

日本酒口コミサイトにおける 推薦アルゴリズムの効果に関する事例研究

新山大地¹ 羽山徹彩¹

概要：推薦システムは様々なオンラインビジネスに利用され、有益さが認識されているものの、それぞれの対象ごとのシステム設計ポリシーについてはほとんど明らかにされていない。その要因のひとつとして、推薦アルゴリズムとそれによるユーザの挙動について、商用プラットフォーム提供側でしか把握できず、その効果を確認できないことが挙げられる。そこで、本研究では日本酒を対象としたオンラインサービスを利用し、複数の推薦アルゴリズムがユーザのオンライン上の振舞いにどのように影響を与えるかを調査した。そのために、1 か月間のシステム運用のなかで、その期間にログインした 858 ユーザに 4 種類の推薦アルゴリズムのうち一種類を割り当て、システムログイン回数、推薦した日本酒銘柄の説明ページへの遷移数、飲みたいリストの日本酒銘柄登録数、および各日本酒銘柄へのコメント投稿数についてシステム履歴をもとに分析した。その結果、日本酒口コミサイトを対象とした推薦アルゴリズムの効果として、アソシエーション分析による推薦が個人の嗜好性と人気度を考慮した日本酒銘柄推薦システムに有効であることが分かった。

キーワード：事例研究、協調フィルタリング推薦システム、実験

A Case Study on the Effectiveness of Recommendation Algorithms for Sake Review Website

DAICHI NIIYAMA^{†1} TESSAI HAYAMA^{†1}

Keywords: Case study, Collaborative Filtering Recommender Systems, Experimentation.

1. はじめに

Twitter や Facebook といったソーシャルネットワーキングサービス (SNS) で、日々の生活習慣や購入品の評価[2][3]について共有する SNS ユーザが増加し、その共有された口コミやサービス上の挙動データには、ユーザの嗜好性やその傾向が含まれていることがわかってきた[1]。さらに、そのような共有データを大量に収集し活用した推薦システムは、ユーザの探索行為を補助し、購買意欲を高めることが示され、様々な商品やサービスに対して活用されている[6][8]。しかしながら、推薦対象となる商品やサービスに対する商用的な推薦システムの設計ポリシーについては、ほとんど明らかにされていないのが現状である。その原因のひとつとして、適用された推薦アルゴリズムとそれがユーザに与える影響については、商用プラットフォームの提供者しか制御やデータ収集ができず、その分析調査を困難にしていることが挙げられる。

これまで、商用プラットフォームを利用した SNS 共有データを分析した事例研究として、ワインやオンラインゲームを対象としてなされてきた。ワインを対象とした SNS データ分析では、920 万のワイン銘柄と 2990 万のレビューの

データに対し、ワインの産地と年代をもとにユーザの嗜好傾向が明らかになることを示した[4][7]。このようなユーザの嗜好傾向は、推薦システムに必要なユーザモデルの設計に有用であるものの、推薦システムの商用効果に影響を与えるユーザの挙動についてはまだ明らかにされていない。その一方で、オンラインゲームを対象とした推薦システムに関する事例研究では、約 2.2 万のユーザに対し、複数の推薦アルゴリズムに対するユーザの挙動を調査した[10]。その結果、推薦アルゴリズムがアイテムの閲覧数、ダウンロード数、売上の増加に影響があることを明らかにした。さらに、いずれの評価項目に対しても、アイテムベース協調フィルタリングの推薦アルゴリズムが有用であることを実証した。このような知見は推薦システムを設計するうえで、大変価値が高いものの、さらに一般化するためには異なる対象とした推薦アルゴリズムの影響に関する実践的な調査が必要である。

そこで、本研究では日本酒を対象とした口コミサイトにおいて、推薦アルゴリズムがユーザに与える影響について調査し、明らかにすることを目的とし、実施した。そのために、ユーザ数 858 人、アイテム数 23,000 件が登録されている日本酒のオンラインサービス「サケアイ」^aの 1 か月の

¹ 長岡科学技術大学
Nagaoka University of Technology

^a <https://sakeai.co.jp/>

運用のなかで、代表的な4種の推薦アルゴリズムがユーザの挙動に与える影響をシステム履歴から分析調査した。その結果、日本酒口コミサイトを対象とした推薦アルゴリズムの効果としては、アソシエーション分析による推薦アルゴリズムが個人の嗜好性と人気度を考慮した日本酒銘柄を推薦し、飲みたいリスト登録数およびページ閲覧数において有効であることが分かった。

2. 先行研究

本研究では事例調査をもとに、日本酒を対象とした推薦アルゴリズムがユーザに与える影響について明らかにすることを目的とする。そのために、顧客データセットおよび実運用による様々な推薦アルゴリズムの比較評価を行った研究について紹介し、調査対象となる有用な推薦アルゴリズムについて検討する。

