

ユーザの書棚整列順序に基づく書籍推薦手法

宮本 達矢[†] 北山 大輔^{††}

[†] 工学院大学大学院工学研究科情報学専攻

〒163-8677 東京都新宿区西新宿 1-24-2

^{††} 工学院大学情報学部システム数理学科

〒163-8677 東京都新宿区西新宿 1-24-2

E-mail: [†]em18014@ns.kogakuin.ac.jp, ^{††}kitayama@cc.kogakuin.ac.jp

あらまし ユーザが書籍を購入した際、多くの場合書籍を書棚に収納する。電子書籍においても、電子上の書棚に書籍を登録することで、ユーザにとってわかりやすく保管することができる。このような書棚への本の配置にはユーザの嗜好が現れていると考えられる。そこで本研究では書籍の並び順を考慮し、その並び順の中に挿入するのにふさわしい書籍を推薦する手法を提案する。まず、レビュー中から書籍の内容及びレビュアーの得た感情を用いて書籍の特徴ベクトルを作成する。次に、書棚の書籍整列順序に合わせて並べた特徴ベクトルから、要素ごとに回帰分析を行うことで推薦予測ベクトルを作成し書籍特徴ベクトルとの類似度を算出することで、書籍並びに挿入すべき書籍を推薦する。

キーワード 書籍推薦, 整列順序, レビュー

A Book Recommendation Method based on User's Book Arrangement

Tatsuya MIYAMOTO[†] and Daisuke KITAYAMA^{††}

[†] Graduate School of Informatics, Kogakuin University

Nishishinjuku 1-24-2, Shinjuku-ku, Tokyo, 163-8677 Japan

^{††} Faculty of Informatics, Kogakuin University

Nishishinjuku 1-24-2, Shinjuku-ku, Tokyo, 163-8677 Japan

E-mail: [†]em18014@ns.kogakuin.ac.jp, ^{††}kitayama@cc.kogakuin.ac.jp

1. はじめに

ユーザが購入した書籍は、多くの場合読み終えたのちにユーザ自身の書棚へと収納する。これは紙の書籍に限ったことではなく、多くの電子書籍サービスでも同じように書棚ごとに書籍を分けて保管することができる。このような書棚は、乱雑に何の基準もなく書籍が収納されるのではなく、多くの場合ユーザ自身が工夫し納得のいく順列で書籍を収納していることが多いと考えられる。たとえば、SF とミステリが並ぶような棚であれば、「本格 SF->SF ミステリ->本格ミステリ->日常ミステリ」のようにジャンルが遷移していくような例があげられる。また、プログラミング言語に関する書籍が並べられた棚であれば、入門書から上級者向け書籍と徐々に難易度が変遷するような順列が考えられる。こうしたユーザの整列意図によって整理された書棚に新たに入る書籍は、既に存在する他の書籍と並んでも違和感のない書籍である必要があると考えた。

しかし、既存の推薦手法では、多くの場合ユーザの最も好む書籍を推薦し、必ずしも特定の書棚の列に当てはまる書籍が推薦されるとは限らない。これは、ユーザの購買履歴や所持書籍を順列関係のない集合として扱っているためである。だが、書棚整列順序にはユーザの嗜好が現れていると考え、整列順序を扱うことでより効果的な推薦ができると考える。

本論文では、ユーザの書籍整列順序に隠された整列意図を汲み取ることで、整列順序により適した書籍の推薦を行う手法の提案を行う。整列順序に合わせた推薦を行うことで、ユーザの興味に合致し、かつユーザにとって受け入れやすい書籍の推薦を行うことが期待できる。提案手法の概要を図 1 に示す。レビューからトピックモデルを用いて特徴ベクトルを作成し、整列順序に沿って並べた特徴ベクトルの各要素に対し回帰分析を行うことで、ある位置の特徴ベクトルを予測する。

