

# 評価値予測タスクを用いたユーザの嗜好に適合したレビューの抽出

坂本 新真<sup>†</sup> 牛尼 剛聡<sup>††</sup>

<sup>†</sup>九州大学 芸術工学部 〒 815-8540 福岡県福岡市南区塩原 4-9-1

<sup>††</sup>九州大学大学院 芸術工学研究院 〒 815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: <sup>†</sup>sakamoto.shinma.626@s.kyushu-u.ac.jp, <sup>††</sup>ushiana@design.kyushu-u.ac.jp

**あらまし** E コマースサイトにおいて、レビューはユーザにとって商品の理解や購入の判断を行うための重要な情報である。掲載されているレビュー数が膨大である場合、それらの中からユーザにとって有用なレビューを推薦することができれば、ユーザの効率的な商品選別を支援できる。しかし、ユーザの嗜好には多様性があり、ユーザの嗜好を考慮してユーザにとって価値の高いレビューを抽出することは困難である。本論文では、ユーザの嗜好に適合したレビューを抽出する手法を提案する。具体的には、レビューを用いてユーザのアイテムに対する評価値を予測する機械学習モデルを訓練させることで、レビュー中に含まれるユーザの嗜好を学習する。そして評価値を正確に予測できるレビューをユーザの嗜好に適合したレビューとして抽出する。評価実験によって提案手法は比較手法と比べてユーザにとって有用なレビューを抽出できることが分かった。

**キーワード** 説明可能性, レビュー, 機械学習, 推薦

## 1 はじめに

Amazon や楽天市場などに代表される E コマースの発展と普及に伴い、商品に対するユーザのレビューは急速に増大している。これらのレビューはユーザにとって、商品の理解や購入の判断を行うための重要な情報である。しかし、商品によっては数百から数千のレビューが掲載されていることも少なくない。ユーザがこのような膨大なレビューの全てに目を通すことは困難であり、ユーザは自分にとって必要な情報を見つけられないことも多い。こうした中で、膨大なレビューの中からユーザに役立つレビューを推薦する研究が行われている。例えば、Paulら [1] は、ニューラルネットワークによって、レビューの有用性を予測することでユーザにレビューを推薦する手法を提案した。

また近年、Amazon の「あなたへのおすすめ」のように、膨大なコンテンツの中からユーザの好みに合うようなアイテムを提示することで、ユーザの意思決定を支援する推薦システムの需要が高まっている。代表的な推薦システムの一つとして、ユーザに過去に購入したアイテムをどれほど好むのかを表す 5 段階の評価値をつけてもらい、そのデータを用いてユーザの好みの分からないアイテムの評価値を推定する手法がある。Chenら [2] は、5 段階の評価値データを正解データとして、ユーザのレビューからニューラルネットワークによって潜在表現を生成し、対象ユーザの未知のアイテムへの評価値を予測する手法を提案している。

深層ニューラルネットワークをはじめとした機械学習技術の発達によって、高精度の推薦モデルが現れる一方で、これらのほとんどのモデルは「なぜこのユーザにこのアイテムを推薦したのか」という推薦理由を提示することができないという問題がある。ユーザに対して精度の良い推薦が行われた場合でも、推薦理由を知ることができなければ、システムへの不信感など

からユーザはその推薦結果を受け入れることができない可能性がある。システムの透明性や説得力を向上させるために、このような解釈性の低いモデルに対して、推薦理由を提示することが求められている [3]。

推薦理由の提示の例として、図 1 のように Netflix のホーム画面では過去に視聴した作品を元にした推薦であることを提示している。このような推薦理由の説明によって、ユーザは推薦に対して納得感を得ることができ、推薦の効果が大きくなる可能性がある。しかし、前述のように解釈性の低いモデルに関しては推薦理由を提示することは困難である。さらに、推薦システムの透明性の向上のためには簡潔な説明文ではなく、より具体的な内容を含む必要がある [4]。

本論文では、推薦理由の提示のためにアイテムのレビューに着目する。レビューは、アイテムに関する情報やレビューを記述した人の嗜好や特性に関する情報を含んでいる。ユーザにアイテムが推薦された際に、ユーザと同様の嗜好を持つ記述者のレビューや推薦の理由となったアイテムの情報を含むレビューを、ユーザに提示することができれば、ユーザはアイテムの推薦に対して、納得感を得ることができ、推薦の効果の増大が期待できる。また、レビュー数が多く、ユーザが役立つレビューを見つけ出すことができないという問題に対しても対処できると考えられる。

本論文では、機械学習モデルを用いてユーザの嗜好に適合し



図 1: Netflix の推薦理由の提示の例

たレビューを抽出する手法を提案する。図2に、提案するシステムの全体像を示す。本システムでは、はじめにアイテムのレビューからユーザーのアイテムに対する評価値を予測する機械学習モデル訓練させる。訓練の際、機械学習モデルはレビューに含まれるユーザーの嗜好を予測の手がかりとして学習していると考えられる。訓練によって評価値を正しく予測できるレビュー、つまり正解データと予測の誤差が小さいレビューは、ユーザーの嗜好に適合するレビューであり、反対に予測の誤差が大きいレビューは、ユーザーの嗜好に適合しないレビューや、「すぐに届きました。」などの具体的な内容のないレビューであると仮定する。この仮定に基づいて図2の下のようにレビューをランク付けし、ランキングの上位を抽出する。抽出したレビューがユーザーにとって役に立つレビューであるかを評価実験によって検証する。