Huang ら[9]は、オンラインの衣料品小売業と書籍のデータセット、および MovieLens プロジェクトbの映画評価データセットを用いて、複数の推薦アルゴリズムの評価を行った。その結果、アイテムベース協調フィルタリングは映画データセットに対し非常に高い精度を示したが、衣料品小売データセットおよび書籍データセットではリンク構造分析を用いた方が良い精度であった。また Matrix Factorization による推薦アルゴリズムは全てのデータセットにおいて平凡な精度であり、ユーザベース協調フィルタリングは他の推薦アルゴリズムよりも低い精度であった。Top-N アルゴリズムはある程度の高精度な結果が得られ、協調フィルタリングアルゴリズムと異なるアイテムが推薦されていた。

Pradel ら[14]は、ホームセンターの購入履歴のデータセットに基づき、複数の協調フィルタリングによる推薦アルゴリズムの比較実験を行った。その結果、アソシエーション分析による推薦が最良の精度が得られた。またユーザの完全な購買履歴を用いた場合には、Matrix factorization による推薦アルゴリズムが良い結果を示し、その要因としてユーザからの頻度の高いフィードバックが重要であると推察している。

Benouaret ら[13]は、フランスの石油企業 TOTAL から提供された顧客の購買履歴のデータセットに基づき、複数の推薦アルゴリズムの比較実験を行った。その結果、アイテムベース協調フィルタリングによる推薦システムの評価指標とランキング指標は最良の精度が得られ、それに次いで、Matrix factorization, アソシエーション分析による推薦アルゴリズムの順に良い精度が得られた。これらは学習データセットの取得時期が推薦アルゴリズムの精度に大きく影響を与えていることが推察された。例えば、アイテムベース協調フィルタリングが有効期限内のデータを、アソシエー

ション分析が直近のデータを、Matrix factorization が6ヵ月程度前のデータを、それぞれ学習データに用いることが有効と報告されている。またいずれの推薦アルゴリズムも単に人気の高いアイテムを推薦する方法よりも良い精度が得られることを示していた。

その一方で、実際に推薦システムを運営しながら、不特定多数のユーザを対象に、複数の推薦アルゴリズムを比較し評価した事例研究がある。Jannach ら[10]は、モバイルインターネットアプリケーションのゲーム推薦システムにおいて、ダウンロード数、売り上げ、およびページ遷移数をもとに、複数の推薦アルゴリズムを比較調査した。その結果として、お勧めリストへの追加およびダウンロード数では、コンテンツベースフィルタリングによる推薦が特に高い結果が得られ、それ以外の推薦アルゴリズムでも Top-N アルゴリズムの結果を上回っていた。また全体的な売上げ率向上においては、コンテンツベースフィルタリングの推薦結果が最良の結果を示していた。このような結果の要因として、ゲームという推薦対象がユーザの偏向性が高い嗜好であることが大きく影響し、また学習データの疎密性も影響することを推察している。さらに、推薦結果に対するユーザの挙動を調査することが商業利用のシステム設計指針に適合し易いことから、実環境で推薦システムを調査することの優位性を挙げている。このような商業利用のための推薦システムの設計指針に対し、一般化するためには異なる対象とした推薦アルゴリズムの実践的な調査が必要といえる。しかしながら、そのような調査については商用プラットフォーム提供側でしか実施することができないため、ほとんど事例調査がなされていないのが現状である。

以上のように、様々なデータセットに対して推薦アルゴリズムの影響が調査した既存研究では、推薦対象やデータセットの内容ごとに優位な推薦アルゴリズムが異なることがわかった。また実環境での調査研究は、推薦アルゴリズムに対するユーザの挙動も調査でき、商用的な推薦アルゴリズムの設計に大きく貢献する可能性があることもわかった。本研究では、既に商業的に運営している日本酒口コミオンラインサービスを利用し、複数の推薦アルゴリズムを比較調査することで、推薦アルゴリズムの効果を明らかにする。

3. 調査方法

3.1 対象

本研究では調査対象として、著者のひとりが2020年5月から運用している日本酒口コミサイト「サケアイ」を利用し、そのなかで日本酒銘柄を推薦する機能が与えるユーザの挙動を分析する。

b <https://grouplens.org/>

サケアイは日本酒の評価やコメントを共有することを目的としたスマートフォンアプリケーションであり、2021年12月時点で、年間アクセス数が23800、登録ユーザー数が9300（そのうち、毎月1回以上利用するアクティブユーザーが800～1000）、日本酒銘柄の登録数が24648、および口コミ投稿数が18018である。またサケアイには表1に示すような、日本酒の銘柄ごとの酒蔵が公開している成分や説明に関するデータベースも含まれている。

ユーザは図1に示すように、スマートフォンにインストールされたサケアイを起動することで、そのトップ画面（図1の(a)）から、おすすめの口コミ投稿（図1の(c)）やユーザが飲みたいリストに登録された日本酒銘柄一覧（図1の(b)）を確認できたり、詳細ページへの遷移（図1の(d)）ができたりする。また、図1の(e)のように、日本酒銘柄ごとに評価づける機能もある。さらに、サケアイには日本酒の銘柄の推薦機能があり、ユーザごとの過去に飲んだ銘柄



(a) トップページ画面



(b) 飲みたいリストの日本酒銘柄一覧



(c) 口コミ投稿の画面



(d) 日本酒銘柄の説明ページ



(e) 日本酒銘柄への評価付け画面



(f) 日本酒銘柄の推薦画面

図1 日本酒口コミサービス「サケアイ」のインタフェース画面

Figure 1 Screenshot of the interface for the sake review service "Sakeai".

