

## 事例研究 [論文]

## 記述式評価データを用いた推薦システムの試作

田中 蘭, 綿川 日菜, 中川 智之, 小林 正弘, 田畑 耕治, 松澤 智史

## 1. はじめに

情報技術の発達により, 大量の情報の中から自分にとって価値のある商品・サービスを見つけ出す手助けとして推薦システムが重要な役割を果たしている. 推薦システムのアルゴリズムの一つである協調フィルタリングは現在, EC サイトアクセスやストリーミング配信サービスなど多くのオンラインシステムに実装されており, ユーザーの評価や行動履歴から興味や嗜好の類似するユーザーを推定し推薦対象アイテムのスコアを算出している [1]. しかし, 収集されるデータであるユーザーの利用アイテムへの数値評価やアイテムの購入履歴・閲覧履歴などのデータからは, ユーザーがアイテムに対して具体的にどんな部分を良いもしくは悪いと思ったかといった情報を得ることができない. そのため, 異なる価値観をもつユーザーでも, 評価値が類似していれば類似ユーザーと見なされてしまう懸念がある. 一方, アイテムへの評価が具体的に記述されたデータが存在し, 推薦システムへの活用が可能である場合, この懸念は解消される.

そこで本研究では, 推薦システムにおける自然言語評価データの有用性を示すことを目的として, 推薦対象アイテムに対する記述式評価データを用いる推薦システムの試作を行い, 記述式評価データを用いない既存の推薦システムとの比較を行う. 記述式評価データを用いない既存の推薦システムは, 「アイテム選出の際に重視する項目」のデータを用いてユーザー間の類似度を算出する協調フィルタリングとしてアイテムのスコアの算出をする. 記述式評価データを用いるシステ

ムは, アイテムへの希望の記述を入力とし, 係り受け解析を用いて記述評価データ中に類似記述があるかどうかを判定しアイテムのスコアを算出する. そして両者の結果の比較を行うことで, 推薦システムにおいて記述式評価データを用いる利点を示す.

## 2. 協調フィルタリング

現在, 多くの EC サイトなどでは, ある対象ユーザーが商品を購入したデータと, 対象ユーザー以外が購入したデータの両方を用いて, その購入パターンからユーザー同士の類似性や商品間の共起性を相関分析し, パersonナライズされた商品を提示する推薦手法が用いられている. この手法は一般的に協調フィルタリングと呼ばれている. 協調フィルタリングは, 1992 年に米ゼロックスパロアルト研究所の「Tapestry」 [2] で登場し, Amazon.com の「この商品を買った人はこんな商品も買っています」で推薦システムの一般的な手法として認知されるようになり, 現在でもなお, 遺伝的アルゴリズムを組み合わせた手法 [3], 類似度算出の改良 [4], 関係性マイニングを組み合わせた協調フィルタリング [5] など, 発展研究も多い.

協調フィルタリングでは, 商品, 評価項目などの多次元の評価値行列をデータとして扱うため, 事前に収集する項目を決定しておく必要がある. データの活用方法が事前に決定している場合では, 事前の項目設定が容易なケースも存在するが, アンケートデータなど活用方法が多岐にわたるデータの場合, 事前の項目設定が困難であり, 使用用途によってはユーザーの嗜好の判断や推薦の基準とするには不十分なことも多い. たとえば, 「店内の明るさ」という項目の5段階評価を収集すると設定した場合, 「明るいため快適だ」と感じたユーザーも, 「暗いため落ち着く」と感じたユーザーも同じ5点をつけてしまう可能性がある. これに対して, 新たに「暗い店舗と明るい店舗のどちらが好きか」といった設問を新たに加えて再度評価を収集するなど対策が挙げられるが, 十分な回答数を収集するまでに時間がかかることや, 追加設問の検討・整備の手間

たなか らん, わたかわ ひな  
東京理科大学大学院理工学研究科情報科学専攻  
〒278-8510 千葉県野田市山崎 2641  
なかがわ ともゆき, たはた こうじ, まつざわ ともふみ  
東京理科大学理工学部情報科学科  
〒278-8510 千葉県野田市山崎 2641  
こばやし まさひろ  
東海大学理学部情報数理学科  
〒259-1292 神奈川県平塚市北金目 4-1-1  
受付 21.7.16 採択 21.11.3