本論文の構成は以下の通りである。2 節では関連研究について述べる。3 節では整列順序に基づく書籍特徴の変化予測につ

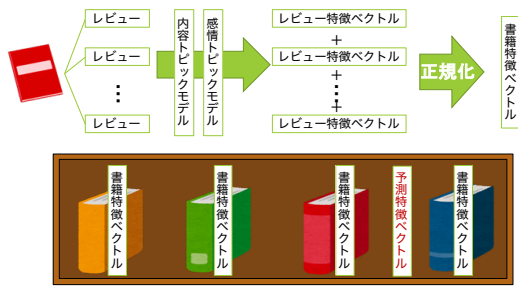


図1 書籍整列順序に基づく書籍推薦手法の概念図

いて述べる．4節では提案手法によって作成した予測ベクトルの予測精度評価を行う．5節では本論文のまとめと今後の課題について述べる．

2. 関連研究

2.1 レビューに含まれるアイテム特徴

ユーザレビューに含まれるアイテムの持つ特徴を利用した推薦や検索は，数多く行われている．

林ら[1]は，レビューから商品の持つポジティブな特徴を抽出することで，同ジャンルの別商品に不満を抱くユーザに不満を解決する代替商品の推薦を行なった．ユーザがレビュー時に商品に対して与える評価値に注目することで，レビュー中に存在する特徴語のポジティブ度合いを算出している．

吉田ら[2]は，主観的特徴を加味した観光スポットの検索を行なった．観光スポットへ投稿されたレビューに対し Paragraph Vector モデルを用いることで，メタデータからは推測困難な主観的特徴を含む検索を可能にしている．

本論文では，トピックモデルを用いることで，レビューから書籍を表す特徴ベクトルの作成を行う．

2.2 トピックモデルを用いた推薦

トピックモデルとは，文書中の単語が文書の潜在的な意味であるトピックに基づき，確率的に出現するという仮定により，文書中の単語からトピックを推定する手法である．代表的な手法として Latent Dirichlet Allocation(LDA) があげられる．

トピックモデルを用いた推薦の例として，加藤ら[3]の手法があげられる．加藤らは，レポートのもつトピックに注目し，レポート間の距離を測ることで，学習者の作成したレポートと同一の主題を扱うレポート検索を行なった．さらに検索したレポートから，学習者のレポートとは構成，表現，オリジナリティが異なるレポートを推薦するシステムを開発した．

2.3 複数アイテムの特徴に基づく推薦

複数のアイテムが持つ特性を組み合わせることにより，新たなアイテムを推薦するというアプローチに関する先行研究について説明する．

奥ら[4]は，ユーザが2つのアイテムを選択することで，元となる2つのアイテムの要素を併せ持つアイテムを推薦するフュージョンベース推薦手法を提案した．材料の組み合わせからユーザは直感的に結果を期待することができ，時に思いがけない結果をもたらすことができることにより，セレンディピティの高いアイテムを推薦している．

本論文においても，整列順序の最低数といえる2冊の書籍に限った場合であれば，2冊の要素を組み合わせたフュージョンベース推薦に近い結果を得ることができると考える．しかし本論文では，より多数の書籍からなる整列順序に注目することで，純粋な集合の要素の組み合わせとは異なる書籍を推薦を行うことを目指している．

また，ユーザの与えた順列から新たなアイテムを推薦する例として，池田ら[5]の手法があげられる．池田らは，楽曲の音響特徴に注目し，プレイリスト中で滑らかに楽曲が遷移するような楽曲の推薦を行なった．楽曲を音響特徴から構成される特徴ベクトルで表し，多次元尺度構成法によって2次元特徴空間へと写像することで楽曲の探索を行なっている．

本論文では，低次元特徴空間への写像を行わずに，特徴ベクトルの個々の要素の値を予測する．また，書籍の持つ特徴のうち，整列順序意図をよく示すような特徴を見つけ出すことで，よりユーザの意図に合致した書籍を推薦することを目指す．

3. 整列順序に基づく書籍特徴の変化予測

ユーザが指定した整列順序は，書籍の持つ特定の特徴に沿った順序であると考えられる．この特徴は列ごとに異なると考えられ，ある列はミステリ作品の謎の難易度に沿っており，また別の列はコメディ要素の多少によって整列されるような例が考えられる．

本論文では，書籍の整列に用いられる特徴は，書籍の持つトピックであると考えた．トピックモデルを用いることで書籍トピックを予測し，整列順序に用いられるトピックを推定し，ある位置に収納されるべき書籍特徴の予測を行う．整列順序に適した推薦を行うことで，ユーザが推薦された書籍と既に所有している書籍との繋がりを理解しやすく，より有意義な推薦を行うことができると考える．