本論文の提案手法によって、ユーザーの嗜好に適合したレビューが抽出される場合、図3のように、推薦モデルが予測した評価値を正解データとして、提案手法の予測モデルを訓練させることで、推薦モデルがユーザーにアイテムを推薦する理由となったユーザーの嗜好を含むレビューを抽出することができる。と考える。

本論文の構成は以下の通りである。第2章で関連研究について述べる。第3章で提案手法の詳細について述べる。第4章では評価実験と結果について述べる。第5章で結果についての考察を述べ、最後に第6章でまとめと今後の課題について述べる。

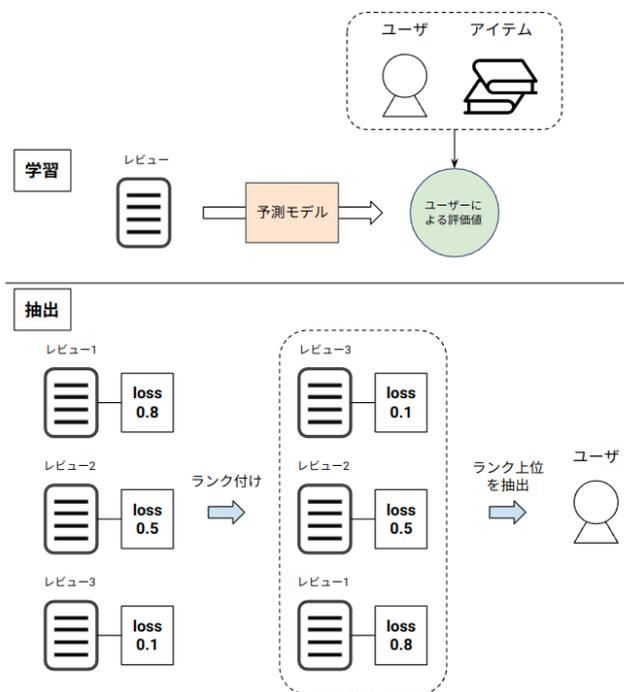


図2: 提案手法の概念図

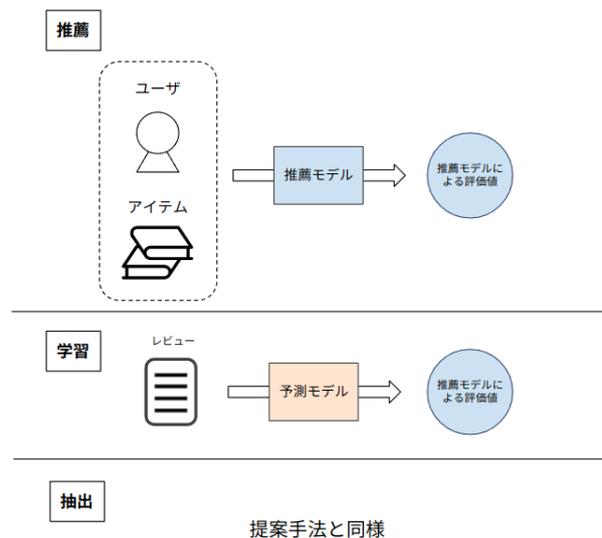


図3: 提案手法を利用して、将来的に実現が期待できるシステム

## 2 関連研究

### 2.1 推薦理由の生成に関する研究

説明可能な推薦システムのアプローチは大きく分けて model-intrinsic approach と model-agnostic approach に分類される [3]. 以下では、それぞれの手法について詳細な説明をする。

#### 2.1.1 model-intrinsic approach

model-intrinsic approach は、推薦モデルに解釈可能なアルゴリズムを採用することで、推薦モデルから意思決定の理由を直接推定する手法である [3]. この手法によって生成される推薦理由は、推薦モデルの結果に対して正確である一方で、説明可能性を担保しながら推薦の精度を維持する必要がある。

Abdollahi ら [5] は、行列分解を用いた推薦手法に対して説明可能性を定式化し、説明可能でない結果に対してペナルティを与えることで説明可能な推薦を実現する手法を提案している。しかし、推薦モデルに対して説明可能性に関する制約を加えているため、推薦の精度が下がってしまう場合がある。

Chen ら [6] は、評価値予測に用いたアテンション機構の重みが大いほど商品の情報を詳細に含む有用なレビューであると予測する手法を提案している。しかし、この手法におけるアテンションの重みはレビューの有用性を表すものであり、ユーザーの嗜好や特性を考慮していないため、推薦理由として提示するには不十分である。

#### 2.1.2 model-agnostic approach

model-agnostic approach は、まず推薦モデルを学習させ、次に推薦理由を説明するために解釈可能なモデルを推薦モデルの入力と出力から別途学習させる手法である [3]. 解釈を行うモデルと推薦モデルは独立しているため、推薦モデルの結果を正確に説明できる保証はない。しかし、推薦モデルに依存せず汎用性が高いという利点がある。

Nóbrega ら [7] は、任意の機械学習モデルに対して入力の特

微量の重要度を求めることができる LIME(Local Interpretable Model-agnostic Explanations) [8] を、推薦システムに利用することで推薦理由を説明する Lime-RS を提案した。

