

推薦システムにおけるユーザから好みを引き出すメカニズムの評価

安藤真之介[†] 加藤 誠^{††}

[†] 筑波大学 人間総合科学学術院 人間総合科学研究群 情報学学位プログラム 〒305-8550 茨城県つくば市春日 1-2

^{††} 筑波大学 図書館情報メディア系 / JST さきがけ 〒305-8550 茨城県つくば市春日 1-2
E-mail: †s2121603@s.tsukuba.ac.jp, ††mpkato@acm.org

あらまし 本研究では、推薦システムにおけるコールドスタート問題を解決するための、ユーザ自身から好みを引き出すフィードバック手法について比較実験を行った。本研究では、タグやアイテムを選択することでユーザの好みを表現する、タグベースとアイテムベースという2つの主要な手法に焦点を当て、これらを比較した。また、既存手法の問題点を解決するために、Tag Splitter と呼ばれる新しいフィードバック手法を提案する。Tag Splitter は、「アニメ AND 恋愛」や「アニメ AND 家族」などのタグの組み合わせを選択することによって自身の好みを表現し、簡単な操作のみでユーザの細かい好みを引き出すことが可能である。実験の結果、単一タグ選択システムは効率的で使いやすいが、他のシステムほど正確に好みを表現できないと感じられた。アイテム選択システムは時間がかかるが、推薦結果の精度が最も高かった。Tag Splitter は使いにくい、自身の細かい好みを最も正確に表現できると感じられた。

キーワード 情報抽出、コールドスタート問題、推薦システム、UI

1 はじめに

推薦システムにおけるコールドスタート問題を解決するため、推薦結果が重要となるサービスの多くでは、ユーザ自身にフィードバックを要求する機能を導入している。ユーザ自身に要求するフィードバックとしては、タグベースでの登録とアイテムベースでの登録の2つが主要な手法となっている。例えば、音声 SNS の Clubhouse では、サインアップ時に、ユーザの興味や好みを表したタグを選択することを求めている。また、Netflix では、新規ユーザに人気の映画や番組のリストを提示し、その中から好きなものを選択させている。タグベースの手法では、タグを選択することでユーザの好みを表現するよう指示し、アイテムベースの手法では、好みのアイテムを選択するような指示が行われている。

これらのフィードバック手法は、いずれも一長一短がある。タグベースの手法では、ユーザはそのドメインのアイテムについて詳しくなくても、フィードバックを容易に行うことができる。しかし、1つ1つのタグに含まれる情報量が少ないことで、ユーザの細かい好みを把握することが難しい場合がある。例えば、映画に関するサービスのフィードバックとして「恋愛」というタグを選択したからといって、ユーザはすべての恋愛映画が好きなのではなく、恋愛を扱ったアニメだけが好きであったり、SF がテーマの恋愛映画だけが好きな場合がある。しかし、細かい好みを引き出すために粒度の細かいタグをリストにして提示すると、ユーザの認知的負担が著しく増大し、サービスの利用を妨げるので、現実的な解決策とは言えない。アイテムベースの手法では、自分の好きなアイテムを選択することで、それぞれのアイテムが持つ豊富な情報を元に、ユーザは自身の好みをより正確に伝えることができるかもしれない。しかし、

サービス内のアイテムについて十分な知識がない場合、アイテムを選択することが困難となる可能性がある。また、タグベースの手法と比較すると、アイテムを選択する意図が必ずしも明確でないため、ユーザの好みを正確に把握することができないという難点がある。さらに、正確な推薦を行うために、ユーザがサービス利用開始時に十分な数のアイテムを選択するよう、長時間の作業を要求することが必要な場合がある。

本研究では、推薦システムのために、ユーザ自身から好みを引き出すフィードバック手法を比較する実験を行い、各手法の特徴を詳細に議論したことを報告する。また、既存のアプローチの問題点を解決するために、Tag Splitter と呼ばれる新しいフィードバック手法を提案する。図1は、Tag Splitter のインターフェースである。本システムでは、まずユーザがタグを一つ選択する。次に、ユーザは選択したタグに関連するアイテムのリストから、好みのアイテムを全て選択する。ユーザからのフィードバックを受けたシステムは、選択されたアイテムと選択されていないアイテムを上手く区別するようなタグの組み合わせを提示し、ユーザはそこから自分の好みを最も的確に表現するタグの組み合わせを選択する。Tag Splitter は、ユーザが最初に選択した粗い粒度のタグから、より細かい好みを表現できるタグの組み合わせを選択させることで、多くのタグやアイテムに煩わされることなく、ユーザの細かい好みを表現できるように設計されている。

