった手法を比較すると、確信度を考慮した手法の方が、考慮しなかった手法よりも、評価の平均値が高く、また評価値の頻度分布においても、評価値4の頻度が考慮しない手法より高く、評価値1～3の頻度は低かった。このことから、レビューの確信度を考慮することが信頼度推定の精度を向上させる上で有効であることがわかった。

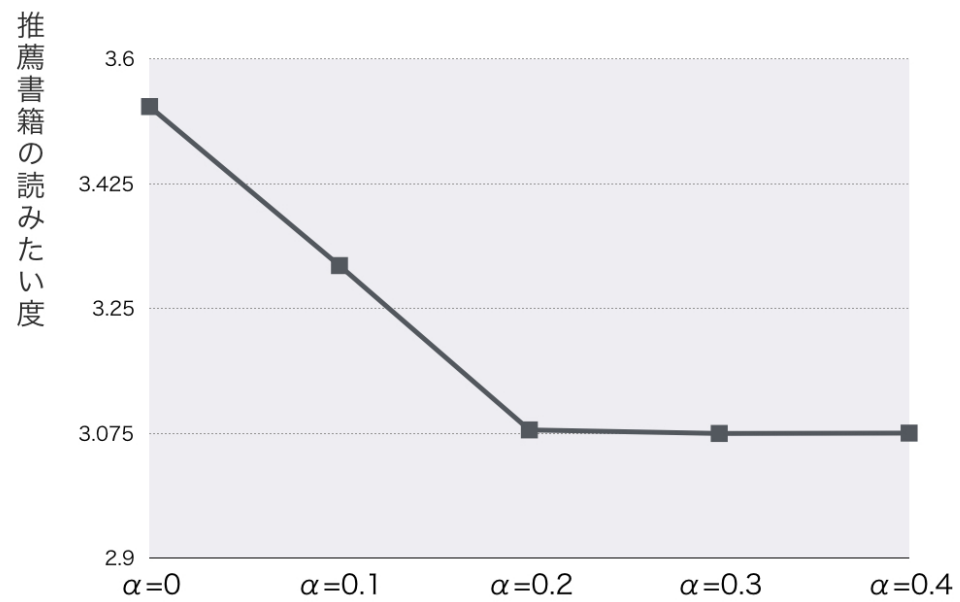
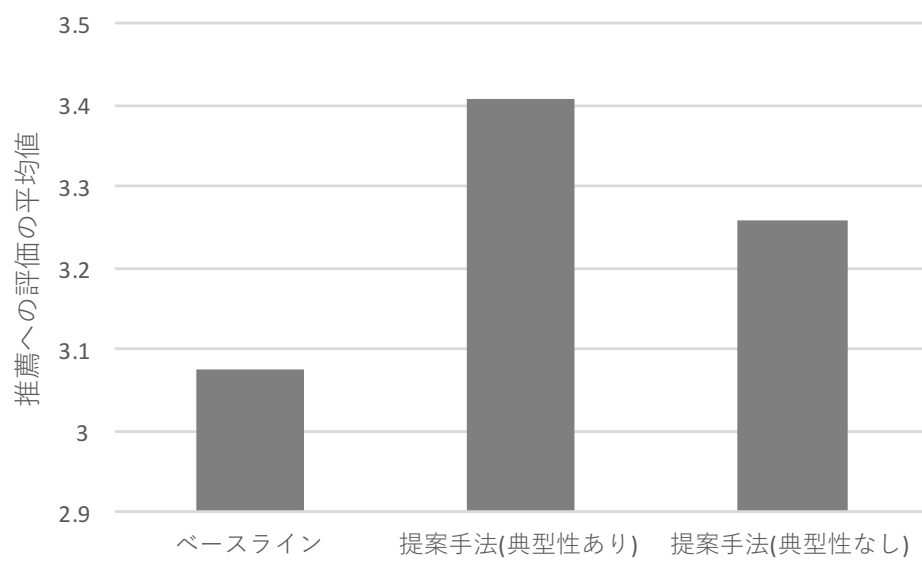
図4.9 β を最大に固定したときのパラメータ推移の結果