また, Morisawa ら [9] は, Lime-RS において特徴量が増えると解釈モデルの学習が困難になるという問題に対して, 特徴量の数をチューニングすることによって解決し, ユーザに提示する推薦理由の生成方法が考慮されていないという問題に対して, テンプレートを用いた説明を生成する手法を提案した。

しかし, テンプレートを用いた説明の生成手法は, 説明が単純であることや同じ説明が繰り返されることから, 説得力に欠けるという欠点がある。本研究ではレビューを用いて推薦モデルと独立なモデルを構築し, 汎用的な手法でユーザに対して柔軟に説明を提示することを目標とする。

## 2.2 レビュー推薦に関する研究

レビュー推薦とは, ユーザに対して購買決定に役立つレビューを提示する処理である。

Lappas ら [10] はレビューの有用性を評価し, 与えられたコーパスの中からレビュー全体を代表するようなレビューのサブセットを抽出する手法を提案している。

Yang ら [11] は, レビューテキストの構造的特徴と意味的特徴を用いることで, レビューの有用性を予測する手法を提案している。

これらの手法はユーザの嗜好を考慮していないため, すべてのユーザに対して一様な基準でレビュー推薦を行う。しかし, レビューの質や有用性はユーザによって異なるため, 本研究ではユーザの推薦理由に基づいてレビュー推薦を行うことで, パersonalized レビュー推薦を実現することを目指す。

## 2.3 レビューを用いた評価値予測に関する研究

ユーザのアイテムに対する評価値を予測する推薦モデルはこれまで数多く研究されている。本節ではレビューを用いた研究について説明する。

McAuley ら [12] は, レビューテキストを組み込んだトピックモデルによってユーザとアイテムの潜在表現を生成することで, 評価値予測の精度を向上させる手法を提案した。

Chen ら [2] は, レビューのテキストから, ユーザの行動やアイテムの特性を学習し評価値を予測するニューラルネットワークを用いた手法である Deep Co-operative Neural Network(DeepCoNN) を提案した。また, [6] ではアテンション機構を用いてレビューの有用性を学習することで DeepCoNN の精度を上回る手法である Neural Attentive Rating Regression(NARRE) を提案した。

これらの研究から, レビューのコンテキスト情報がユーザのアイテムに対する評価値を予測する際に非常に有用であることがわかる。

## 3 提案手法

### 3.1 提案手法の概要

本論文では, 対象とするアイテムに対するレビューの中から,

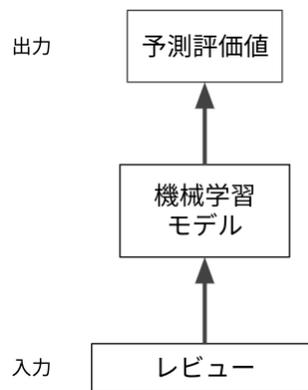


図 4: 機械学習モデルによる評価値予測の概念図

機械学習を用いてユーザの嗜好に適合したレビューを抽出する手法を提案する。

はじめに, 対象ユーザの評価値が既知のすべてのアイテムのレビューを入力として, それぞれのアイテムに対するユーザの評価値を予測する機械学習モデルを訓練させる。図 4 は, 提案する機械学習モデルの概念図である。

この機械学習モデルは, 入力するレビューによってうまく予測ができる場合と, そうでない場合が存在する。表 1 は, 小説「同志少女よ, 敵を撃て」<sup>1</sup>に関するレビューの例である。レビュー 1 は「戦争」「エンターテインメント性」などの小説のテーマや性質に関する内容や, 「イリーナ隊長」などの小説に関する具体的な語を含んでいる。それに対して, レビュー 2 は小説に関する内容を含んでおらず, どの本にも出現する可能性のあるレビューである。それぞれのレビューを図 4 の機械学習モデルに入力した場合, レビュー 1 は文中の内容を, 予測の手がかりとなるユーザの嗜好として学習することができるが, レビュー 2 はユーザの嗜好を学習することができない。

本論文では上記の考えに基づき, 評価値予測モデルを訓練させることでユーザの嗜好を学習させたのち, 正解データとの誤差が小さいものを, ユーザの嗜好に適合したレビューとして抽出する。

表 1: 「同志少女よ, 敵を撃て」のレビュー例

レビュー 1	戦争という難しいテーマをととても読みやすく, またエンターテインメント性を大事にした, 小説である。個人的にはイリーナ隊長がかっこよくて, たまに見せる優しさでホッコリする。
レビュー 2	全く本に興味を持たなかった中学生の娘が初めて買って欲しいと言ってきた本です。夜寝る前に読んでいるようで今半分くらいはよんだようです。

### 3.2 機械学習による評価値の予測

提案手法における, 機械学習を用いた評価値の予測について説明する。図 5 は機械学習モデルによる評価値予測の全体像である。以下に機械学習を行う際の詳細な説明をする。

1: 逢坂冬馬『同志少女よ, 敵を撃て』(早川書房, 2021)