3.1 語の抽出

書籍の特徴を抽出する媒体として，今回はレビューに注目した．ユーザが自由に投稿することのできるレビューには，書籍自体の内容に触れる文章や，書籍を読んだことで得た感想が述べられている．よって，レビュー中に含まれる単語を用いることで，書籍トピックの推定が可能であると考えた．

まず，レビュー中から語を抽出するため，レビューに対し形態素解析を行うことで単語単位にレビューを分割する．同時に，単語の表記揺れを減少させるため，単語を原型に戻し分割を行う．形態素解析器には MeCab を利用した．

本論文では，書籍特徴を表す語として一般名詞，固有名詞，動詞，形容詞，形容動詞，副詞を用いた．しかし，レビューに含まれる登場人物名は書籍固有のものであり，トピック作成には向かないと考えられるため，固有名詞であっても人名は除外した．

3.2 トピックモデル作成

書籍特徴を表すためのトピックモデルを作成する．ある書籍に対し投稿された3.1節の処理を施したレビュー全てをまとめて1文書とし，推薦候補である全ての書籍から作られた文書の集合を全文書とする．この全文書に対し，LDAを用いてトピッ

表 1 LDA による小説トピックの例

topic 1	topic 2	topic 3	topic 4
王	舞台	手紙	ミステリ
女王	戯曲	書簡	ミステリー
騎士	劇	書く	トリック
いる	悲劇	メール	探偵
面白い	ドイツ	往復	いる
ある	冒険小説	やりとり	謎
てる	演劇	いる	ある
バトル	喜劇	送る	面白い
竜	いる	やり取り	推理
主人公	物語	届く	作品

ク抽出を行う。トピック数は予備実験により 125 とした。

表 1 は、推薦候補書籍を小説とした場合のトピック抽出例である。トピック所属確率の高い語から順に、各トピック 10 単語を示している。

3.3 書籍特徴ベクトル作成

作成したトピックを用い、書籍特徴ベクトルを作成する。まず、レビューごとに 3.2 節で作成したトピック t への所属確率 p_t を算出し、所属確率 p_t を各要素としたレビュー特徴ベクトルを作成する。レビュー i のレビュー特徴ベクトル r_i は、式 1 のように表せる。

$$r_i = [p_1, p_2, \dots, p_{125}] (\sum p_i = 1) \quad (1)$$

その後、書籍ごとにレビューのレビュー特徴ベクトルの要素を合算し、L2 正規化を行う。書籍 j の特徴ベクトル b_j は式 2 のように表せる。

$$b_j = \frac{\sum r_k}{\|\sum r_k\|} \quad (2)$$

3.4 書籍特徴ベクトル予測

3.3 節で作成した書籍特徴ベクトルから、回帰分析によってある位置の書籍特徴ベクトルを予測する。回帰分析は特徴ベクトルの要素ごとに行う。説明変数には、書籍をユーザの書棚整列順序に並べたときの書籍位置を順に整数で表した値を与える。目的変数は、予測特徴ベクトルの各要素の値とする。回帰式の算出には、最小二乗法による線形回帰と、ガウスカーネルを用いたカーネルリッジ回帰の 2 つの方法を用いる。

線形回帰を用いた場合、図 2 のように要素ごとに決定係数の差が現れる。図 2 左は決定係数の高くなる要素の例である。決定係数が高くなる要素では、その要素に対応するトピックは、書籍整列順序をよく表すトピックであると考えられる。図 2 右は決定係数の低くなる要素の例である。決定係数が低くなる要素では、その要素に対応するトピックは、書籍整列順序とはあまり関係のないトピックであると考えられる。

作成した回帰式に、整列順序の中の書籍を推薦する位置を説明変数として与えることで、特徴ベクトルの要素の値とする。書籍特徴ベクトルの要素全てにおいて上記の操作を行うことで、書棚整列順序のある位置に入る書籍の予測特徴ベクトル b' を作