がかかるとの懸念がある。

また、アイテムの利用目的もユーザーによってさまざまであるため、具体的な記述データが存在し直接的または間接的に別ユーザーが利用することが可能となれば便利である。たとえば、「店内の明るさ」について「景色を楽しむ際に適している明るさだが、読書をするには暗い」といった記述評価がある場合、景色を楽しみたいユーザーはこの店舗を検討し、読書をしたいユーザーは別店舗を検討するといった判断ができる。このように、記述式評価データの活用が可能になると特定の利用目的がある場合や利用する店舗・サービスにこだわりをもつ場合、アレルギーなどの制約のある場合など、さまざまなユーザーにとって有用である。しかし、記述式評価データは件数が多いほど、また1件あたりの文字数が多いほど閲覧の労力がかかるため、ユーザーの希望に合うアイテムを機械的に処理し提示することが望ましい。そこで本研究では、記述式評価データに対して自然言語解析を用いることで、協調フィルタリングでは行えない推薦システムを構築し、協調フィルタリングを用いた推薦システムの代替システム、補完システムとしての検討を行う。本研究で用いるデータには、アイテムの評価項目や実際の評価データ、また評価項目をもとに「良かった点」、「悪かった点」を具体的に記述するデータが存在するため、それらを用いて推薦システムの作成を行う。

### 3. データの概要

本研究では、令和2年度データ解析コンペティションにて提供された消費者満足度調査データを用いた。この調査は、過去一定期間以内もしくは一定期間以上継続して利用している特定業界（例：カフェ、定額制動画配信）の店舗・サービスを利用した人を対象に、年代・性別・居住地などの回答者の属性のほか、主に『過去一定期間内に最も利用した店舗・サービス（以後、本研究では最利用店舗と表現する）』について尋ねる内容となっている。具体的には、特定業界の最利用店舗の一つを選択し、総合的な満足度や設定されている評価項目（カフェ業界の例：店内の雰囲気、商品のおいしさ、店員の明るさなど）について利用後の評価や利用前の期待度を10段階で評価する設問、また評価項目のうちどれを重視しているか（以後、本研究では重視項目と表現する）などの設問で構成されている。なお、最利用店舗以外に過去一定期間以内に利用した店舗・サービスについて任意で回答できる設問も存在する。また、一部業界の調査については、最

利用店舗について良かった点や悪かった点を、評価項目を参考に自由記述式で回答する設問も含まれている。

本研究では、記述式評価データが存在し、かつ回答数他家業よりも多いカフェ業界の消費者満足度調査を用いて推薦システムの試作を行う。

## 4. 試作推薦システムの概要

本研究では、記述式評価データを用いないシステムとして従来の推薦アルゴリズムを用いた重視項目推薦、また記述式評価データを用いるシステムとして自由入力推薦を試作した。それぞれの概要は以下のとおりである。

### 1. 重視項目推薦

評価項目（消費者満足度調査の設問と同一）のうち重視する項目をすべて選択したものを入力とする推薦システム

### 2. 自由入力推薦

1の選択肢の具体的な内容や、1の選択肢以外の店舗・サービスの特徴や希望を記述したものを入力とする推薦システム

試作した両システムはそれぞれ上記のような入力を受け取って店舗ごとのスコアを算出し、上位10位までのランキングを出力する。

#### 4.1 重視項目推薦

重視項目推薦では、推薦アルゴリズムの一つとして広く用いられている協調フィルタリングが「アイテムの評価が似ている2人のユーザーは嗜好が似ている可能性が高い」という考えに基づき設計されていることから、「店舗選びの際に重視する項目が似ている本推薦システムのユーザーと消費者満足度調査回答者は嗜好が似ている可能性が高い」と仮定して店舗へのスコアリングを行う。

重視項目推薦は、ユーザーによる重視項目の入力（選択式・複数選択）を受けとった後、協調フィルタリングのアルゴリズムにならない、「類似者抽出」を行ってから「店舗ごとのスコア算出」という流れで作成する。

##### 4.1.1 類似者抽出

「推薦システムのユーザー」と「消費者満足度調査における各回答者」間の重視項目の類似度を算出する。そして、類似度が閾値以上の類似者データのみを抽出し、4.1.2節の店舗ごとのスコア算出に利用する。

本研究では類似度の計算に二つのベクトルのなす角を表現するコサイン類似度（式(1)）を用いる。



図1 類似者抽出の例

$$\cos(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}| |\vec{y}|} \quad (1)$$

また、店舗選びで重視すると選択している項目を1、選択していない項目を0と表現することで、システムユーザー・消費者満足度調査各回答者ベクトルをそれぞれ構成している。

類似者抽出の例を図1に示す。図1の例は、「快適さ」、「美味」、「店員」、「コスパ（コストパフォーマンス）」、「便利さ」の五つの評価項目に対して、重視するかしないかを二値で表現した様子である。システムのユーザー  $x$  の入力  $\vec{x}$  が (1, 0, 1, 1, 0)、回答者  $a$  の消費者満足度調査回答データが (0, 1, 1, 0, 1) であるため、類似度を式(1)を用いて算出すると0.33となる。同様にしてユーザー  $x$  と回答者  $b$  の類似度算出、ユーザー  $x$  と回答者  $c$  の類似度算出、…と消費者満足度調査の全回答者分の類似度を算出する。そして類似者抽出の閾値が0.7である場合、類似度が0.7以上である回答者  $c, d, e, f$  が抽出され、店舗スコアの算出時に利用される。