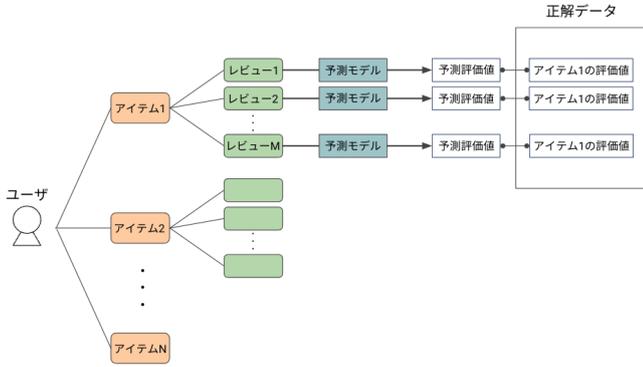


図 5: 機械学習モデルによる評価値予測の全体像

いま, ユーザ集合  $U = \{u_1, u_2, \dots, u_{|U|}\}$  と, アイテム集合  $I = \{i_1, i_2, \dots, i_{|I|}\}$ , レビュー集合  $R = \{r_1, r_2, \dots, r_{|R|}\}$  が与えられているものとする. ユーザ  $u \in U$  が評価したアイテムの集合を  $I(u) \subseteq I$  とする. またアイテム  $i$  に対するレビュー集合を  $review(i)$  と表す. ここで,  $review(i) \subseteq R$  であり,  $i_n \neq i_m$  のとき,  $review(i_n) \cap review(i_m) = \emptyset$  である. また, ユーザ  $u \in U$  のアイテム  $i \in I$  に対する評価 (スコア) を  $score(i, u)$  と表記する.

入力  $x$  に対する正解ラベル  $y$  を  $(x, y)$  と表記するとき, 訓練データ集合を以下の式 (1) ように定義する.

$$D_u = \{(x, y) | i \in I(u), x \in review(i), y = score(i, u)\} \quad (1)$$

以上のように定義された教師データを用いて機械学習による評価値予測を行う.

本手法では, 機械学習モデルとして Devlin ら [13] によって提案された BERT (Bidirectional Encoder Representations from Transformers) を用いた回帰モデルを利用する. 図 6 は, 本研究で利用するレビューを入力とした BERT モデルの構造である. 最終層の Transformer Encoder から単語ごとに分割した文と, 入力の際に挿入した特殊トークンに対応する分散表現を得る. 特殊トークンである CLS トークンに対する分散表現は, 文章全体を表す分散表現として利用することができるため, これを全結合層に入力し最終的な予測評価値を得る. 損失関数には, 平均二乗誤差を用いて実際の評価値と予測評価値の差を求める.  $x \in review(i), i \in I(u)$  のときの予測評価値を  $prediction(x, u)$ , バッチサイズを  $N$  とすると平均二乗誤差は以下の式 (2) で表される.

$$L = \frac{1}{N} \sum_{n=1}^N (prediction(x, u) - score(i, u))^2 \quad (2)$$

また BERT モデルについては, 事前学習済み BERT モデルを fine-tuning することで学習を行う. 事前学習済み BERT モデルには, 東北大学乾研究室による日本語 Wikipedia 学習済みモデル [14] を用いる.

### 3.3 ユーザの嗜好を表す文の抽出

3.2 節で訓練した機械学習モデルを用いて, ユーザ  $u \in U$  の嗜好に適合するレビューを抽出することを考える.

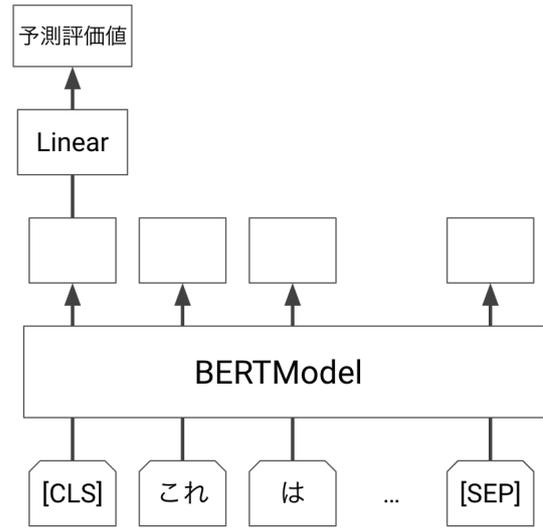


図 6: BERT による評価値予測の構造

あるレビュー  $x \in review(i), i \in I(u)$  を入力したとき, モデルの出力  $prediction(x, u)$  がユーザのアイテムに対する評価値  $score(i, u)$  を正しく予測することができた場合, 入力したレビューはユーザの嗜好に適合していると仮定する. 一方, 「商品届いています. ありがとうございます。」などの内容とは関係のないレビューや, ユーザの嗜好に適合しないレビューは評価値を正しく予測することができない.

この仮定に基づいて, ユーザが評価した各アイテムのすべてのレビューに対して, 機械学習モデルの出力  $prediction(x, u)$  を得た後, 式 (3.3) が小さいレビュー順に並び替え, 上位を抽出しユーザに提示する.

$$L(x) = (prediction(x, u) - score(i, u))^2 \quad (3)$$

## 4 評価実験

評価実験では, 提案手法によって抽出されたレビューがユーザの嗜好に適合しているかを検証するための実験を行った.