Tag Splitter の第一の利点は、ユーザが自身の細かい好みを直接的に表現できることである。アイテムベースの手法は協調フィルタリングを行うために十分なユーザ数が必要になる。一方、本システムでは、ユーザが直接細かい好みを登録できることで、ユーザ数が少ない段階から効果的な推薦を容易に行うことが可能である。本システムの第二の利点は、登録過程においてユーザを煩わせないインターフェースである。Tag Splitter

は、ユーザが最初に選択した粒度の粗いタグを、そこからより細かい好みを表すタグの組み合わせに分割する。概念的には、親が単一の粒度が粗いタグ、子がタグの組み合わせであるツリー構造にユーザの好みが整理できる。これによって、一画面の要素数を少なくすることを実現し、ユーザの認知負担を軽減している。

本研究では、既存のアプローチと Tag Splitter を、30 名の参加者による比較実験により評価した。その結果、単一タグ選択システムは効率的で使い勝手が良いが、主観的・客観的に精度が低く、アイテム選択システムは時間がかかるが推薦に有効であり、Tag Splitter は使い勝手が悪いがきめ細かい好みを表現するのに有効であるという、それぞれのフィードバック手法の特徴が明らかになった。

本研究の貢献は以下のようにまとめられる。(1) 従来の問題を克服した新しいフィードバック手法 Tag Splitter を提案し、(2) 推薦システムにおけるフィードバック手法を定量的、定性的に比較し、(3) 異なるフィードバック手法の特性を分析した。

本論文は以下のような構成となっている。2 章では、コールドスタート問題と推薦システムのインターフェースに関する先行研究を簡単にまとめている。3 章では、Tag Splitter という本研究で提案するフィードバック手法と、比較対象となるタグベース、アイテムベースのフィードバック手法について説明する。4 章で実験で用いたデータセットと、実験設定について紹介する。5 章では、ユーザ調査の結果とシステムの実験結果について述べる。最後に 6 章では、本論文の結論について述べている。

2 関連研究

本研究に関連する研究は、コールドスタート問題に関する研究と、推薦システムにおけるインターフェースに関する研究に分けられる。

2.1 推薦システムにおけるコールドスタート問題

推薦システムにおいて、新規ユーザに行う最初の推薦は非常に重要である。なぜなら、システムが初期段階でユーザの好みでない推薦をした場合、ユーザがサービスから離れる可能性が高いからである [18]。また、近年の推薦アルゴリズム（協調フィルタリング [16], [19] など）は、ユーザのサービス利用履歴に大きく依存するため、初期に良い推薦を提供するのは基本的に難しい問題である。

コールドスタート問題を解決するための手法には、3 つのカテゴリがある。1 つ目のカテゴリは、コンテンツベースフィルタリングと協調フィルタリングを統合したものである。いくつかの研究では、各表現にユーザやアイテム固有の用語を含めるために行列分解を拡張し [1], [13], [14], [20]、さらに、画像やテキストの表現を効果的に学習するために深層ニューラルネットワークを適用した [12], [21]。Covington らはディープラーニングモデルを適用して大規模な動画推薦を行った [6]。また、短い Web セッション中にアイテムを推薦する問題を、行列分解

ではなく RNN アーキテクチャを用いて解決しようとした研究がいくつか存在する [10], [11]。2 つ目のカテゴリは、ユーザの好みを正確に推定するために効果的な個別のアイテムをユーザに対して積極的に推薦し、ユーザが好みのアイテムを選択できるようにするものである [8]。第三のカテゴリは、コールドスタート問題を解決するためにメタ学習を利用するものである [4], [7], [15], [22]。メタ学習とは、各タスクの学習サンプルが少なく、様々なタスクにまたがって学習する手法であり、各タスクをユーザの好みの推定と見なし、コールドスタート問題に適用している。