#### 4.1.2 店舗ごとのスコア算出

4.1.1 節で抽出した類似者を最利用店舗ごとに分類し、店舗ごとのスコア算出に利用する。本研究における店舗ごとのスコアは、より類似している回答者の評価が反映されるよう、最利用店舗に対する満足度と4.1.1 節で算出した類似度との加重平均とした。最利用店舗が  $A$  である類似者  $i$  の類似度を  $w_i$ 、類似者  $i$  による店舗  $A$  への総合満足度を  $x_i$ 、最利用店舗が  $A$  である類似者の数を  $n$  として、店舗  $A$  へのスコア  $\bar{x}$  は式(2)で与えられる。

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (2)$$

店舗スコア算出の例を図2に示す。図2の三つの表



図2 店舗スコア算出の例

はそれぞれ、抽出された類似者を最利用店舗 (A, B, C) ごとに分類したものである。そして各類似者はデータとして「4.1.1 節で算出したユーザーとの類似度」、「最利用店舗への総合満足度」の二つをもつ。それらのデータを用いて店舗スコアを式(2)によって算出すると、店舗Cのスコアが8.65、店舗Bのスコアが8.62、店舗Aのスコアが8.55となるため、推薦結果は1位が店舗C、2位が店舗B、3位が店舗Aとなる。

## 4.2 自由入力推薦

自由入力推薦では、ユーザーの入力した希望に合致する店舗へ加点、合致しない店舗へ減点ができるよう、記述式評価データを用い希望と合致するか判定を行い店舗へのスコアリングを行う。

具体的には、ユーザーの希望が「コーヒーがおいしい」であった場合に、「珈琲が美味しい」、「美味なコーヒー」などのデータが存在する店舗へ加点、「コーヒーが不味い」、「コーヒーが美味しくなく」などのデータが存在する店舗へ減点、「店内がおしゃれ」、「店員が親切」などの入力と無関係であるようなデータに対しては特に操作を行わないことを想定する。

### 4.2.1 合致判定

ユーザーの入力した内容と記述式評価データの内容が合致するかの判定は、以下の流れで行った。

**Step1** 類似単語判定

**Step2** 係り先・係り元の類似単語判定

**Step3** 対義語判定

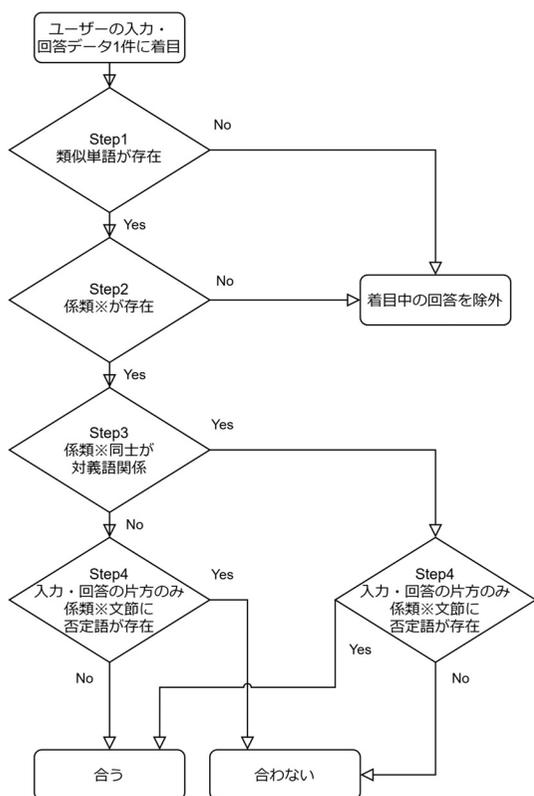
**Step4** 否定語判定

なお、この流れや各段階で使用する技術について株式会社オプティムの意見抽出の取り組み [6] を参考にしている。

また、消費者満足度調査に含まれる記述式評価データの設問は、最利用店舗の「他者推奨意向」や「よく行く理由」など複数存在するが、最利用店舗の「良かった点」、「悪かった点」の2問は設問文中に評価項目を参考に記述するよう指定があるため、本研究ではこれ

表 1 入力・記述式評価データ間の合致判定における各段階の状態

	Step1	Step2	Step3		Step4 (最終判定)	
	合う/合わない 可能性	合う/合わない 可能性	合う 可能性	合わない 可能性	合う 可能性	合わない 可能性
入力: <u>コーヒー</u> が <u>おいしい</u>	—	—	—	—	—	—
例 1: <u>美味しい</u> 珈琲	○	○	○	—	○	—
例 2: 店内がおしゃれ	—	—	—	—	—	—
例 3: おかわりの <u>コーヒー</u> がある	○	—	—	—	—	—
例 4: <u>コーヒー</u> の <u>不味さ</u>	○	○	—	○	—	○
例 5: <u>コーヒー</u> が <u>美味しく</u> ない	○	○	○	—	—	○
例 6: <u>コーヒー</u> が <u>不味く</u> ない	○	○	—	○	○	—