図4.10 推薦書籍への平均評価値

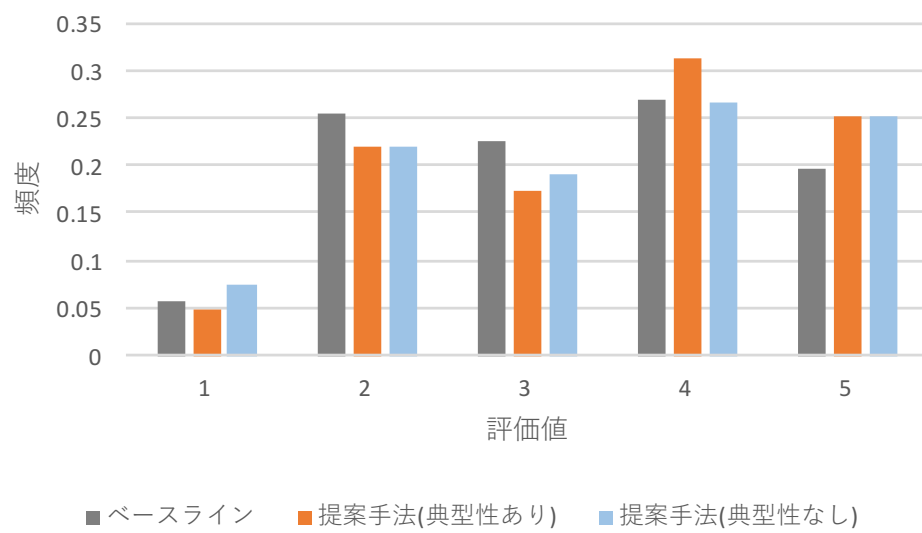


図4.11 評価値の頻度分布

第5章

レビュー者の信頼度に基づくアイテムの品定め支援

5.1 アイテムの品定めにおけるレビュー推薦の必要性

アイテム推薦によって、検討候補のアイテムが絞り込まれた後、ユーザは対象のアイテムが本当に自分にとって価値のあるものであるのかについて品定めを行うことが考えられる。未知のアイテムの品定めをする際に、レビューはアイテムの使用感や微細な情報を得る上で有用である。しかし、現在オンラインレビューサイト上には莫大な数のレビューが存在し、有用なレビューを効率的に発見することは難しいという問題がある。また、投稿されているレビューは対象ユーザにとって役に立つレビューばかりではなく、中には役に立たないレビューも存在する場合がある。そこで、対象ユーザにとって役に立つレビューを推薦することにより、推薦アイテムの価値判断をより効率的に行うことができると考えられる。

5.2 レビュー者の信頼度に基づくレビュー推薦手法

コンテンツの品定めをする上で有用なレビューは、対象ユーザと感覚が近いレビュー者のレビューであると考えられる。本章では、ユーザの共感できるレビューを推定し、上位に表示することで、効率的な品定めを支援する手法を提案する。図5.1は、提案手法の概念図である。既知のアイテムAのレビューに対する評価のフィードバックに基づいて、4.2で得られたレビュー者に対する信頼値を用いて、未知のアイテムBのレビューの価値を推定する。信頼値の高いレビュー者と似ているレビュー者も信頼できるという仮定に基づいて、対象ユーザ u に対するレビュー d の推薦スコアは以下の式で計算される。

$$\text{score}(u, d) = \frac{\sum_{r \in RVR(PR, NR)} \text{trust}(u, r) \times \text{sim}(r, r_d)}{\sum_{r \in RVR(PR, NR)} \text{sim}(r, r_d)} \quad (5.1)$$

ここで、 PR はユーザ u の投稿したレビュー集合のうち、対象ユーザからポジティブな評価を受けたレビュー集合を表し、 NR はユーザ u の投稿したレビュー集合のうち、対象ユーザからネガティブな評価を受けたレビュー集合を表す。 r_d はレビュー d を投稿したレビュー者を表しており、4.2.1で推定したレビュー r の信頼値に対して、レビュー r とレビュー r_d の類似度を掛け合わせ、重み付き平均を計算することで、ユーザ u に対するレビュー d の推薦スコアを計算している。レビュー r 同士の類似性 $\text{sim}(r, r_d)$ は、文章の分散表現を獲得できるDoc2Vec [41] を用いて計算される。具体的には、レビュー r が投稿した全てのレビューを一纏めにした文書をDoc2Vecによってベクトル化し、ベクトル同士の類似度をコサイン類似度によって計算する。レビュー集合 RVR は以下の式で決定される。

$$RVR(PR, NR) = \bigcup_{d \in PR \cup NR} \text{reviewer}(\text{item}(d)) \quad (5.2)$$

ここで、 $\text{item}(d)$ はレビュー d が投稿されているアイテム集合を表し、 $\text{reviewer}(\text{item}(d))$ はレビュー d が投稿されているアイテム集合にレビューを投稿しているレビュー者集合を表している。