上記の研究の多くは、コールドスタート問題に対するアルゴリズム的な解決に焦点を当てたものであり、ユーザの好みを効果的に引き出すための仕組みやインターフェースについては考慮されていない。

2.2 推薦システムにおけるインターフェース

一方で、推薦のた目にフィードバックを得るための手法もいくつか提案されている。ユーザからの初期フィードバックを得るための手法は大きく分けて二つある。タグベースの手法 [2] とアイテムベースの手法 [5], [9], [17] である。タグベースの手法では、タグの集合から好みのタグを選択することで、ユーザに自身の好みを表現してもらう。一方、アイテムベースの手法では、ユーザは人気のあるアイテムの中から好みのアイテムを選択することで自身の好みを表現する。また、Facebook や Twitter などの外部リソースからユーザの好みを取得する手法もある [3]。本研究では、推薦システムにおける主なフィードバック手法を、推薦の有効性とユーザの満足度という観点から比較検討する。我々の知る限り、推薦システムにおけるフィードバック手法を比較検討した研究は、これまで行われていない。

3 システムデザイン

推薦システムにおけるフィードバック手法を比較するために、3 種類のシステムを実装した。3 種類のシステムとは、単一タグ選択システム、アイテム選択システム、Tag Splitter である。ユーザの好みが多様であり、ユーザが自身の好みを明確に説明するのが難しい推薦タスクであることから、実験では映画の推薦を対象とした。3.1 節で 3.3 節では、開発した各システムについて簡単に説明する。

3.1 単一タグ選択システム

単一タグ選択システムは、ユーザがタグを選択することで、自分の好みを表現する。図 1(a) に示すように、システムはタグの一覧を表示する。今回の実験において使用したタグは全部で 76 個あり、8 つのカテゴリーに分類されている（詳細は 4 章参照）。ユーザは、タグの一覧から自分の好みに合ったタグを選択することができる。ユーザがタグをクリックすると、タグの背景色がオレンジ色に変わり、そのタグを選択されたことを示す。オレンジ色のタグをもう一度クリックすると、背景色が白に変わり、そのタグが非選択の状態になったことを示す。ユーザが、自身の好みを十分に表現できたと感じるまでタグの選択