※係類……係り先・係り元類似単語

図 3 入力・記述式評価データ間の合致判定の流れ

ら 2 問の回答データを用いた。

図 3 に合致判定の流れ、表 1 に例文と各段階での判定の状態を示す。以下、各段階の判定の説明を行う。

### Step1：類似単語判定

入力と記述式評価データ間の名詞単語同士の類似度を算出し、類似度が閾値以上のものを類似単語として扱う。表 1 の例文中の下線部が類似単語判定段階に見つかる類似単語の例である。例文 2 以外のように類似

単語が存在する場合は、ユーザーの希望と合致するもしくは合致しない可能性があるとして抽出し Step2 へと進む。

また、今回は単語の分散表現として株式会社ホットリンクが配布している日本語大規模 SNS+Web コーパス Word2Vec モデル [7]、分散表現同士の類似度算出にはコサイン類似度 (式 (1)) を用いた。

### Step2：係り先・係り元の類似単語判定

Step1 で類似単語が存在すると判定された場合、入力・記述式評価データそれぞれの類似単語の『係り先・係り元文節の自立語 (名詞、動詞、形容詞、形容動詞) 単語』に着目し、再び類似単語が存在するかを判定する。Step1 との区別のために、係り先・係り元類似単語と表現する。

表 1 の例文中の太字部分が、Step1 での類似単語の係り先・係り元文節の自立語単語を表す。さらに、長方形で囲われている単語が係り先・係り元類似単語を表す。入力と例文 3 について考えると、各文の類似単語「コーヒー」の係り先・係り元文節自立語単語は入力「おいしい」に対して、例文 3 は「おかわり」、「ある」の二つが存在する。そして、「おいしい」、「おかわり」の類似度、「おいしい」、「ある」の類似度をそれぞれ計算し、片方でも類似単語であった場合はユーザーの希望と合致するもしくは合致しない可能性があるとして抽出する。

またこのとき、表 1 に挙げた例文中の『係り先・係り元文節の自立語単語』である「美味しい」、「おかわり」、「ある」、「不味さ」、「美味しく」、「不味く」について、直感的には「美味しい」、「美味しく」のみが「おいしい」と類似単語となるように思われる。しかし、Word2Vec は「ある単語の意味は周辺に出現する単語によって形成される」という分布仮説にもとづき学習をして分散表現を得ていることから、周辺に出現する

単語が似通いやすい対義語の類似度も高くなる性質がある [8]. そのため, この段階では例文 1, 5 のほかに例文 4, 6 も抽出されることが想定され, 例文 3 は入力内容とは無関係として, ユーザーの希望と合致するもしくは合致しない可能性のある候補からは除外する.

また, 本システムでの係り受けの解析や Step1 における形態素解析は, mecab-ipadic-neologd 辞書 [9] を用いて係り受け解析ツール Cabocha [10] で行った.

### Step3: 対義語判定

入力内容と Step2 で抽出された記述式評価データそれぞれの係り先・係り元類似単語が対義語関係であるか判別する.

まず, ユーザーの入力内容の係り先・係り元類似単語の対義語が何であるかを, 対義語が収録されている辞書を用いて検索する. そして, 検索結果の単語が, 記述式評価データの類似単語もしくは係り先・係り元類似単語と一致した場合, 該当の記述式評価データは「ユーザーの希望と合致しない可能性がある」と判定する. 表 1 に挙げた例文中では, 「おいしい」の対義語を検索して「不味い」が出た場合, 例文 4, 6 を「ユーザーの希望と合致しない可能性がある」と判定することが想定される.

本研究では対義語が収録されている辞書として英語 WordNet [11] を, また検索したい単語を英語 WordNet へ対応させるために日本語 WordNet [12] を用いた.

### Step4: 否定語判定

入力と記述式評価データの「係り先・係り元類似単語の存在する文節」同士に着目し, 片方の文節に否定語 (形容詞の「無い」や打消しの助動詞の「ない」, 「ぬ」など) が存在する場合, Step3 までの判定結果を反転させる. もし入力・記述式評価データ両方の「係り先・係り元類似単語の存在する文節」に否定語が存在した場合は判定結果の反転はしない. 表 1 に挙げた例文中では, 入力の係り先・係り元類似単語「おいしい」に否定語がついていないのに対して, 例文 5, 6 の係り先・係り元類似単語「美味しく」, 「不味く」の文節中には打消しの助動詞「ない」が存在するため, 例文 5 の判定を「希望に合致する可能性がある」から「希望に合致しない可能性がある」へ, 例文 6 は「希望に合致しない可能性がある」から「希望に合致する可能性がある」へと変更する. この段階での判定結果を最終結果とする.