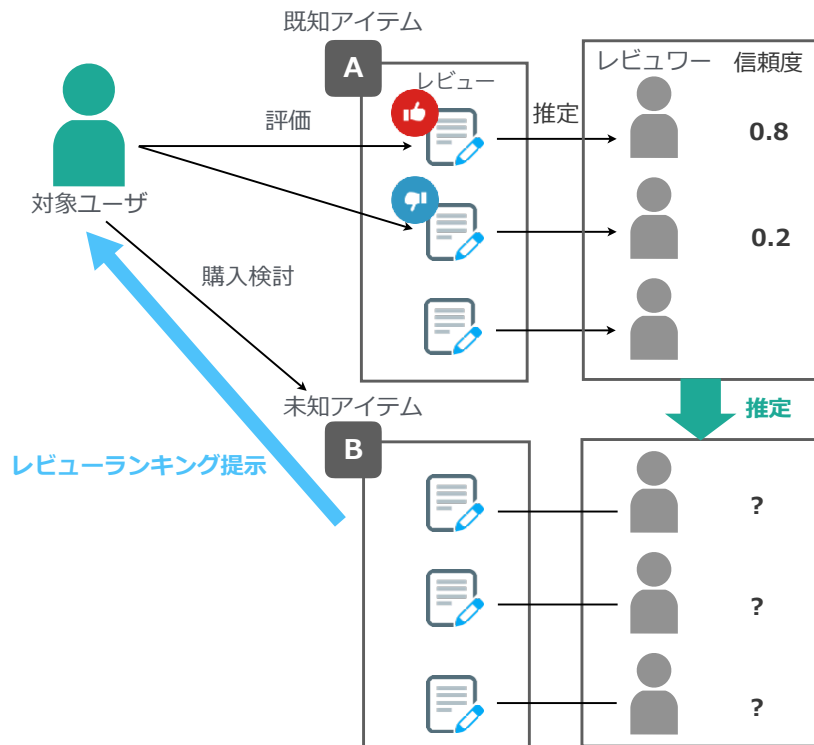


図5.1 提案手法の概念図

5.3 著者推定モデルによるレビューの確信度計算

5.3.1 Deep Learningを用いた著者推定

4.2.3では、評価されたレビューがどれだけそのユーザーらしい発言であるかという確信度を信頼推定モデルの中に組み込むことで、計算される信頼度の確信性を向上させた。本節では、Deep Learningに基づくレビューの著者推定モデルを用いて、レビューの確信度の計算を行う手法について述べる。図5.2は、著者推定のモデル図である。具体的には、一件のレビューの意味情報と語彙情報をDoc2Vec、文字に基づく1-gramを用いて表現した文書ベクトルを入力データとし、ユーザーに紐づいたID情報を正解値として、隠れ2層のニューラルネットワークにより学習を行う。学習が収束したモデルに対して、任意のレビューを入力することで得られる出力は、各レビュー者に対する確率値であり、任意のレビューがレビュー者らしい文書である時に高くなり、レビュー者らしくない場合には低くなると考えられる。

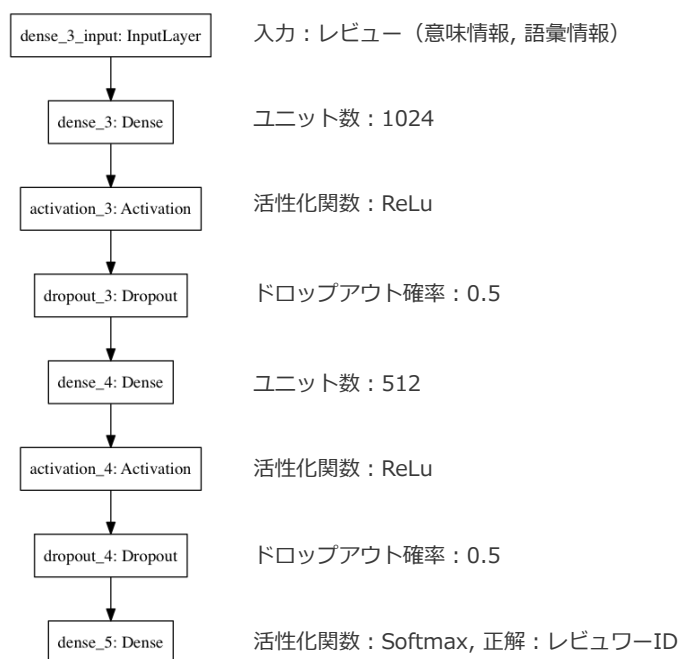


図5.2 著者推定モデル

5.3.2 予備実験

5.3.1の著者推定モデルに基づいて、オンラインレビューサイトのレビュー者に関する著者推定の精度について、検証を行なった。オンライン書籍レビューサイト「読書メーター」のレビュー

データを用いた。対象としたレビュー者は、2015年に出版された小説、「火花」(著: 又吉 直樹)にレビューを投稿したレビュー者498人である。1レビュー者あたりのレビューの平均投稿件数は399件であった。精度指標はAccuracy, 誤差関数はCategorical Crossentropyを用いた。図5.3は, エポック数40で学習させた時の, 精度の推移を示した結果である。図5.4は, エポック数40で学習させた時の, 誤差の推移を示した結果である。検証データにおいては, 15エポック学習させた時点で, 精度および誤差が収束している。これらの結果より, 数百人程度の著者推定タスクであれば, 3割程度の精度で著者推定モデルを構築できることがわかった。