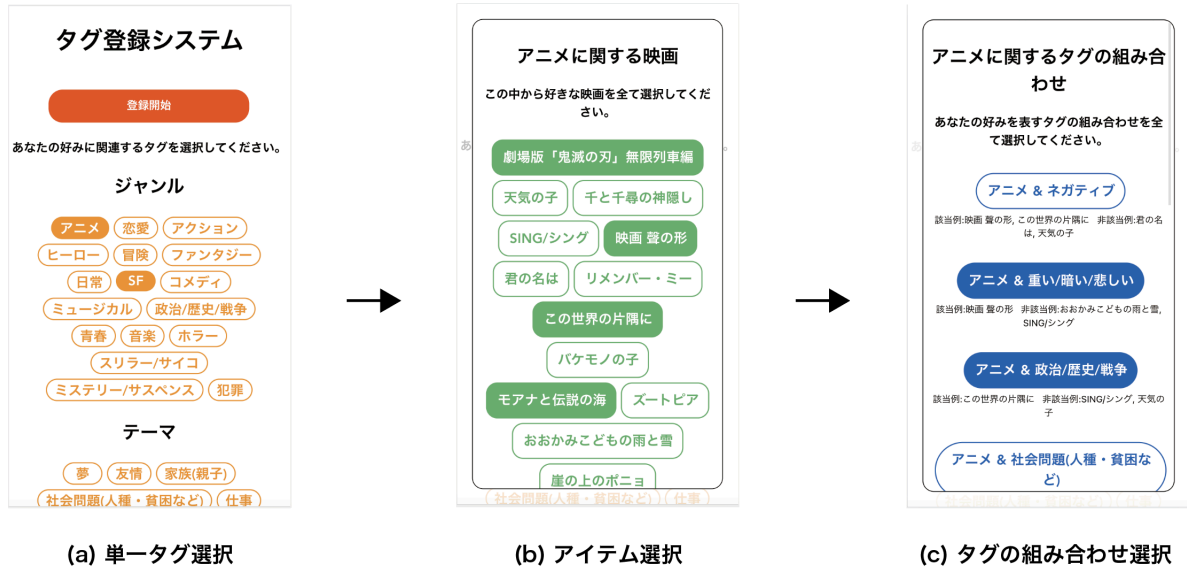


図 1 Tag Splitter のインターフェース。

を繰り返し、最後に送信ボタンをクリックして作業を完了する。

3.2 アイテム選択システム

アイテム選択システムは、ユーザが好みのアイテムを選択することで自身の好みを表現する。システムの最初のページには、単一タグ選択システムで表示されるタグリストと同じものが表示されている。ユーザがタグを一つ選択すると、図 1 (b) ように、選択したタグに関連するアイテムのリストが表示される。本実験で用いられる映画は 100 本であり、各映画には人力でタグが付与され、そのタグに基づいた表示がなされている。ユーザは、リストの中から自分の好みに合った映画を全て選択する。ユーザは、選択したタグのアイテム選択が完了した後、送信ボタンをクリックすると、再びタグリストを含む最初のページに戻る。ユーザはこのタグ選択とアイテム選択を繰り返すことで、自身の好みを表現する。

3.3 Tag Splitter

a) システム概要

Tag Splitter は、タグの組み合わせを選択することで、自身の好みを表現する。システムのインターフェースは、アイテム選択後の手順を除き、アイテム選択システムと同じである。ユーザが好みの映画を全て選択した後、Tag Splitter は図 1 (c) に示すようなタグの組み合わせのリストをユーザに提示する。タグの組み合わせは、ブール式のクエリに似ており、複数のタグと、AND と NOT (¬) 演算子から構成される。本システムでは、1 回の単一タグ選択に対して、ユーザに提示するタグの組み合わせは 10 個までとしている。ユーザは自分の好みを表すタグの組み合わせを全て選択する。選択が完了すると、ユーザは送信ボタンをクリックし、タグのリストが提示される最初のページに戻る。以降ユーザはこの手順を、自身の好みを十分に表現できたと感じるまで繰り返す。

b) タグの組み合わせ

単一のタグではユーザの好みを正確に表現できない可能性があるため、ユーザの好みを正確に伝える手段として、タグの組み合わせを考案した。本実装では、タグの組み合わせ C は、タグまたは否定されたタグ（例：“NOT アニメ”）の句の集合として定義される。すなわち、 $c_j = t$ または $c_j = \neg t$ (t はタグ) であり、 $C = \{c_1, c_2, \dots, c_k\}$ と定義する。) タグの組み合わせは、そのタグ節の接続として解釈される。つまり、タグの組み合わせが提示または処理されるとき、タグ節は AND 演算子で結合される。たとえば、 $C = \{\text{アニメ}, \neg \text{家族}\}$ は、「アニメ AND ¬ 家族」と表すことができる。形式的には、あるアイテム i がタグ組み合わせ C に関連すると判断されるのは、 i に割り当てられたタグ T_i が以下を満たす場合である。

$$\forall t \in T, t \in C \Rightarrow t \in T \wedge \neg t \in C \Rightarrow t \notin T, \quad (1)$$

が成り立つ。ここで、 T はタグの全セットである。

本研究では、複雑なタグの組み合わせを避けるため、 k を 2 または 3 に制限し、否定されるタグの数を最大 1 に制限している。従って、可能なタグの組み合わせは次のように限定される：A AND B, A AND ¬B, A AND B AND C, and A AND B AND ¬C, なお A, B, C はタグである。タグの組み合わせにはもう一つの制約があり、システムの最初のページでユーザによって選択されたタグを含むタグの組み合わせでなければならないと規定している。たとえば、ユーザが最初に「アニメ」タグを選択した場合、システムは「アニメ」を含む「アニメ AND ロマンс」または「アニメ AND ¬ 家族」などのタグの組み合わせを提案しなければならない。

c) タグの組み合わせのランキング

タグの組み合わせは多数存在するため、ユーザとの関連性に基づいてタグの組み合わせをランク付けし、ユーザに提示する数を絞ることが必要である。この場合、関連性は、タグの組み合わせがどれだけ正確にユーザの好みのアイテムを表している