#### 4.2.2 店舗ごとのスコア算出

4.2.1 節の操作を行うと, 最終的にユーザーの希望に合致する記述式評価データ・合致しない記述式評価

データの個数が得られる. それらを店舗ごとに集計しスコアを算出する. 最利用店舗 A への回答数を  $m$  とし, そのうち「ユーザーの希望に合致する記述式評価データ」の個数を  $m_{match}$ , 「ユーザーの希望に合致しない記述式評価データ」の個数を  $m_{unmatch}$  とし, 店舗 A へのスコア  $s$  は式 (3) で与えられる.

$$s = \frac{m_{match} - m_{unmatch}}{m} \quad (3)$$

## 5. 実装

本研究ではカフェ業界のアンケートデータ (回答数: 9,001 件) を用いて両システムの試作を行い, 令和 2 年度データ解析コンペティション OR 部会最終発表にて, 参加者向けに本システムの公開を行った. 重視項目推薦における類似者抽出時の閾値, 自由入力推薦における Step1 や Step2 で類似単語であるかを判定するための類似度の閾値は, システムのユーザーが決定できる形式で公開した. なお, 閾値を満たすカフェが存在しない場合は「指定の条件ではカフェが存在しません。」と表示される. また, 結果に表示されるスコアについては, 式 (2) や式 (3) で算出したスコアを 100 点満点に換算して表示を行った.

図 4 のように, 本システムは Ubuntu 上で Apache と PHP を連携させたシステムである. Apache を Ubuntu 18.04 LTS 上で動かし, ユーザーから入力されたデータを PHP で受け取る. PHP 上で Python ファイルを呼び, 受け取ったデータと CSV ファイルを用いて Python で計算して, ユーザーに推薦結果を表示する.

図 5 はユーザーの入力画面である. ユーザーは図 5 上部 I でカフェを選ぶ際に重視する項目を計 26 項目の中から当てはまる項目すべてを選択する. また, I の選択肢の具体的な内容や, I の選択肢以外の店舗・サービスの特徴や希望を図 5 下部 II のテキストボックスに日本語で入力する.

図 6 はユーザーへの出力画面である. 図 6 左部 I で重視項目推薦結果, 右部 II で自由入力推薦結果を提示する.

## 6. 評価

本節では, 重視項目推薦と自由入力推薦について, 推薦・検索分野における評価指標を用いた精度比較と, 具体的な結果例による比較について述べる.

### 6.1 重視項目推薦と自由入力推薦の比較

重視項目推薦と自由入力推薦の精度を比較するため

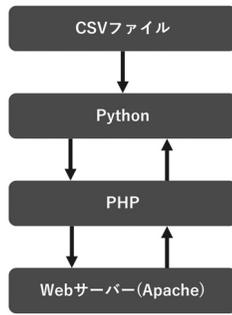


図 4 実行環境

図 5 入力画面例

重要項目ランキング I		自由入力ランキング II	
1位 モリバドコーヒー	89.0	1位 espressamente Rly	75.0
2位 カフェ・ド・クリエ	84.0	2位 カフェ・ド・クリエ アラタ	67.0
3位 タリアン・ド・マド カフェ	83.0	3位 珈琲 珈琲	67.0
4位 コスモ (Kona's Coffee)	82.0	4位 珈琲 HONOLULU'S Cafe	67.0
5位 スターバックスコーヒー	80.0	5位 UCCカフェラテ	59.0
6位 珈琲の珈琲	80.0	6位 セカフレド・ザネティ・エスプレッソ	57.0
7位 タリクスコーヒー	79.0	7位 上野珈琲店	56.0
8位 トナムコーヒーショップ	79.0	7位 珈琲	56.0
9位 珈琲珈琲	78.0	7位 珈琲	56.0
9位 ショートムズドコーヒー	78.0	7位 珈琲珈琲	56.0

図 6 結果画面例

以下のような被験者実験を行った。

- 以下のデータを被験者から収集
  - カフェに対する具体的な希望
  - (a) を入力とした際推薦されてほしい (適合) カフェの組合せ
- (a) を重視項目 (消費者満足度調査の設問と同一・複数選択) にあてはめ重視項目推薦へ入力
- (a) を自由入力推薦へ入力

表 2 重視項目推薦における類似者抽出閾値別の MAP

類似者抽出閾値	MAP
0.75	0.2561
0.70	0.0938
0.65	0.0895

表 3 自由入力推薦における Step1, Step2 閾値別の MAP

Step1 閾値	Step2 閾値		
	0.60	0.55	0.50
0.80	0.0924	0.1066	0.0762
0.75	0.0870	0.1107	0.0615
0.70	0.0727	0.1042	0.0784