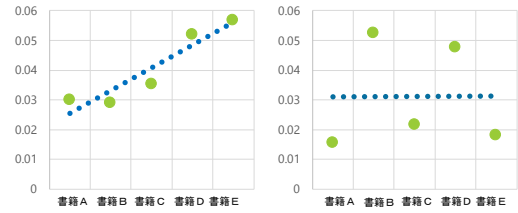


図 2 順序を表すトピックと表さないトピック

成する。

3.5 書籍類似度の算出

予測特徴ベクトルと、推薦対象書籍との類似度を算出する。類似度の算出には、以下の 3 種類の方法を用いる。

- 重み付き cos 尺度による類似度
- 決定係数上位要素のみを用いた cos 尺度
- cos 尺度による類似度

書籍 j の書籍特徴ベクトル b_j と予測特徴ベクトル b' の重み付き cos 尺度は、式 3 のように表せる。特徴ベクトルの要素 k に、要素 k 予測に使用した回帰式の決定係数 w_k を与えることで、整列順序をよく表すと考えられる要素がより強く影響した類似度が算出される。

$$\cos(b', b_j) = \frac{\sum w_k^2 b'_k b_{j,k}}{\sqrt{\sum w_k b'_k} \cdot \sqrt{\sum w_k b_{j,k}}} \quad (3)$$

決定係数上位要素のみを用いた cos 尺度では、予測特徴ベクトル算出時に決定係数の高かった要素のみを抽出した特徴ベクトルを作成する。本論文では、全要素の 10% にあたる 12 要素のみからできる特徴ベクトルを作成する。類似度の算出には、式 3 を用い、常に $w_k = 1$ として計算する。

cos 尺度による類似度では、式 3 を用い常に $w_k = 1$ として予測特徴ベクトルと書籍特徴ベクトルをそのまま用いて類似度を算出する。この場合要素ごとの重みは一切影響せず、全ての要素が等しく類似度に影響する。

3.4 節で述べた 2 種類の回帰分析と、本節で述べた 3 種類の類似度算出手法の組み合わせとして、以下の 4 種類の組み合わせ手法が存在する。

- 線形回帰・要素重み付き cos 尺度（線形重み付き）
- 線形回帰・決定係数上位要素のみを用いた cos 尺度（線形上位）
- 線形回帰・cos 尺度（線形）
- カーネルリッジ回帰・cos 尺度（カーネルリッジ）

カーネルリッジ回帰では、書棚の整列順序空間から非線形写像を行なうため、算出された決定係数はトピックの整列順序への影響度合いとは考えない。よって、カーネルリッジ回帰では決定係数を用いた類似度算出手法は除いている。各手法で類似度を算出したのち、類似度の高い書籍を順にユーザへと推薦する。

4. 評価実験

提案手法の有効性を検証するため、評価実験を行なった。3.5

表 2 各セットに用いた書籍の選定基準

セット	選定基準
小説ランキング	honto - 発表!! 2017 年年間ランキング【小説】: 電子書籍 [6] [7]
小説ミステリ	このミステリーがすごい! 2018 年版, 2017 年版 国内編 [8] [9]
実用書ランキング	「2017 年 年間ベストセラー」 トーハン調べ【単行本 実用書】 [10]
実用書料理レシピ	料理レシピ本大賞 in Japan 2014, 2015, 2016, 2017 料理部門 大賞, 準大賞, 入賞 [11]

表 3 各セット中の書籍に投稿された最大・最小・平均レビュー数

	最大レビュー数	最小レビュー数	平均レビュー数
小説ランキング	7350	1685	3394.3
小説ミステリ	3423	118	1210.6
実用書ランキング	351	18	130.9
実用書料理レシピ	176	41	105.1

表 4 セットごとの書籍特徴ベクトルの要素ごとの標準偏差の平均

セット	要素ごとの標準偏差
小説ランキング	0.040
小説ミステリ	0.038
実用書ランキング	0.013
実用書料理レシピ	0.009

節で提示した 4 種類の提案手法に加え, 比較手法として書棚に整列された書籍の集合平均特徴ベクトルを作成し, \cos 尺度による類似度を算出し推薦書籍を決定する手法を用意した. 集合平均を用いるため, この比較手法では順位情報が用いられない.

4.1 実験方法

被験者に提示する 10 冊の書籍群は 4 セット用意し, 各セットに入る書籍は表 2 のようにした. 小説, 実用書それぞれで, ランキングと同一ジャンル書籍の 2 セットを用意した. 選択基準の中から, 同一著者作品とシリーズ作品の 1 巻以外を除いた 10 冊を 1 セットとする.