### 4.1 被験者アンケートによる訓練用データセットの作成

#### 4.1.1 アンケートによるデータの取得

本論文では, 被験者によるアンケートを行い, アイテムとアイテムに対する評価値についてのデータを取得した. アンケート対象のアイテムは漫画, 雑誌, 学習参考書を除く書籍を対象とした. アンケートは, 「本のタイトル」と「あなたの本に対する総合評価」について 1 から 5 の 5 段階評価で答えてもらった. また, 書籍の同一性を担保するために, 「本の ISBN もしくは Amazon リンク」についても回答してもらった. アンケートの回答数は 1 人あたり 25 以上 30 以下とした. 次に, 回答してもらった書籍についてのレビューを Amazon<sup>2</sup> から取得した. アンケートの被験者数は 7 人で, 各ユーザに対して取得した総レビュー数は表 2 のようになった.

2: <https://www.amazon.co.jp/>

表 2: アンケート結果

被験者	評価された本の数	総レビュー数
1	26	9519
2	25	9910
3	30	5883
4	25	4999
5	27	7945
6	25	1889
7	25	3192

#### 4.1.2 データの前処理

4.1.1 節で抽出したレビューを用いて訓練用のデータセットを作成した。抽出したレビューは、BERT の入力長に合わせて単語数が 512 以下となるようにした。また、図 7 のように url 等のテキスト中のノイズは除去し、sentencepiece [15] を用いて単語に分割をした。単語の正規化のために、数字はすべて 0 に置き換え、アルファベットはすべて小文字で統一し、半角文字はすべて全角文字に変換した。さらに、「最高!」や「よかったです。」など明らかに具体的な内容を含んでいないレビューを訓練データにしないために、単語数が 20 以下のレビューを除外した。

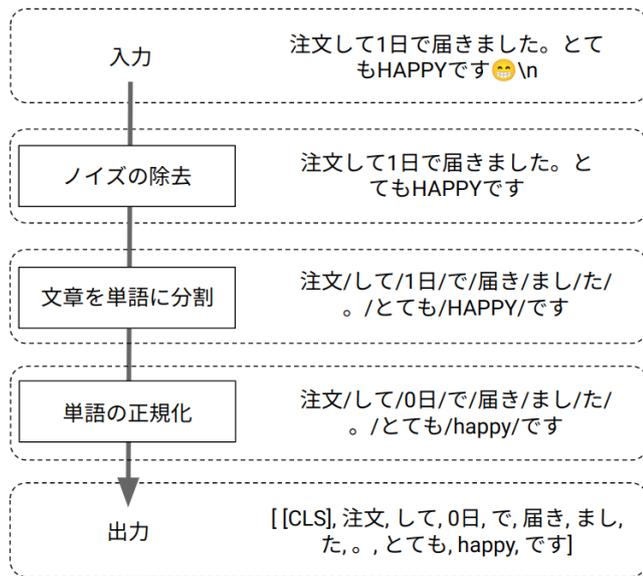


図 7: レビューの前処理

#### 4.2 比較手法

提案手法の有効性を確かめるために以下の 2 つの比較手法を用いた。

- random: ユーザーが評価した各アイテムにつき 5 件をランダムに抽出し、ユーザーに提示する。
- tf-idf: tf-idf 法は、単語の重要度を測る手法である。tf(term frequency, 単語の出現頻度) は各文書においてその単語がどのくらい出現したのかを表す。idf(inverse document frequency, 逆文書頻度) は単語がいくつかの文章内で共通して使われて

いるかを表す。  $x \in review(i), i \in I(u)$  に含まれる単語のうち、名詞と形容詞に属する単語の集合を  $W(x) = \{w_1, w_2, \dots, w_{|W(x)|}\}$  とするとき、tf-idf 法に基づくレビュー  $x$  のスコア  $L_{tfidf}(x)$  を式 (4.1) に定義する。

$$L_{tfidf}(x) = \frac{\sum_{w \in W(x)} tfidf(w)}{|W(x)|} \quad (4)$$

各アイテムのすべてのレビューに対して  $L_{tfidf}(x)$  を算出し、最もスコアの高かった上位 5 件をユーザーに提示する。

#### 4.3 実験方法

各ユーザーに対して 4.1 節のデータセットを利用して評価値予測モデルの訓練を行った。

本論文の目的は、未知のアイテムのレビューに対する評価値を予測するのではなく、ある商品のレビュー集合からユーザーの嗜好に適合したレビューを抽出することである。そのため、アンケートによって評価されたアイテムのレビューはすべて訓練データとして利用するが、対象レビューに対してスコアを求める際はそのレビューを訓練データとして利用しないために、交差検証によってテストデータとして利用された際の予測値を利用した。

交差検証には層化 K 分割交差検証を用いた。層化 K 分割交差検証は図 8 のように、データを K 個に等しく分割する際に、正解ラベル (今回はユーザーのアイテムに対する評価値) の分布が、訓練データとテストデータでできるだけ均等になるようにする手法である。式 (3) の  $L(x)$  を求める際はレビュー  $x$  がテストデータに含まれるときの予測値を利用した。

提案手法によって抽出された上位 5 件のレビューと、比較手法によって抽出されたレビューについてアンケートを用いて評価実験を行った。評価の際の質問項目は以下の 3 つである。回答は 5 段階のリッカート尺度 (5. そう思う, 4. ややそう思う, 3. どちらとも言えない, 2. あまりそう思わない, 1. そう思わない) を用いた。