4. 2 と 3 の結果と (b) を用いて、評価指標 MAP にて精度比較

なお、MAP とは、ユーザー (被験者) ごとに算出した AP (式 (4)) について全ユーザーで平均をとった評価指標である。

$$AP = \frac{1}{|Rel|} \cdot \sum_{i=1}^k relevant(i) \cdot P@i \quad (4)$$

ただし、

$$P@i = \frac{|Rel \cap Res[1, \dots, i]|}{|Res[1, \dots, i]|}$$

であり、Rel は適合アイテムの集合、Res[1, ..., i] は i 位までの推薦アイテムの集合、relevant(i) は推薦結果 i 位のアイテムが適合アイテムであれば 1、そうでなければ 0 を返す関数を表現している。また、本研究では推薦結果として 10 位までのアイテムを提示するため、k = 10 である。

上記の手順で 20 件のデータ<sup>1</sup>を収集し、重視項目推薦において類似者抽出の閾値ごとに MAP を算出した結果を表 2、自由入力推薦において Step1 の類似単語判定の閾値、Step2 の係り先・係り元類似単語判定の閾値ごとに MAP を算出した結果を表 3 に示す。

重視項目推薦では類似者抽出の閾値が 0.75 の際の MAP が 0.2561、自由入力推薦では Step1 の閾値が 0.75、Step2 の閾値が 0.55 の際 MAP が 0.1107 で最大となっている。また、被験者ごとに両システムの AP の差をとることで「対応のある 2 標本の平均の差の検定」を行ったところ、本実験結果では両システムの性能に有意な差があると確認できていない。そのため、次

<sup>1</sup> 本データは 2021 年 7 月 14 日～9 月 19 日の期間に行った Web 上アンケート結果より収集し、回答者は年齢 20～60 代の男女 20 名であった。

**表 4** 「店内・店外の雰囲気」、「店内・店外の清潔さ」、「立地・アクセスの良さ」、「商品のおいしさ」、「価格に対する商品の質」が入力の重視項目推薦結果

順位	店名	スコア (100 点満点)
1	珈琲館	82
1	星乃珈琲店	82
1	丸福珈琲店	82
4	スターバックスコーヒー	79
5	カフェ・ド・クリエ	78
5	シアトルズベストコーヒー	78
5	PRONTO (プロント)	78
5	むさしの森珈琲	78
9	エクセルシオールカフェ	77
9	コメダ珈琲店	75
9	ドトールコーヒーショップ	75

節にて具体的な推薦結果例を用い筆者らの考察によって評価を行う。また、定量評価の実現のために今後も継続的にデータ収集を行う。

## 6.2 結果例

「ゆったりした雰囲気のきれいな店内で、駅の近くにあり、料理がおいしく、価格が高くないカフェ」という希望をもつユーザーの場合の推薦結果について取り上げる。

重視項目推薦で選択できる項目のうち、該当する項目は「店内・店外の雰囲気」、「店内・店外の清潔さ」、「立地・アクセスの良さ」、「商品のおいしさ」、「価格に対する商品の質」となる。上記 5 項目を入力とし、類似者抽出の閾値を 0.7 とした場合の結果（有効回答数 219 件）を表 4 に示す。なお、同一の入力で類似者抽出の閾値を 0.75 とした際の有効回答数は 121 件、0.65 とした際の有効回答数は 464 件である。

また、自由入力推薦において、「ゆったりした雰囲気、きれいな店内、駅の近く、料理がおいしい、価格が高くない」と入力し、類似単語判定の閾値を 0.8、係り先・係り元類似単語判定の閾値を 0.55 とした場合の結果を表 5 に示す。

表 4 の重視項目推薦では上位に珈琲館、星乃珈琲店、丸福珈琲店がある一方、表 5 では STREAMER COFFEE COMPANY、珈琲館 蔵、すなば珈琲が上位にあり、珈琲館や星乃珈琲店は順位を下げている。この要因の一つとして、自由入力ランキングは店舗スコア算出（式 (3)）の分母に店舗への回答数を用いていることから、回答数が多い店舗は点数が上がりやすく、反対に回答数が少ない店舗はスコアが上がりやすい。そのため、回答数が 20 未満であるカフェが上位 3 位を占めている。

**表 5** 「ゆったりした雰囲気、きれいな店内、駅の近く、料理がおいしい、価格が高くない」が入力の自由入力推薦結果

順位	店名	スコア (100 点満点)
1	STREAMER COFFEE COMPANY	75
2	珈琲館 蔵	57
3	すなば珈琲	56
4	イタリアン・トマト カフェジュニア	52
4	PRONTO (プロント)	52
4	星乃珈琲店	52
4	UCC カフェプラザ	52
8	イタリアン・トマト カフェ	51
8	エクセルシオール カフェ	51
8	倉式珈琲店	51
8	珈琲館	51
8	サンマルクカフェ	51
8	ベックスコーヒーショップ	51
8	ペローチェ	51
8	むさしの森珈琲	51