実験に使用するレビューには, 読書メーター^(注1)から収集した 18,805,119 件のレビューを用いる. 各セットごとの書籍へ投稿されたレビューの最大, 最小, 平均は表 3.5 のようになっている. 1 冊ごとの平均レビュー数は, 小説の方が圧倒的に多くなっている. また, トピックモデルは小説, 実用書それぞれ別に作成した.

クラウドソーシングサービスであるクラウドワークス^(注2)を利用し, 40 名の被験者から 1 人 1 件, 4 セットの書籍群から各 10 件ずつ仮想書籍整列データを用意した. 仮想書籍整列データは, 以下のように作成した. 被験者は, まず提示された 10 冊の書籍の中から, 読みたいと思った書籍 5 冊を選択する. 次に, 選択した 5 冊の書籍を, 被験者自身の書棚の同じ列に入れると仮定し並び替えを行う. この際, 被験者は整列理由を記述する. この書籍 5 冊の整列データを, 仮想書籍整列データとする. 得られた仮想書籍整列データに対し, 並び替えた 5 冊から 1 冊を抜き出す. その後, 仮想書籍整列データ作成の際に用いた 10 冊から被験者に選ばれなかった 5 冊と抜き出した 1 冊による推薦ランキングを作成し, 評価を行なった.

4.2 結果と考察

実験結果は, 表 5 のようになった. 各セットでの各手法で作成した推薦ランキングに対し, 元の並びから抜き出した書籍を正解とした平均逆順位によって評価した.

まず, 小説ランキングセットでは他のセットに比べ, 線形上位を除いた手法全てにおいて精度が低い. これは, 小説ランキ

ングセットで用いた書籍の特徴ベクトルで, 共通する要素に意味のある値が少なくなってしまったためと考えられる. LDA による各トピックへの所属確率は, 全体のトピック数に比べ少数トピックに対し高い所属確率を示し, その他のトピックへの所属確率は非常に低くなる. よって, 意味のある高い値を示す共通要素が存在しなければ, 予測は非常に難しくなってしまうと考えられる. 表 4 は, 各セットで用いた書籍の特徴ベクトルの要素ごとの標準偏差の平均値である. 標準偏差が最も高くなったのは, 小説ランキングである. 次いで小説ミステリ, 実用書レシピ, 実用書ランキングの順に高い値となった. 標準偏差が高いほど書籍ごとに所属確率の高いトピックは異なり, 滑らかな要素の遷移は難しくなる. それに対し, 標準偏差の低い実用書では, 共通する要素に意味のある比較的高い値が出現しているため, 整列順序から比較的精度の高い特徴ベクトルを予測できたと考えられる. この問題を解決するためには, 特徴ベクトル作成方法を変更する必要がある.

また, 線形回帰において決定係数の高くなったトピックの例は, 表 6 のようになった. 特徴ベクトルは L2 正規化を行っており, L2 正規化を行なった際のベクトル要素の最大値は 1 である. 表中には 0.01 前後の値が多いが, これは特徴ベクトルの要素値の中でも非常に小さな値である.

表 8 は, 作成した推薦ランキングにおいて決定係数が高かった上位 5 要素のうち, 回帰式作成に用いた整列書籍の対応する要素に, 1 冊でも 0.1 を超える値が存在した個数である. 最大は小説ランキングの 71 個だが, 各セット決定係数は 250 個存在するため非常に少ないといえる. このようになってしまった原因として, 極端に偏った特徴ベクトルの特性があげられる. 今回作成した特徴ベクトルは, 要素の大半が極端に小さな値になり, 一部の要素のみが高い値を示す. そのためユーザの整列順序意図に沿ったトピックであっても, 所属確率の高い書籍が含まれる場合には, 値はなだらかな変化をせずに極端な変化をしてしまう. これにより決定係数は下がってしまったと考えられる.