質問 1: この文章は本の内容に関する特徴を含んでいる

質問 2: この文章の内容に共感した

質問 3: この本をまだ読んでいない他の人が読もうとするとき、この文章は役に立つ

その後、提案手法と比較手法で質問の回答に有意差がみられるかの検定を行った。有意差を求める際は、Welch の t 検定を用いて、p 値を算出した。

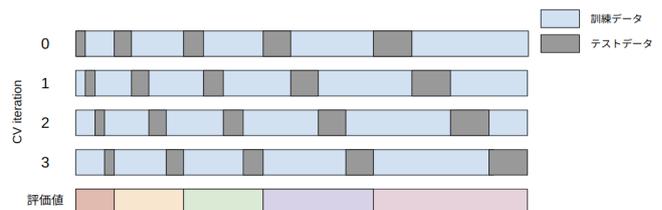


図 8: 層化 K 分割交差検証

#### 4.4 実験結果

図9は、各質問に対する各評価値の割合である。また図10は、各質問に対する評価値の平均値、及び提案手法と各比較手法の Welch の t 検定による有意差の有無を表すグラフである。質問ごとに p 値を求めたところ、random については質問1では有意水準 5% で有意であることが確かめられ、質問3では、有意水準 1% で有意であることが確かめられたが、質問2では p 値が 0.05 以上となり有意差は認められなかった。tf-idf についてはすべての質問において有意水準 1% で有意であることがわかった。

また、図11と図12は、同じ『永遠の0<sup>3</sup>』を評価した被験者に対して抽出されたレビューの例である。それぞれのレビューについて、入力する特徴量の予測への貢献度を求めることが可能な SHAP [16] を用いてレビュー中の単語の、予測に対する貢献度を求めた。