次に、回答数が同程度の規模である珈琲館と PRONTO (プロント) に着目すると、自由入力ランキングでは珈琲館が重視項目ランキングより順位を下げた一方、PRONTO (プロント) は順位を上げている。この要因として、たとえば両店舗の消費者がおいしいと評価した商品の違いが挙げられる。「良かった点」、「悪かった点」の設問において商品のおいしさに言及している回答のうち料理（パスタ、サンドイッチなど）に言及している回答の個数と料理以外（コーヒー、紅茶など）に言及している回答を人手で計数したところ、珈琲館では前者が 9 件、後者が 22 件であった。一方、PRONTO (プロント) では前者が 11 件、後者が 6 件であった。このことから、PRONTO (プロント) は商品の中でも料理をおいしいと評価した消費者のほうが多く、珈琲館は料理以外をおいしいと評価した消費者のほうが多いといえる。そして、スコア算出に利用すべき回答のうち複数件を抽出できなかったものの、料理をおいしいと評価した消費者の方が多く PRONTO (プロント) の方が順位が高いことから、おいしいと評価された具体的な商品のデータがある場合はそれを利用した推薦システムのほうが、「商品のおいしさ」という項目を選択する推薦システムよりもユーザーにとって好ましい推薦をすることが可能といえる。なお、珈琲館で料理のおいしさに言及している回答 9 件中、合致判定アルゴリズムで抽出できたのは 1 件である。これは、9 件のうち「サンドイッチ」、「ランチ」、「モーニング」などがおいしいと記述されていたが、それらと「料理」が類似単語に判定されなかったためである。そして、PRONTO (プロント) で料理のおいしさに言及している回答 11 件中、合致判定アルゴリズムで

抽出できたのは5件である。抽出できなかった回答のほとんどは、直接係り受け関係のある2文節間に着目する本合致判定アルゴリズムでは対応できないケースであった。詳細は6.4.2節にて述べる。

### 6.3 被験者による感想

令和2年度データ解析コンペティション OR 部会最終発表時に得られたシステムへの被験者による感想を述べる。まず重視項目推薦について、「自分の行きつけの店が実際に上位に出た」、「自分と重視する項目の組み合わせが完全に一致する人はいないとしても、ある程度似ている人が行って評価した店舗を知ることができ楽しい」といった評価を得た。また自由入力推薦について、「『コーヒーがおいしい』と入力した際、コーヒーではなくエスプレッソが評判の有名チェーン店が上位に上がらなかったため直感に合っている」といった評価を得た。このことから、試作ランキング結果に対する直感的な納得感や、まだ馴染みのない店舗への興味を誘発できたことなどがうかがえる。

### 6.4 自由入力推薦の合致判定アルゴリズムの考察

本節では、自由入力推薦の合致判定アルゴリズムの考察として、良い点と悪い点について述べる。

#### 6.4.1 良い点

本研究の合致判定アルゴリズムをアンケートデータに適用することで、「良かった点」の回答データから減点要素、「悪かった点」の記述式評価データから加点要素を抽出することに成功した。本消費者満足度調査の「良かった点」、「悪かった点」の回答データには、「コーヒーは美味しくないが店内の居心地がよい」、「コーヒーはおいしいのに店員が暗い」など、良い点をたずねる設問に不満な点の記述や、反対に悪い点を尋ねる設問に満足な点を記述しているケースが一定数みられる。そのため、「良かった点」、「悪かった点」の回答データを設問ごとに別々に集計せずに、一度に両方集計できると効率的である。今回用いたアルゴリズムは「良かった点」、「悪かった点」といった設問の区別に関係なく、入力内容に存在する係り受け関係と「類似している係り受け関係」もしくは「反対の意味になっている係り受け関係」が回答の記述にも存在するかどうかに着目しており、一度に両方集計できるため適している。

#### 6.4.2 悪い点

本システムでは「コーヒー以外が美味しい」という記述式評価データに対して「コーヒーがおいしい」という入力にも希望が合うと判定してしまう。本システムではStep1で着目する単語は「文節内の接頭詞（「お味噌汁」の「お」など）を除いて一番最初に出現する

単語」としている。そのため、「コーヒー以外が/美味しい」という記述で着目される単語は「コーヒー」のみであり、「以外」は無視しており、望ましくない判定結果となる。

また、本研究の合致判定アルゴリズムでは直接係り受け関係のある2文節間のみに着目しているため対応しきれない例が存在する。たとえば入力が「優しくて明るい店員がいる」であった場合、「店員が明るい」という記述式評価データのほかに「不愛想な店員がいる」、「仲の良い店員がいる」といった記述式評価データも「ユーザーの希望と合う可能性がある」と判定され、「ベテランの店員がいない」が「希望に合致しない可能性がある」と判定されてしまう。このように判定されてしまうのは「店員が/いる」の2文節間に係り受け関係が存在するためである。そのため、直接係り受け関係の存在する2文節間に着目する方法ではすべてのケースに対応できない。このような複雑な係り受け関係に対応した合致判定アルゴリズムの検討は今後の課題とする。