かによって測定できる。これはさらに、Precision と Recall に分けることができる。

前述したように、Tag Splitter を使用する場合、ユーザはシステムの最初のページでタグを選択し、選択したタグを持つアイテムの集合から好みのアイテムを選択するよう指示される。選択されたアイテムを P 、タグの組み合わせ C に関連するアイテムの集合を R_C とする。タグの組み合わせ C の Precision と Recall はそれぞれ次のように定義される。Precision = $|P \cap R_C|/|R_C|$ 、Recall = $|P \cap R_C|/|P|$ 。

パラメータ β を持つ F 値は、タグの組み合わせをランク付けするための単一のスコアを生成するために使用される。 β の値を変えてタグの組み合わせのランキングを調べたところ、再現性を重視する $\beta = 2$ が妥当な結果を生むことがわかった。さらに、大きなタグの組み合わせ（すなわち、 $k = 3$ ）は、一般に小さなタグの組み合わせよりも高い F 値になるが、ユーザの好みを表すには特殊すぎる場合が多いことも確認された。そこで、タグの組み合わせのサイズが大きい場合には、スコアにペナルティを課し、以下のようなスコアリングでタグの組み合わせのランキングを行うことにした。

$$\text{Score} = \alpha_k F_2 = 3\alpha_k \frac{\text{Precision Recall}}{4 \text{ Precision} + \text{Recall}}, \quad (2)$$

ここで、 $\alpha_2 = 1$ 、 $\alpha_3 = 0.7$ である。可能なすべてのタグの組み合わせをランク付けすることで、タグの組み合わせのランク付けリストを得ることができ、上位 10 個のタグの組み合わせをユーザに提示することが可能となる。

3.4 推薦アルゴリズム

本研究では、それぞれのシステムにおける異なるフィードバック結果を元に、それぞれの推薦精度を算出してその有効性を比較する。公平に比較するためには、推薦アルゴリズムがシンプルで、かつ、どのシステムでも同じであることが望ましい。そこで、タグを主要な手段として、各アイテムをスコアリングし、推薦精度を算出する。

まず、各タグとタグの組み合わせについて、以下のように IDF 値を定義する。IDF(t) = $\log(N/\text{df}_t)$ 、ここで t はタグまたはタグの組み合わせ、 N はアイテム数、 df_t は t に関連するアイテム数である。ただし、タグが映画に割り当てられている場合、映画はタグに関連し、タグの組み合わせに対する関連性は式 1 で定義される。

タグに基づいてアイテムをスコアリングする主な考え方は、文書検索のベクトル空間モデルに類似している。ユーザのタグのセット T_u がある場合、アイテム i の関連性を以下のように推定する。

$$r_i = \sum_{t \in T_u \cap T_i} \text{IDF}(t) \quad (3)$$

ここで、 T_i はアイテム i に割り当てられたタグの集合である。同じスコアリング関数がすべてのシステムに適用されるが、 T_u の定義はシステムによって異なる。単一タグ選択システムでは、 T_u はユーザによって選択されたタグと定義する。アイテム選択システムでは、 T_u は、システムの最初のページで選択され

たタグと、選択された映画の中の頻出タグの和と定義する。タグの個数は、タグ選択システムで選択されたタグの平均値である 35 個に設定した。Tag Splitter では、 T_u は、アイテム選択システム用に定義されたユーザのタグと、ユーザが選択したタグの組み合わせの和と定義している。