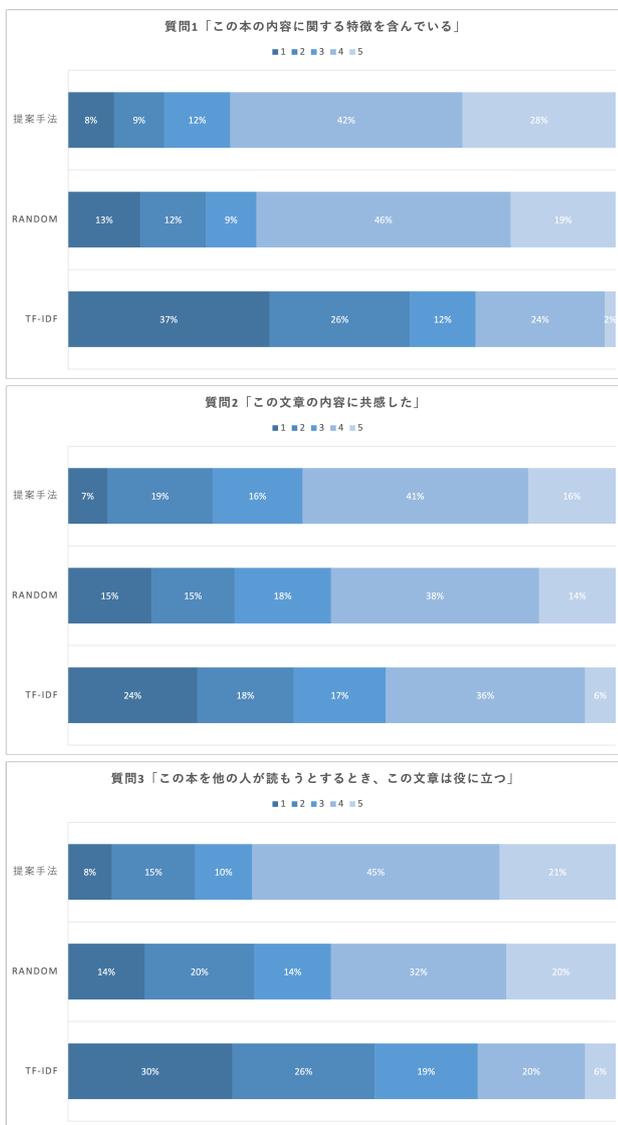


図9: 各質問の回答の割合

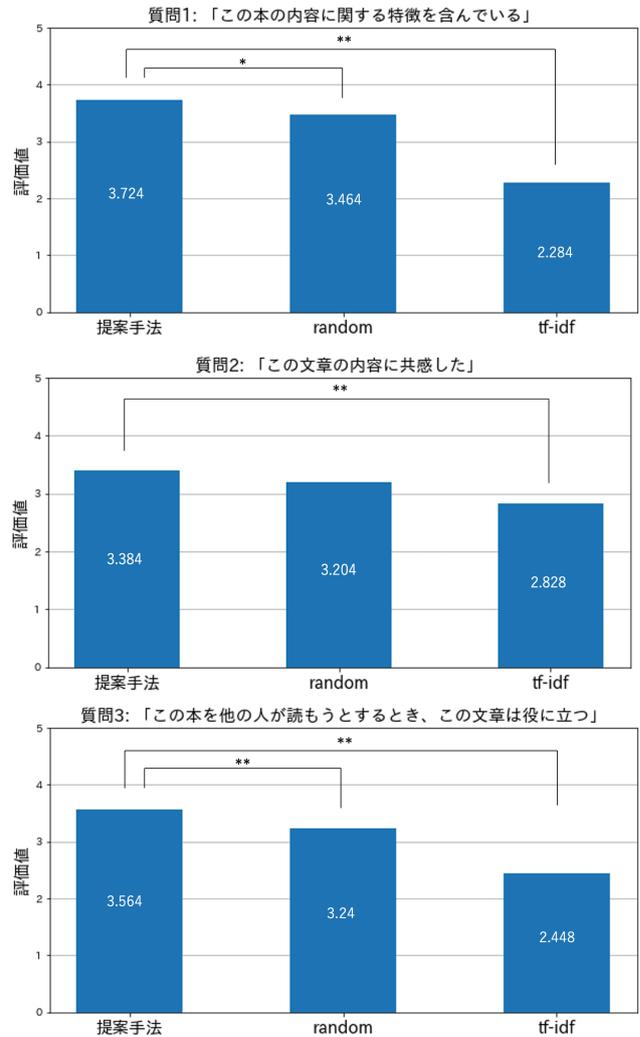


図10: 各質問の評価の平均値 (\* $p < 0.05$ , \*\* $p < 0.01$ )

## 5 考察

評価実験の結果について、図10より、提案手法が比較手法と比べて評価値の平均が比較的高いことが分かる。特に、質問1と質問3については t 検定により有意差があることが示された。

質問1の結果から、提案手法によって本に関する具体的な内容を含むレビューを抽出できることが分かった。また、表3のように、提案手法によってレビューをランク付けした際に、下位となったレビューは本についての具体的な内容が書かれていないレビューが多くあったことから、「どの本にも出現する可能性のあるレビューは、ユーザの評価値をうまく予測することができない」という仮定は概ね正しかったと考える。

また、図11と図12から、同じ本に対して異なる被験者にレビューを抽出をした場合、被験者1は「特攻」や「戦争」、作者の「百田」という単語を含むレビューを重視していることが shap 値から予測される。被験者2は被験者1と同様の「戦争」という単語を重視したレビューが抽出されているが、「泣ける」「感動」「涙」という単語を含むレビューが抽出されており、そのような単語を重視していることも shap 値から予測される。

3: 百田尚樹『永遠の0』(講談社文庫、2009)

分厚い本を読むのが苦手な私が、こんなにも集中して読んだ本は初めてなくらい、内容に引き込まれました。そして、じわじわと流れ出てくる涙を押さえるのが大変でした。。喫茶店や、電車で読んでいても、涙が止らず、溢れ出します。特攻隊の話というキーワードだけを聞いて、読み始めたのですが、まさかと思う展開ばかりが続き、読んでみた人にだけわかる、ドキドキ感やじわじわ感があります。友達にも勧めているほどです。

「売れてるけど、零戦とか特攻の話か～重いな～。でも、ナイトスクープの百田さんだしな～きつとおもしろいんだろうな～」と思いながら、ついに買いました。すぐに読み終わりました。壬生義士伝ですな。展開も読める。パカにしてんのかくらい朝日新聞な男が出てくるところなんて、読んでるこっちが恥ずかしい。。でも、エンタメとしてはイイ!だってナイトスクープの百田さんだもの!しかしながら、視聴者と読者って違うと思いますよ。

初めは、飛行機の話で??ばかりだったけど、後半...読むスピードがとまりませんでした。t\_t戦争を知らない若い世代の人に呼んで欲しい!!作品だと思います。最後は、びっくりの連続で、読んでよかった!!と気持ちが満たされていました。

図 11: 被験者 1 に対して抽出された「永遠の 0」のレビュー<sup>4</sup>

すぐに買って読みました。久々に泣ける本に出会いました。「生きる」言う事を真剣に見つめなおすチャンス頂きました。

感動でした。tvで知って購入したんですが、買って良かったです。

読み出したら止まりませんでした。しかも、途中からは涙も、そして鼻水も笑止まらなくなりティッシュ箱を手放すこともままならなくなりました。どうしてこう都合良く、時系列順に宮部の話が聞けたのか?現代を生きる主人公、姉のストーリーは出来過ぎではないのか?色々ごちなさを感じてしまう点はあったのですが、宮部を取り巻く壮絶なストーリーの迫力を前に、そんなものは吹っ飛んでしまいました。個人的な話なのですが・・・私は今まで過去の日本について肯定的にとらえることができませんでした。あのナチス・ドイツと手を組んで、カミカゼを信じて戦った、狂気としか思えない時代を。ですが、この本を読んで、あの時代も、私と同じように人が、愛や喜びや悲しみ、人間的な感情を持って生きていたこと、それを強く感じられた気がします。この本は現代に書かれたフィクションですが・・・この本を読んで、その時代を必死に生きてきた父や祖父に関して、肯定的な感情を持つだけではなく敬意を感じる事が出来るようになった気がします。こんな変化をもたらすことができるなんて、読書は素晴らしい体験だ!!!と思うと同時に、それを可能にさせる、この本の凄さ、素晴らしさを改めて感じます。しかし・・・やはり・・・こんな愚かな戦争は二度とするべきではない、二度と過ちを繰り返してはならない、そう改めて心に強く思いました。