### 6.5 実用面の課題

協調フィルタリングは購入履歴や閲覧履歴といった暗黙的に収集できるデータや、購入した商品への5段階評価といった手軽な数値評価データを利用できる利点から普及している。また、それらのデータを利用することで、ユーザーが改めてアイテムの希望を入力する必要なく「この商品を買った人はこんな商品も買っています」といった小窓を出すことが可能である。これに対して、本研究の自由入力推薦ではユーザーの入力の負担が大きい記述式評価データを収集する必要があり、推薦アイテムを提示する際はユーザーが改めて希望を言語化して入力する必要がある。これらの問題を軽減するには、商品を検索できるサービスの場合その検索ログを利用する、レビューを検索・レビューを評価（「このレビューは役に立ちましたか？」に対する「はい」「いいえ」など）できるサービスの場合閲覧したレビュー内容や評価を用いるといった方法でユーザーの嗜好データを収集することで、ユーザーの特別な入力を必要としない推薦システムへと応用できる。

また、重視項目推薦の類似者抽出時・自由入力推薦の類似単語判定時、係り先・係り元類似単語判定時に閾値を決定する必要があり、本研究ではプロトタイプとしてそれらをユーザーが自由に設定できるシステムを作成したが、この閾値によって推薦結果が変化する。そのため、閾値を固定した実用的なシステムを作成する際は閾値決定の方法を検討する必要がある。たとえ

ば、同一の自由入力に対して閾値をずらしながら推薦結果を被験者に提示し、最も再現率と適合率の調和平均や *MAP* が大きい閾値を採用する方法などが挙げられる。

## 7. おわりに

本研究では、推薦システムにおける記述式評価データの有用性を示すことを目的として、推薦対象アイテムに対する記述式評価データを用いた推薦システムの試作を行い、記述式評価データを用いない既存の推薦システムとの比較を行った。希望アイテムを具体的に記述し入力することができる記述式評価データを用いた推薦システムでは、予め設定された項目を入力とする推薦システムよりも具体的な希望が反映された結果を得ることができた。このことから推薦システムにおいて記述式評価データを用いることの利点を示すことができた。

**謝辞** この度消費者満足度調査データを提供いただきましたオリコン株式会社様、日本語大規模 SNS+Web コーパス Word2Vec モデルを提供いただきました株式会社ホットリンク様、データ解析コンペティションを運営いただいた方々に御礼申し上げます。また、論文を読んでいただき丁寧にコメントいただきました2名の査読者に御礼申し上げます。

## 参考文献

- [1] 岡本一志, 藤井流華, “協調フィルタリング入門,” 知能と情報, **31**, pp. 5–9, 2019.
- [2] D. Goldberg, D. Nichols, B. M. Oki and D. Terry, “Using collaborative filtering to weave an information tapestry,” *Communications of the ACM*, **35**, pp. 61–70, 1992.
- [3] B. Alhijawi and Y. Kilani, “A collaborative filtering recommender system using genetic algorithm,” *Information Processing & Management*, **57**, 102310, 2020.
- [4] A. Gazdar and L. Hidri, “A new similarity measure for collaborative filtering based recommender systems,” *Knowledge-Based Systems*, **188**, 105058, 2020.
- [5] 山脇淳一, 工藤康生, 村井哲也, “関係性マイニングと協調フィルタリングを用いた情報推薦手法,” 日本感性工学会論文誌, **17**, pp. 481–488, 2018.
- [6] OPTiM TECH BLOG, 「本気で自然言語処理やらないエンジニアでもできる, イベントアンケートの意見抽出」, <https://tech-blog.optim.co.jp/entry/2019/07/04/173000> (2021年7月9日閲覧)
- [7] 松野省吾, 水木栄, 榎剛史, “日本語大規模 SNS+Web コーパスによる単語分散表現のモデル構築,” 人工知能学会全国大会論文集 (第33回), 4Rin1-13, 2019.
- [8] M. Ono, M. Miwa and Y. Sasaki, “Word embedding-based antonym detection using thesauri and distributional information,” *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*, pp. 984–989, 2015.
- [9] 佐藤敏紀, 橋本泰一, 奥村学, “単語分ち書き辞書 mecab-ipadic-neologd の実装と情報検索における効果的な使用方法の検討,” 自然言語処理学会第23回年次大会発表論文集, pp. 875–878, 2017.
- [10] 工藤拓, 松本裕治, “チャンキングの段階適用による日本語係り受け解析,” 情報処理学会論文誌, **43**, pp. 1834–1842, 2002.
- [11] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross and K. J. Miller, “Introduction to WordNet: An online lexical database,” *International Journal of Lexicography*, **3**, pp. 235–244, 1990.
- [12] H. Isahara, F. Bond, K. Uchimoto, M. Utiyama and K. Kanzaki, “Development of the Japanese WordNet,” In *Proceedings of the Sixth international conference on Language Resources and Evaluation*, pp. 2420–2423, 2008.