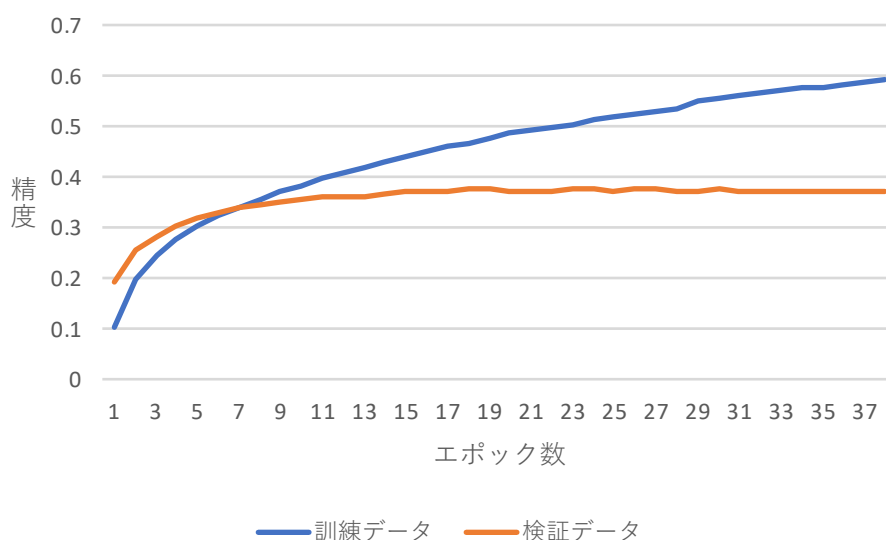


図5.3 学習による精度の推移

5.4 評価実験

5.4.1 実験環境

ユーザに提示するレビューおよび推薦に利用するアイテム評価データは, オンライン書籍レビューサイト「読書メーター」を用いた。日本語解析にはMeCab [43]を用いて, 辞書はmecab-ipadic-neologd [44]を使用した。著者推定モデルの学習エポック数は30とした。

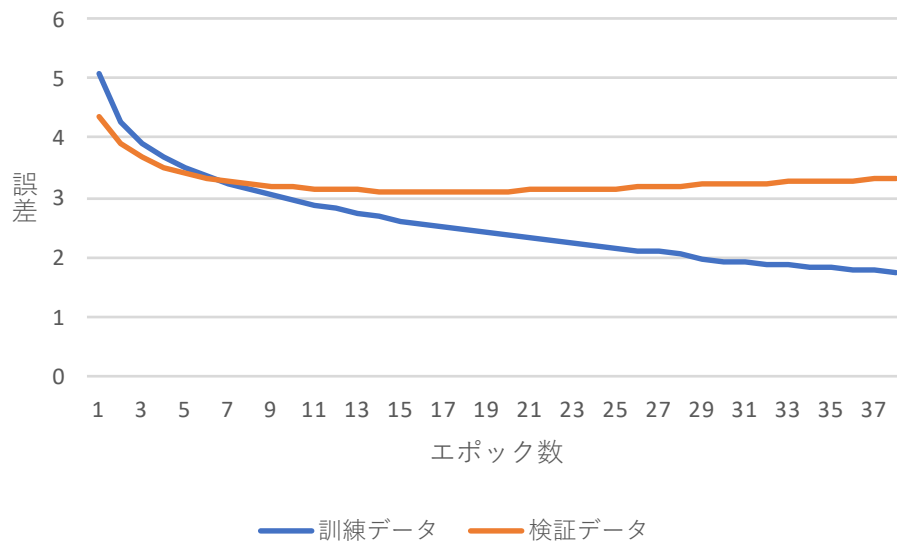


図5.4 学習による誤差の推移

5.4.2 評価手法

被験者には、最近読んだ書籍の中から、自由に書籍を2冊選んでもらい、2冊の書籍のレビューに対して、評価のフィードバックを入力してもらう。評価については「このレビューに共感するか」という基準で「共感する」、「わからない」、「共感しない」の3段階の評価とし、それぞれ100件のレビューを評価してもらう。片方の書籍のレビューに対する評価データを学習データとして、もう一方の書籍のレビューに対する推薦スコアを予測する。提案手法を通して得られた推薦結果上位10件に含まれる、「共感する」、「わからない」、「共感しない」の割合を算出した。実験の対象とした被験者は20代の男女6人で、学習用データと検証用データを交差させた12サンプルに対する推薦結果の平均値を算出した。

5.4.3 比較手法

比較手法は、ランダムサンプリング、読書メーターに設置されている投票ボタンによる投票数の多い順、サポートベクトル回帰、ランダムフォレスト回帰とする。



図5.5 読書メーターにおけるレビューの表示画面

投票数の多い順

図5.5は、読書メーターにおける、レビューの表示画面である。読書メーターでは、レビュー文の左下に表示されているアイコンが評価の入力ボタンであり、ユーザはボタンをクリックすることによって、レビューを評価することができる。多くのユーザから投票を獲得したレビューは、多くの人が良いと思ったレビューであり、グローバルな品質が高いと考えられる。本実験では、投票数の多い上位10件と提案手法による上位10件のレビューを比較することで、提案手法が対象ユーザの好みに合ったレビューを提示できているかについて検証する。