書籍特徴として比較的薄い部分であっても, ユーザによっては書籍の主題ではない一部の要素にのみ注目して書籍を整列しているという例も考えられる. よって, 所属確率の低い値の決

(注1) : <https://bookmeter.com/>

(注2) : <https://crowdworks.jp>

表 5 各セットでの各手法の実験結果

	小説ランキング	小説ミステリ	実用書ランキング	実用書料理レシピ
線形重み付き	0.37	0.49	0.41	0.49
線形上位	0.41	0.35	0.41	0.46
線形	0.36	0.50	0.48	0.46
カーネルリッジ	0.37	0.53	0.47	0.40
集合平均	0.31	0.54	0.43	0.42

表 6 セットごとの決定係数が高かった要素の例

	決定係数	傾き	書籍 A	書籍 B	書籍 C	書籍 D
小説ランキング	0.9628	0.0038	0.0030	0.0046	0.0100	0.0140
小説ミステリ	0.9457	0.0033	0.0018	0.0032	0.0066	0.0136
実用書ランキング	0.9944	-0.0003	0.0024	0.0022	0.0019	0.0015
実用書料理レシピ	0.9970	0.0015	0.0021	0.0036	0.0050	0.0068

表 7 一部整列理由を省いた場合の実験結果

	小説ランキング	小説ミステリ	実用書ランキング	実用書料理レシピ
線形重み付き	0.37	0.46	0.41	0.51
線形上位	0.42	0.36	0.41	0.46
線形	0.34	0.46	0.48	0.46
カーネルリッジ	0.36	0.50	0.47	0.39
集合平均	0.30	0.52	0.43	0.42

表 8 各セットにおける推薦において決定係数が高い 5 要素の中で整列書籍中に 0.1 を超える値があった件数

	要素数
小説ランキング	71
小説ミステリ	41
実用書ランキング	14
実用書料理レシピ	19

定係数が高くなってしまふことに必ずしも問題はないと考える。しかし、本論文で作成した特徴ベクトルでは、極端に所属確率の低いトピックに関してはほとんど意味がないほど小さな値となっている。そのため、書籍の主題ではない一部の要素によって整列された意図を汲み取ったとは考えづらい。極度に偏ったような特徴ベクトルを用いる場合、決定係数以外にも回帰式の傾きや、特徴ベクトル中の値の大きさを考慮する必要があると考えられる。

また、提案手法の精度が高くなった実用書のランキング、料理レシピの整列理由には以下のようなものがあつた。括弧内は、その被験者の平均逆順位である。

- 右から左に、自分が現在興味を持っている順に並べました。F と J はいずれもヨガ関係ですが、ずぼらさ＝負担が少ない順を右から左に加えました。(線形重み付き：0.73)
 - 役に立つと思う順に左から右に並べました。主菜を作り置きできる本を左から並べ、副菜であるサラダの作り置き本、最後に作り置き本ではない本にしました。(線形重み付き：0.67)
- 負担が少ない順という明確な難易度の推移や、主菜から副菜の作り置きのように共通する条件を持ちながら要素が推移する場合に、提案手法の精度が高くなると考えられる。しかし、明確な難易度の推移であっても精度が低下する場合が存在した。健

康に関する書籍を難易度順に整列したとき、ヨガに関する書籍の間に異なる健康法の書籍を入れるような、同一の主題を持つ書籍の間に異なる主題を持つ書籍を入れた場合に精度は低下した。原因としては、あくまで書籍の持つ最も強い特徴に強く影響されており、同一の特徴が最も強い特徴であった場合に初めて細かい特徴の変化を考慮したということが考えられる。書籍の持つ細かな特徴に基づき整列される場合に適切な推薦をするためには、より重みの付け方を工夫する必要があると考える。

また、整列理由として著者の五十音順や判型の大きさ順という理由が見られた。五十音順や書籍の大きさは、提案手法の特徴ベクトル作成方法では、取り出すことのできない情報であると考えられる。上記のような整列理由は実用書ランキングを除くセット 30 件中 5 件存在し、表 7 はこの 5 件を除いた場合の実験結果である。なお、取り出すことのできない整列理由で並べられた仮想書籍整列データを省いても、提案手法の精度が著しく向上することはない。

比較手法である集合平均に対し、一部勝る箇所は存在するものの、同じ手法であつて他のセットでは下回る箇所も存在している。よって、提案手法が比較手法に対し確実に勝っているということとはできない。しかし、極度に偏った特徴ベクトルながら同一トピックへの所属確率の高かつた実用書では、料理レシピのカーネルリッジを除き集合平均よりも高い精度を出すことができた。このことから、適切な特徴ベクトルを作成することができた場合、提案手法によりユーザの意図する整列順序に沿った書籍を推薦することができる可能性がある。