4 実験設定

本章では、本実験で用いたデータセットと、実験手法について紹介する。

4.1 データセット

前述したように、今回の実験では、映画推薦のシナリオを使用した。映画は、国内で人気のある映画レビューサイト Filmmarks¹ から入手した。実験参加者の多くが観たことのある人気映画を選ぶため、2001 年以降に公開された映画で、2021 年 8 月 9 日時点でビュー数 75,000 回以上のものをランダムに選択した。その結果、149 本の映画を実験用に入手した。タグは、Filmmarks で使用されているジャンルや制作国のタグを抽出し、「夢」「正義」「動物」など、レビューで頻出するキーワードを手手で特定し、5 作品以上のレビューに登場するキーワードに限定して準備した。その結果、ジャンル、テーマ、感情、ストーリー、評価、主人公、制作国、制作会社のいずれかに分類された計 76 個のタグを得ることができた。タグの付与は、国内のクラウドソーシングサービスである Lancers² を利用し、各映画に精密に付与した。各映画には 3 名のアノテーターが割り当てられ、Filmmarks に掲載されている映画のレビューを精査した上で、関連するタグを選択するように指示した。タグは、2 人以上のアノテーターによって選択された場合に、その映画に関連しているとみなす。その結果、各映画に平均 17.0 個のタグが割り当てられた。また、Fleiss の kappa 係数で測定した評価者間の一致度は 0.497 であり、これは中程度の一致度である。

4.2 実験手法

単一タグ選択システム、アイテム選択システム、Tag Splitter の 3 つのシステムについて、定量的・定性的な評価を行った。過去 1 年間に 5 本以上の映画を鑑賞したことのある 30 名（男性 19 名、女性 11 名、平均年齢 22.9 歳、標準偏差 1.71）を被験者として募集した。参加者には映画に関する一定の知識が求められるため、筆者らの知人から年間で映画を 5 本以上観ている人を純粋にサンプリングしたものである。実験は遠隔で行われ、被験者間調査であり、各参加者は 3 つのシステムのいずれかにランダムに割り当てた。本研究は、筑波大学の倫理審査委員会によって承認された。

データセット中の 149 本の映画のうち、100 本はシステムの利用（アイテム選択）に使用し、残りの 49 本はユーザフィードバックに基づく推薦精度の評価に使用した。映画セットはランダムに 10 通りのグループ分けを行い、各システムグループの

1: <https://filmmarks.com/>

2: <https://www.lancers.jp/>

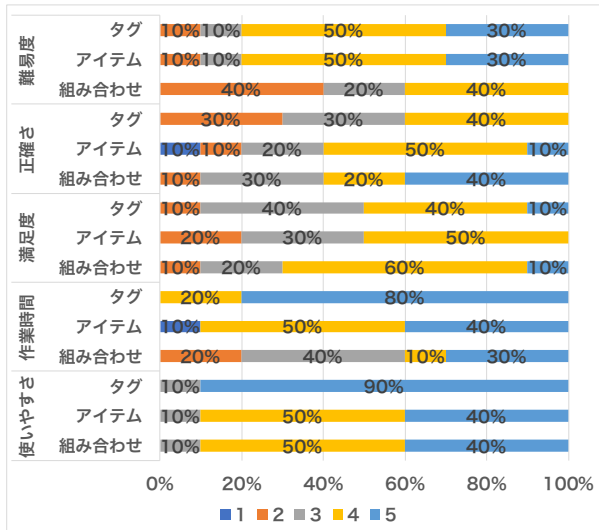


図2 ユーザによるアンケート回答の結果。

参加者はそれぞれ異なる映画セットを体験した。使用したセットは、3つのシステムで共通である。

実験では、まず参加者に同意書を読んでもらい、同意する場合は署名してもらうように指示した。その後、各システムに関する文書と、タスクに関する詳細について説明した。タスクの指示は各システム共通で、次のようなものであった。「タグ（またはアイテム）を選択することで、あなたの映画の好みを網羅的に表現してください」。

タスク終了後、参加者はアンケートに回答し、実験に使用したシステムに関する自由記述を行う。アンケートでは、実験に使用したシステムについて、1～5点で評価する。5は優れていることを示し、評価項目は、表現の難易度、表現の正確さ、満足度、作業時間の妥当性、使いやすさである。また、システムに対する評価に加えて、フィードバックに使用しなかった49本の映画についても、参加者の好みを示すように指示した。好みは、以下のいずれかによって評価する。「見ていない」(評点:0)、「悪い」(0)、「良くない」(1)、「良い」(2)、「素晴らしい」(3)。49本の映画について、これらの評点を用いて、各システムで提供されるフィードバックに基づく推薦精度の評価を行った。

5 実験結果

5.1 アンケート回答

図2は、参加者によるシステム評価のアンケート結果である。結果として、Tag Splitterは「表現の正確さ」と「満足度」において最も高いスコアを獲得したが、両スコアにおいて統計的に有意な差はなかった。そこで、自由記述欄のコメントから、Tag Splitterに対する好感度の理由を追加で調査した。