サポートベクトル回帰・ランダムフォレスト回帰

本実験では、レビュー推薦手法の中における、提案手法の有効性を検証するために、回帰分析を用いたレビューのスコア予測による手法との比較を行う。回帰分析とは、1つあるいは複数の変数の値（説明変数）を用いて、ある一つの変数の値（目的変数）を予測するために用いられる多変量解析手法の一つである。今回は、比較手法として、サポートベクトル回帰（SVR）、ランダムフォレスト回帰（RFR）を用いる。SVRは、分類学習における代表的な手法である、サポートベ

クターマシン（SVM）を応用した回帰分析手法である。RFRは、データから抽出されるルールを、データ分類やモデリング、値の予測などに応用する決定木を弱学習器として、アンサンブル学習するアルゴリズムである。

回帰分析の説明変数は、単語のTF-IDFベクトル、レビューに含まれる各品詞の割合（合計13品詞）、単語の合計数、単語の種類数とした。目的変数は、被験者からレビューに対して与えられた、「共感する」、「どちらでもない」、「共感しない」の評価を、それぞれ1点、0点、-1点として、レビューに与えた尺度変数とした。回帰分析によって、検証データの説明変数を入力して得られる目的変数の予測値が高いほど、より「共感できる」と予測されたレビューであると考えられ、予測値の高い順に並び替えることで、回帰分析に基づく、レビュー推薦を行うことができる。

5.4.4 実験結果

提案手法と比較手法により並び替えられたレビューランキング上位10件のうち、「共感する」という評価を入力されたレビューが存在する割合を図5.6に示す。縦軸は「共感する」という評価の割合であり、横軸は左からランダムサンプリング、投票数の多い順、サポートベクトル回帰（SVR）、ランダムフォレスト回帰（RFR）、提案手法である。提案手法とランダムサンプリング、投票数順との間で、有意水準5%において有意な差が見られた。また、提案手法と比較手法により並び替えられたレビューランキング上位10件のうち、「共感しない」という評価を入力されたレビューが存在する割合を図5.7に示す。提案手法とランダムサンプリング、サポートベクトル回帰との間で、有意水準5%において有意な差が見られた。

5.5 議論

提案手法に基づいて並び替えられた上位10件レビューは図5.6に示す全ての手法の中で、「共感する」という評価を入力されたレビューの割合が最も大きいことから、対象ユーザの品定め役に立つレビューを上位に提示する上で、有効であることがわかった。また、提案手法は図5.7に示す全ての手法の中で、「共感しない」という評価を入力されたレビューの割合において、提案手法は最も少ないことから、対象ユーザにとって品定め役に立たないレビューをフィルタリングする上で、有効であることがわかった。