5. ま と め

本論文では、ユーザの書籍整列順序に隠された整列意図を汲

み取することで、整列順序により適した書籍の推薦を行う手法の提案を行なった。整列順序に合わせた推薦を行うことで、ユーザの興味に合致し、かつユーザにとって受け入れやすい書籍の推薦を行うことを目的としている。

提案手法では、書籍レビューからトピックモデルを作成し、レビューごとのトピック所属確率に対し合算と正規化を行うことで書籍特徴ベクトルを作成した。書籍特徴ベクトルを整列順序に合わせて並べ、要素ごとに回帰分析を行うことで、ある位置に入る書籍の特徴ベクトルを予測し、予測特徴ベクトルとの類似度を算出することで書籍推薦を行なった。提案手法の評価実験として、クラウドソーシングを用いて仮想書籍整列データを作成し、1種類の比較手法と4種類の提案手法の比較実験を行なった。

これらの結果、提案手法の有効性を確認することはできなかった。しかし、一部の理想的な特徴ベクトルを用いた場合、ユーザの意図する整列順序に沿った書籍を推薦できる可能性があることがわかった。

今後の課題としては、まず、特徴ベクトルの作成方法の改善があげられる。現在のLDAを用いたトピックモデルによる特徴ベクトルでは、極度に偏った特徴ベクトルが作成されてしまい、本手法における予測には不向きであるといえる。適切な特徴ベクトルを作成することで、より精度の高い予測が可能であると考えられる。

一方、今回の評価実験では、予測精度に関する評価実験のみを行なった。被験者を用い仮想書籍整列データを作成したが、あくまで書籍セットを指定した仮想的な整列データであり、かつ実際にユーザが本手法によって書籍を推薦された場合の精度に関しては評価していない。今後、ユーザが自由に書籍を指定し作成した整列順序データから書籍を推薦し、ユーザが推薦された書籍を気に入るかについて調査することを計画している。

謝辞 本研究の一部は、平成30年度科研費基盤研究(C)(課題番号:18K11551)によるものです。ここに記して謝意を表すものとします。

文 献

- [1] 林 利憲, 王 元元, 河合 由起子, 角谷 和俊 “E-Commerceにおける不満情報とレビューに基づいた不満解決商品推薦手法の提案”, 研究報告情報基礎とアクセス技術 (IFAT), Vol. 2017-IFAT-128, No. 23, pp.1-6, 2017.
- [2] 吉田 朋史, 北山 大輔, 中島 伸介, 角谷 和俊 “ユーザレビューの分散表現を用いた主観的特徴の意味演算による観光スポット検索システム”, DEIM Forum 2017, P6-5, 2017 年.
- [3] 加藤 嘉浩, 石井 隆稔, 宮澤 芳光, 植野 真臣 “Latent Dirichlet Allocation を用いたレポート推薦システム”, 電子情報通信学会論文誌 D, Vol. J99-D, No. 2, pp.152-164, 2016.
- [4] 奥 健太, 服部 文夫 “セレンディビティ指向情報推薦のためのフュージョンベース推薦システム”, 知能と情報, Vol. 25, No. 1, pp.524-539, 2013.
- [5] 池田 翔武, 川越 恭二, 奥 健太 “楽曲遷移の滑らかさを考慮したプレイリスト推薦”, 情報処理学会論文誌データベース (TOD), Vol. 9, No. 4, pp.1-5, 2016.
- [6] honto “honto - 発表!! 2017 年年間ランキング【小説】:電子書籍”, <https://honto.jp/cp/hybrid/recent/ranking/ebook/04.html>
- [7] honto “honto - 発表!! 2017 年年間ランキング【小説 11~50 位】:電子書籍”, <https://honto.jp/cp/hybrid/recent/ranking/ebook/04/11-50.html>

- [8] 『このミステリーがすごい!』編集部編 “このミステリーがすごい! 2018 年版”, 宝島社, 2017
- [9] 『このミステリーがすごい!』編集部編 “このミステリーがすごい! 2017 年版”, 宝島社, 2016
- [10] TOHAN “トーハン調べ「2017 年 年間ベストセラー」”, http://www.tohan.jp/bestsellers/upload_pdf_past/171201_bestseller_2017y.pdf.pdf
- [11] 料理レシピ本大賞 in Japan 実行委員会 “料理レシピ本大賞 in Japan”, <http://recipe-bon.jp/>