- **単一タグ選択システム:** 王道映画の中でも、私の好きなものと、嫌いなものがあるで、「王道」タグを選択することによって、私の好みがうまく表現されているかが心配になった。
- **単一タグ選択システム:** 自分の好みを表現するためには、より細かいタグが望ましいが、今より細くなると網羅的に選択するのが大変になりそう。

表1 3つのシステムによる推薦結果のnDCG。

システム名	nDCG@3	nDCG@5	nDCG@10
単一タグ選択システム	0.478	0.484	0.455
アイテム選択システム	0.438	0.494	0.535
Tag Splitter	0.480	0.428	0.439

- **Tag Splitter:** 選択したタグをより細かいタグに分解することで、自分の好みを認識することができた。

粒度の細かい嗜好表現のニーズが、Tag Splitterの「表現の正確さ」や「満足度」の高さにつながっているのかもしれない。予想通り、単一タグ選択システムは最も表現が不正確であると認識された。アイテム選択システムは、嗜好表現の精度は高いが、満足度はTag Splitterほど高くはない。

単一タグ選択システムは、作業時間の妥当性や操作性の点で好意的に受け止められている。一元配置分散分析の結果、Q4ではシステム効果が統計的に有意であり($F(2, 27) = 4.23$, $p < 0.05$)、Tukey HSDテストでは単一タグ選択システムとTag Splitterの間に統計的に有意な差異があったが、アイテム選択システムとTag Splitterには統計的に有意な差異がなかった。このように、単一タグ選択システムは効率的で使い勝手がよさそうだが、アイテム選択システムほど正確ではないことが、後の定量的にも示されている。

また、Tag Splitterが提示するタグの組み合わせは、ユーザのコメントにより、さらに改良が必要であることがわかった。特に、否定されたタグについては、Tag Splitterを利用した10人中5人が「否定を含むタグの組み合わせに戸惑った」と回答しており、好意的に受け止めていなかった。また、ある参加者は「感動 AND 驚き」と「感動 AND 驚き AND 愛」のどちらを選択するべきかで迷ったと回答している。これらは「表現の難易度」に関する設問で統計的に有意な差が生じた要因であると考えられる。一元配置分散分析の結果、システム効果は統計的に有意で合った($F(2, 27) = 3.75$, $p < 0.05$)。

5.2 推薦結果

表1は、ユーザ調査のアンケート回答で得られた正解データを用いて、各システムが生成した推薦結果のnDCGを示したものである。nDCG@3ではTag Splitterが最も高い精度を示したが、nDCG@5とnDCG@10ではそうでなかった。この結果は、Tag Splitterがユーザの狭く深い好みを表現するためには有効であるが、広く浅井興味をカバーするためには適していないことを示唆している。一方、アイテム選択システムは、nDCG@5とnDCG@10でより良い性能を示した。このように、フィードバック時間と推薦精度はトレードオフの関係にあると考えられる。アイテム選択システムとTag Splitterでは、フィードバックに時間がかかるが、そのフィードバックはより良い推薦結果を得るために非常に有益である。

また、タグ選択システムにおいて、ユーザのドメイン知識と推薦性能の間に強い相関があることが示された。また、映画に関する知識として、過去1年間に観た映画の本数を用いている。

nDCG@3 と映画鑑賞回数の相関係数は、タグ選択システムで 0.740, アイテム選択システムで 0.103, Tag Splitter で 0.174 であった。この結果は、決定的なものではないが、タグ選択システムで合理的な推薦結果を得るためには、ある程度のアイテム知識が必要であるという、タグ選択システムの欠点を指摘するものである。

6 結 論

本研究では、推薦システムのフィードバック手法の比較に関する調査を報告した。本研究では、タグベースとアイテムベースの手法、および Tag Splitter の 3 つの手法を検討した。その結果、Tag Splitter のシステムには改善の余地があることがわかった。しかし、この研究は、フィードバック手法のさまざまな特性に関する様々な知見を提供し、コールドスタート問題を解決するためのより良いフィードバック手法をさらに探求するために有用なものであると思われる。