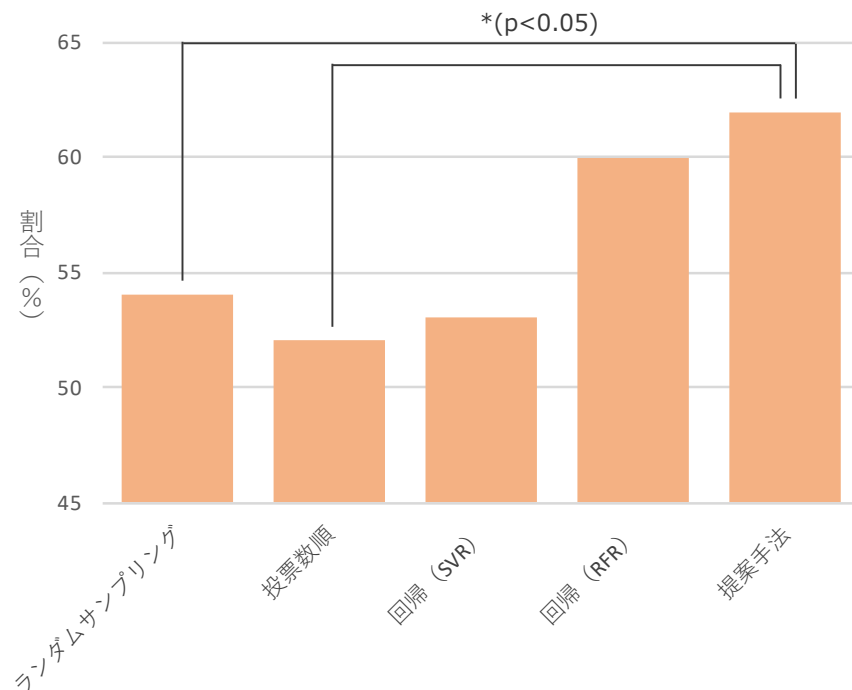


図5.6 各手法上位10件のレビューの中に含まれる「共感する」の割合

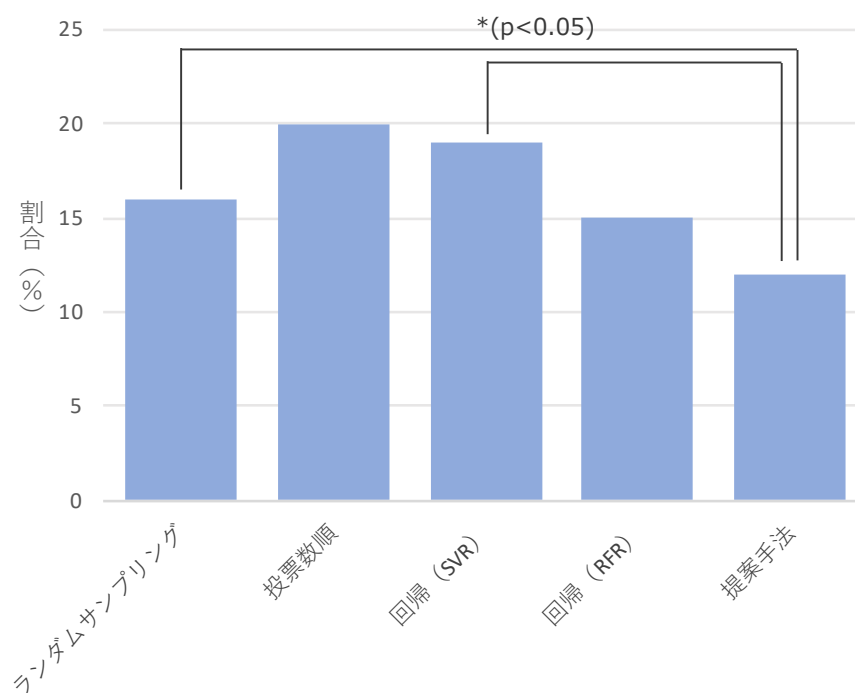


図5.7 各手法上位10件のレビューの中に含まれる「共感しない」の割合

第6章

まとめ

本研究では、インターネット上のアイテムが膨大になり、レビューの役割が重要になっている状況において、レビューの価値はユーザの嗜好に合わせて変動させるべきであるという考えに基づいて、アイテムの絞り込みと品定め観点から、アイテムの選別を支援することを目的とした。具体的には、アイテムの絞り込みにおいては、協調フィルタリングによるアイテム推薦の最適化、アイテムの品定めにおいては、レビュー推薦を行った。

対象ユーザにおけるそれぞれのレビューの価値を計算するため、3章で示した既知のアイテムのレビューに対する評価のフィードバックを入力として、4.1では、レビューの観点の違いを考慮したレビューの価値推定手法を提案し、4.2では、レビューがレビューらしい発言であるかという確信性を考慮したレビューの価値推定手法を提案した。4.3において、それぞれの手法により得られたレビューの価値を考慮した協調フィルタリングによるアイテム推薦手法を提案した。提案手法のプロトタイプシステムを実装し、システムで出力された推薦結果の有効性を検証するため、読書メーターのレビューと書籍の登録データを収集し、レビューの重みが均一である協調フィルタリングをベースラインとして、被験者実験を行なった。実験の結果、より精度の高いアイテム推薦を行うために、それぞれの提案手法が有効であることが示された。

また、5章では、既知のアイテムのレビューに対する評価のフィードバックに基づいて、未知のアイテムのレビューに対する価値を推定する手法を提案した。提案手法によって並び替えられたレビューランキングがユーザの品定め役に立つかを検証するため、読書メーターのレビューを用いて、被験者実験を行なった。実験の結果、提案手法は対象ユーザにとって、共感できるレビューを上位に提示する割合が比較手法より高く、共感できないレビューは上位に提示する割合が低かったことから、アイテムの品定めにおいて、提案手法は有効であることが示された。

今後の課題として、まず、書籍以外のオンラインレビューサイトのデータを用いた検証実験が考えられる。書籍を始め、映画や音楽などは、レビューに対する好みがユーザによって別れることが考えられるため、提案手法を適応できると考えられるが、一般的な商品などにおいては、レ

ビューの客観的な信頼性の方が重要な要素である可能性がある。そのため、客観的な価値と主観的な価値の両方を考慮した手法の開発を検討している。また、SNSやシェアリングエコノミーなどにおいても、提案手法を適応できる可能性がある。具体的には、SNSにおけるタイムラインのフィルタリングや、シェアリングエコノミーにおけるアイテム推薦などが考えられる。