図 12: 被験者 2 に対して抽出された「永遠の 0」のレビュー<sup>4</sup>

このことから、提案モデルが被験者の嗜好を学習し、それを評価予測に利用していることが推測される。

表 3: 提案手法の下位のレビューの例<sup>5</sup>

説明通りの商品で満足しております。また利用させていただきます。ありがとうございました。

よくインスタでこの本を載せて承認欲求を満たしてる人いますが、そういう子が大好きです

面白かった。カバーの絵も内容に合っているのだが、自分の好みでは無かったためなぜか手に取ることがなかった。でも気にはなっていた。audibleに出ている事を知り、聴いてみたらスピード感もあり、引き込まれた。ナレーションも上手く役を分けてくれているので理解しやすい。いや悲しくも面白い。

推薦システムの推薦理由の提示として、提案手法が有効であるかという点については、質問 3 の結果から未知のアイテムが推薦された場合に、ユーザにとって役に立つレビューを提示す

ることができる可能性があることが示された。しかし、質問 2 の結果は比較手法と比べて評価値の平均は高かったものの、有意差は認められなかったことから提案手法によって抽出されたレビューが、ユーザの嗜好に適合したレビューであるとは言えないため、推薦理由の提示としては不十分である可能性がある。

被験者に対して共感できるレビューを抽出できなかった理由として、肯定的なレビューと否定的なレビューを区別せずに利用していたことが挙げられる。実際に、いくつかの例では被験者が高評価をつけていた本に対して、否定的なレビューを抽出していたり、反対に低評価の本に対して、肯定的なレビューを抽出していることが分かった。提案手法に加えて、肯定的・否定的の判断をすることができれば、ユーザに対してより共感することができるレビューを抽出することができるのではないかと考える。

## 6 おわりに

本論文では、ユーザの嗜好に適合したレビューを提示するために、ユーザのアイテムに対する評価値をアイテムのレビューから予測する機械学習モデルを訓練させることで、ユーザの嗜好を学習し、予測値と正解データとの誤差の小さいレビューを

4: レビューの出典 (<https://www.amazon.co.jp/dp/406276413X>)

5: レビューの出典 (<https://www.amazon.co.jp/dp/4478025819>)

<https://www.amazon.co.jp/dp/4046047992>

<https://www.amazon.co.jp/dp/4152100648>)

抽出する手法を提案した。評価実験から、提案手法によって抽出されたレビューが、比較手法と比べてアイテムの特徴を捉えた役立つ文章であることが分かった。今後の課題として、ユーザの嗜好をより正確に捉えるために、肯定的・否定的などの感情に関する分析を用いることを検討している。また、実際に推薦システムの結果を用いて訓練することで、推薦理由の提示として有効であるかを検証する予定である。

## 謝 辞

本研究は JSPS 科研費 19H04219 の助成を受けたものです。

## 文 献

- [1] Debanjan Paul, Sudeshna Sarkar, Muthusamy Chelliah, Chetan Kalyan, and Prajit Prashant Sinai Nadkarni. Recommendation of high quality representative reviews in e-commerce. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, RecSys '17, p. 311–315, New York, NY, USA, 2017. Association for Computing Machinery.
- [2] Lei Zheng, Vahid Noroozi, and Philip S. Yu. Joint deep modeling of users and items using reviews for recommendation. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, WSDM '17, p. 425–434, New York, NY, USA, 2017. Association for Computing Machinery.
- [3] 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, 2020.
- [4] Priscila Valdiviezo-Diaz, Fernando Ortega, Eduardo Cobos, and Raúl Lara-Cabrera. A collaborative filtering approach based on naïve bayes classifier. *IEEE Access*, Vol. 7, pp. 108581–108592, 2019.
- [5] Behnoush Abdollahi and Olfa Nasraoui. Using explainability for constrained matrix factorization. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, RecSys '17, p. 79–83, New York, NY, USA, 2017. Association for Computing Machinery.
- [6] Chong Chen, Min Zhang, Yiqun Liu, and Shaoping Ma. Neural attentional rating regression with review-level explanations. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, p. 1583–1592, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.
- [7] Caio Nóbrega and Leandro Marinho. Towards explaining recommendations through local surrogate models. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, SAC '19, p. 1671–1678, New York, NY, USA, 2019. Association for Computing Machinery.
- [8] Marco Ribeiro, Sameer Singh, and Carlos Guestrin. “why should I trust you?”: Explaining the predictions of any classifier. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, pp. 97–101, San Diego, California, June 2016. Association for Computational Linguistics.
- [9] Shun Morisawa and Hayato Yamana. *Faithful Post-hoc Explanation of Recommendation Using Optimally Selected Features*, pp. 159–173. Springer International Publishing, Cham, 2021.
- [10] Theodoros Lappas, Mark Crovella, and Evimaria Terzi. Selecting a characteristic set of reviews. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, p. 832–840, New York, NY, USA, 2012. Association for Computing Machinery.
- [11] Yinfei Yang, Yaowei Yan, Minghui Qiu, and Forrest Bao. Semantic analysis and helpfulness prediction of text for online product reviews. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pp. 38–44, Beijing, China, July 2015. Association for Computational Linguistics.
- [12] Julian McAuley and Jure Leskovec. Hidden factors and hidden topics: Understanding rating dimensions with review text. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, p. 165–172, New York, NY, USA, 2013. Association for Computing Machinery.
- [13] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [14] 東北大学乾研究室. Pretrained Japanese BERT models. <https://github.com/cl-tohoku/bert-japanese>.
- [15] sentencepiece. <https://github.com/google/sentencepiece>.
- [16] Scott M. Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, p. 4768–4777. Curran Associates Inc., 2017.