謝辞 本研究は JSPS 科研費 21H03775 の助成を受けたものです。ここに記して謝意を表します。

文 献

- [1] Deepak Agarwal and Bee-Chung Chen. Regression-based latent factor models. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 19–28, 2009.
- [2] Fabiano M Belém, Jussara M Almeida, and Marcos A Gonçalves. A survey on tag recommendation methods. *Journal of the Association for Information Science and Technology*, 68(4):830–844, 2017.
- [3] Svetlin Bostandjiev, John O'Donovan, and Tobias Höllerer. Tasteweights: A visual interactive hybrid recommender system. In *Proceedings of the Sixth ACM Conference on Recommender Systems*, RecSys '12, page 35–42, New York, NY, USA, 2012. Association for Computing Machinery.
- [4] Fei Chen, Zhenhua Dong, Zhenguo Li, and Xiuqiang He. Federated meta-learning for recommendation. 02 2018.
- [5] Konstantina Christakopoulou, Filip Radlinski, and Katja Hofmann. Towards conversational recommender systems. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 815–824, 2016.
- [6] Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*, RecSys '16, page 191–198, New York, NY, USA, 2016. Association for Computing Machinery.
- [7] Manqing Dong, Feng Yuan, Lina Yao, Xiwei Xu, and Liming Zhu. Mamo: Memory-augmented meta-optimization for cold-start recommendation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 688–697, 2020.
- [8] Mehdi Elahi, Francesco Ricci, and Neil Rubens. A survey of active learning in collaborative filtering recommender systems. *Computer Science Review*, 20:29–50, 2016.
- [9] Mark P. Graus and Martijn C. Willemsen. Improving the user experience during cold start through choice-based preference elicitation. In *Proceedings of the 9th ACM Conference on Recommender Systems*, RecSys '15, page 273–276, New York, NY, USA, 2015. Association for Computing Machinery.
- [10] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. Session-based recommendations with recurrent neural networks. *CoRR*, abs/1511.06939, 2016.
- [11] Balázs Hidasi, Massimo Quadrana, Alexandros Karatzoglou, and Domonkos Tikk. Parallel recurrent neural network architectures for feature-rich session-based recommendations. In *Proceedings of the 10th ACM conference on recommender systems*, pages 241–248, 2016.
- [12] Donghyun Kim, Chanyoung Park, Jinoh Oh, Sungyoung Lee, and Hwanjo Yu. Convolutional matrix factorization for document context-aware recommendation. In *Proceedings of the 10th ACM conference on recommender systems*, pages 233–240, 2016.
- [13] Yehuda Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 426–434, 2008.
- [14] Xuan Nhat Lam, Thuc Vu, Trong Duc Le, and Anh Duc Duong. Addressing cold-start problem in recommendation systems. In *Proceedings of the 2nd international conference on Ubiquitous information management and communication*, pages 208–211, 2008.
- [15] Hyeop Lee, Jinbae Im, Seongwon Jang, Hyunsouk Cho, and Sehee Chung. Melu: Meta-learned user preference estimator for cold-start recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1073–1082, 2019.
- [16] Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, 7(1):76–80, 2003.
- [17] Benedikt Loepp, Tim Hussein, and Jüergen Ziegler. Choice-based preference elicitation for collaborative filtering recommender systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 3085–3094, New York, NY, USA, 2014. Association for Computing Machinery.
- [18] Sean M. McNee, John Riedl, and Joseph A. Konstan. Being accurate is not enough: How accuracy metrics have hurt recommender systems. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, page 1097–1101, New York, NY, USA, 2006. Association for Computing Machinery.
- [19] Andriy Mnih and Russ R Salakhutdinov. Probabilistic matrix factorization. In *Advances in neural information processing systems*, pages 1257–1264, 2008.
- [20] David Stern, Ralf Herbrich, and Thore Graepel. Matchbox: Large scale online bayesian recommendations. pages 111–120, 01 2009.
- [21] Chong Wang and David M Blei. Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 448–456, 2011.
- [22] Liang Zhao, Yang Wang, Daxiang Dong, and Hao Tian. Learning to recommend via meta parameter partition. *ArXiv*, abs/1912.04108, 2019.