謝辞

三年間，本研究に際して，熱心なご指導をいただきました牛尾剛聡准教授に心から感謝いたします。

また，様々なアドバイスをいただきました牛尾研究室の皆様，本当にありがとうございました。

参考文献

- [1] 楽天市場, <http://www.rakuten.co.jp>
- [2] <http://www.rakuten.co.jp/com/inc/rc/info.html>
- [3] “Google Books:Books of the world, stand up and be counted! All 129,864,880 of you.”, <http://booksearch.blogspot.jp/2010/08/books-of-world-stand-up-and-be-counted.html>, (2017/01/22アクセス)
- [4] Christopher D. Manning, Prabhakar Raghavan, Hinrich Schutze, “Introduction to Information Retrieval”, Cambridge University Press, 2008.
- [5] @cosme, <http://www.cosme.net/>
- [6] 読書メーター, <http://bookmeter.com/>
- [7] Yahoo!映画, <http://movies.yahoo.co.jp/>
- [8] ”GDP に 現 れ な い ICT の 社 会 的 厚 生 へ の 貢 献 に 関 す る 調 査 研 究”, http://www.soumu.go.jp/johotsusintokei/linkdata/h28_04.houkoku.pdf
- [9] 土方 嘉徳, “利用者の好みをとらえ活かす-嗜好抽出技術の最前線- : 1.嗜好抽出・情報推薦の基礎理論 1) 嗜好抽出と情報推薦技術”, 情報処理, Vol.48, No.9, pp.957-965, 2007.
- [10] M.L.Anderson, J.R.Magruder. ”Learning from the crowd: Regression discontinuity estimates of the effects of an online review database”, Economic Journal, 122(563):957-989, 2012.
- [11] Foltz, P.W.: Using Latent Semantic Indexing for Information Filtering, Proc. of ACM Conference on Office Inforamtion Systems, pp.40-47, 1990.
- [12] 土方 嘉徳, “情報推薦・情報フィルタリングのためのユーザプロファイリング技術,” 人工知能学会誌, vol.19, no.3, pp.365-374, 2004.
- [13] 顔 洪, 牛尼 剛聡, “スマートフォンでの効率的な商品選別を目的としたユーザの振舞いに基づく閲覧リスト最適化手法”, 情報処理学会論文誌データベース (TOD), Vol.8, No.4, pp.1-15, 2015.
- [14] A Statistical Analysis of 1.2 Million Amazon Reviews:

- <http://minimaxir.com/2014/06/reviewing-reviews>
- [15] S. Moghaddam, M. Jamali, and M. Ester. ETF: Extended Tensor Factorization model for personalizing prediction of review helpfulness. In *Procs. of the WSDM*, Seattle, WA, USA, 2012.
- [16] Raghavan, S., Gunasekar, S., Ghosh, J.: Review quality aware collaborative filtering. In: *Proceedings of the 6th ACM Conference on Recommender systems*, pp. 123-130, 2012.
- [17] 伊木 淳, 亀井 清華, 藤田 聡, “レビューを対象とした信頼性判断支援システムの提案”, *情報処理学会論文誌*, Vol.55, No.11, pp.2461-2475, 2014.
- [18] Mukherjee, A., Liu, B. and Glance, N.: Spotting Fake Reviewer Groups in Consumer Reviews, *Proc. 21st International Conference on World Wide Web*, pp.191-200 2012.
- [19] Xie, S., Wang, G., Lin, S. and Yu, P.S.: Review SpamDetection via Temporal Pattern Discovery, *Proc. 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.823-831, 2012.
- [20] Wang, G., Xie, S., Liu, B. and Yu, S.: Review Graph based Online Store Review Spammer Detection, *Proc. 11th IEEE International Conference on Data Mining*, pp.1242-1247, 2011.
- [21] Garcia Esparza, S., O’ Mahony, M.P., Smyth, B.: Effective product recommendation using the real-time web. In: *Proceedings of the 30th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, Cambridge, UK, Springer, pp. 5-18, 2010
- [22] Garcia Esparza, S., O’ Mahony, M.P., Smyth, B.: A multi-criteria evaluation of a user-generated content based recommender system. In: *Proceeding of the 3rd Workshop on Recommender Systems and the Social Web in RecSys’ 11*, Chicago, Illinois, USA, pp. 49-56, 2011
- [23] Leung, C.W.K., Chan, S.C.F., Chung, F.: Integrating collaborative filtering and sentiment analysis: A rating inference approach. In: *Proceedings of the ECAI 2006 Workshop on Recommender Systems*, Riva del Garda, Italy, pp. 62-66, 2006
- [24] Zhang, W., Ding, G., Chen, L., Li, C., Zhang, C.: Generating virtual ratings from chinese reviews to augment online recommendations. *ACM Trans. Intell. Syst. Technol.* 4(1),9 , 2013
- [25] Poirier D, Fessant F, Tellier I.: Reducing the cold-start problem in content recommendation through opinion classification. In: *Proceedings of the 2010 IEEE/WIC/ACM In-*

- ternational Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, Toronto, Canada, IEEE Computer Society, WI-IAT' 10, pp. 204-207, 2010
- [26] Musat, C.C., Liang, Y., Faltings, B.: Recommendation using textual opinions. In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, AAAI Press, IJCAI' 13, pp. 2684-2690, 2013
- [27] McAuley, J., Leskovec, J.: Hidden factors and hidden topics: Understanding rating dimensions with review text. In: Proceedings of the 7th ACM International Conference on Recommender Systems, Hong Kong, China, ACM, RecSys' 13, pp. 165-172, 2013
- [28] Seroussi, Y., Bohnert, F., Zukerman, I.: Personalised rating prediction for new users using latent factor models. In: Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia, Eindhoven, The Netherlands, ACM, HT' 11, pp 47-56, 2011
- [29] Wang, Y., Liu, Y., Yu, X.: Collaborative filtering with aspect-based opinion mining: A tensor factorization approach. In: Proceedings of the IEEE International Conference on Data Mining, Brussels, Belgium, IEEE Computer Society, ICDM' 12, pp. 1152-1157 (2012)
- [30] Liu, H., He, J., Wang, T., Song, W., Du, X.: Combining user preferences and user opinions for accurate recommendation. *Electron. Commer. Res. Appl.* 12(1), 14-23, 2013
- [31] Chen, L., Wang, F.: Preference-based clustering reviews for augmenting e-commerce recommendation. *Knowl. Based Syst.* 50, 44-59, 2013
- [32] Levi, A., Mokryn, O., Diot, C., Taft, N.: Finding a needle in a haystack of reviews: Cold start context-based hotel recommender system. In: Proceedings of the 6th ACM International Conference on Recommender Systems, Dublin, Ireland, ACM, RecSys' 12, pp 115-122, 2012
- [33] Li Chen, Guanliang Chen, Feng Wang, Recommender systems based on user reviews: the state of the art. *User Modeling and User-Adapted Interaction*, Vol. 25, pp. 99-154, 2015
- [34] 林 貴宏, 尾内 理紀夫, 「Web上のレビューを利用した映画推薦システム」, 人工知能学会論文誌, Vol.30, No. 1, pp.102-111, 2015.
- [35] Koren, Y., Bell, R. and Volinsky, C.: Matrix factorization techniques for recommender systems, *Computer* 42(8), pp.30-37, 2009.
- [36] Melville, P., Mooney, R. J. and Nagarajan, R.: Content boosted collaborative filtering for improved recommendations, *Proc. AAAI*, pp.187-192, 2002.

- [37] Wang, C. and Blei, D.M.: Collaborative topic modeling for recommending scientific articles, Proc. ACM SIGKDD, pp.448-456, 2011.
- [38] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.: GrxfoupLens: An Open Architecture for Collaborative Filtering of Netnews, Proc. ACM Conf. on Computer Supported Cooperative Work (CSCW' 94), North Carolina, pp.175-186, 1994.
- [39] David M Blei, Andrew Y Ng, and Michael I Jordan. "Latent dirichlet allocation", the Journal of machine Learn-ing research, Vol. 3, pp. 993-1022, 2003.
- [40] Thomas L. Griffiths, Mark Steyvers : Finding Scientific Topics, Proc. National Academy of Sciences, vol.101, pp.5228-5235, 2004.
- [41] Quoc V, Leand Tomas Mikolov. DistributedRepresentations of Sentences and Documents, In Proceedings of The 31st In- ternational Conference on Machine Learning, pp. 1188-1196, 2014.
- [42] Tomas Mikolov, Greg Corrado, Kai Chen, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space. In Proceedings of the International Conference on Learning Representations (ICLR 2013), pp.1-12, 2013.
- [43] <http://mecab.sourceforge.net/>
- [44] <https://github.com/neologd/mecab-ipadic-neologd/blob/master/README.ja.md>

業績一覧

- [1] 南 大智, 牛尼 剛聡, “書評SNSにおけるレビュワーの観点の違いを考慮したフィードバック型協調フィルタリング”, 第8回データ工学と情報マネジメントに関するフォーラム, 2016
- [2] 南 大智, 牛尼 剛聡, “書評レビューへのフィードバックを用いた推薦書籍の最適化手法”, 第60回システム制御情報学会研究会発表講演会, 2016
- [3] 南 大智, 牛尼 剛聡, “レビュワーに対する共感度を考慮した協調フィルタリングによるアイテム推薦手法”, 第9回Webとデータベースに関するフォーラム, 2016
- [4] Daichi Minami, Taketoshi Ushiyama, “A collaborative filtering method based on empathy with reviewers”, ACM Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication, Article No. 33, 2017
- [5] 牛島美桜, 南 大智, 牛尼 剛聡, “視聴者はドラマの「どこ」に「どう」反応しているのか - 実況ツイートを利用したドラマのシーン特徴の抽出 -”, 第9回データ工学と情報マネジメントに関するフォーラム, 2017
- [6] 南 大智, 牛尼 剛聡, “このユーザは信頼に値するか? - SNSにおける協調的な信頼推定モデル -”, 第9回データ工学と情報マネジメントに関するフォーラム, 2017
- [7] Daichi Minami, Taketoshi Ushiyama, “Can you trust the user? Collaborative Trust Estimation Model for Recommendations”, IEEE Twelfth International Conference on Digital Information Management, 2017
- [8] Daichi Minami, Mio Ushijima, Taketoshi Ushiyama, “How do viewers react to drama? Extraction of scene features of dramas from live commentary tweets”, ACM Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication, 2018
- [9] 南 大智, 牛尼 剛聡, “推薦アイテムの品定め支援のための共感レビュー提示手法”, 第10回データ工学と情報マネジメントに関するフォーラム, 2018 (予定)