表 1 日本酒のデータベース

Table 2 The configuration of template file.

項目
純米 or 醸造アルコール添加
大吟醸 or 吟醸 or その他
原料米 (麴米)
原料米 (掛米)
精米歩合 (麴米)
精米歩合 (掛米)
酵母
日本酒度
酸度
アルコール度数
アミノ酸度

リストやお気に入りの銘柄リスト, およびそれらの評価値をもとに, システムが推薦アルゴリズムを使用して, システム起動ごとに日本酒銘柄 3 種が推薦される (図 1 の(f)).

サケアイのシステム履歴には主に, ユーザのログイン日時, システム内の閲覧ページとそのアクセス日時, および日本酒商品の登録データとその日時が含まれる. ユーザごとに, 飲んだことがある日本酒銘柄リスト, マイリスト (お気に入り登録した日本酒銘柄リスト) も, そのユーザが各日本酒銘柄に付与したコメントと評価値とともに含まれる.

3.2 評価方法

3.1 節で述べた対象プラットフォームで, 複数の推薦アルゴリズムを適用し運営することで, システム履歴をもとにユーザの挙動で評価する. 適用する推薦アルゴリズムは 2 章で述べた先行研究のなかから主要なアルゴリズムとして, 「コンテンツベースフィルタリング」, 「アイテムベース協調フィルタリング」, 「アソシエーション分析」, および「Matrix factorization」を採用する. また推薦アルゴリズムごとのユーザの挙動を評価するために, 一定期間内でユーザが推薦結果を見てシステム操作する項目として「飲みたいリストに登録された日本酒銘柄数」, 「日本酒銘柄の詳細ページへの遷移数」, および「各日本酒銘柄への口コミ投稿数」を利用する.

本実験のなかで用いた各推薦アルゴリズムの実装方法について, 以下の(1)-(4)に示す.

(1) コンテンツベースフィルタリング

コンテンツベースの推薦システムでは, ユーザが既に興味を持っているアイテムと類似するアイテムを推薦する [5].

本研究が対象とする日本酒では, 表 1 の日本酒商品の登録データを使用し, 式(1)をもとにユーザが既に登録した日本酒商品とそれ以外の日本酒商品との類似度を算出し, その類似度が高い順に推薦する.

$$d(S, F) = \sqrt{\sum_{i=1}^n (S_i - F_i)^2} \quad (1)$$

ここで, $d(S, F)$, S_i , および F_i はそれぞれ, あるユーザが過去に高評価した日本酒の銘柄とある日本酒銘柄との類似度, ある日本酒銘柄の登録された i 番目の項目変数, およびある日本酒の登録された i 番目の項目変数を表す. それぞれの項目変数では 0/1 のダミー変数が用いられるが, 精米歩合, 日本酒度, 酸度, アルコール度数, およびアミノ酸度の項目では正規化した値が用いられる. n は表 1 が示すように日本酒商品ごとに 11 項目が含まれているため, 11 となる. つまり, $d(S, F)$ が 0 に近いほど, 日本酒銘柄間の距離が近く, 類似しているため, この数値が小さい順に日本酒銘柄が推薦される.

(2) アイテムベース協調フィルタリング

アイテムベース協調フィルタリングを用いた推薦システムでは, あるユーザが過去に評価したアイテムの評価と各ユーザが近い評価付けしたアイテム間の類似度を用いて, 未評価のアイテムの評価を予測する. その結果をもとに, ユーザごとに予測された評価が高いアイテムの順に推薦される [11].

本研究で対象とする日本酒銘柄の推薦では, 日本酒銘柄の i と j の類似度を, 式(2)をもとに算出し, あるユーザが過去にシステム中で評価した日本酒銘柄と類似度が高い銘柄を推薦する.

$$\text{sim}(i, j) = \frac{R_i^T R_j}{\sqrt{R_i^T R_i} \sqrt{R_j^T R_j}} \quad (2)$$

ここで $\text{sim}(i, j)$ は日本酒銘柄の i と j の類似度を表し, R_i は日本酒銘柄の評価の行列 R の i 番目列, R^T は日本酒銘柄の評価の行列 R の転置行列を表す.

(3) アソシエーション分析

アソシエーション分析を用いた情報推薦システムでは, アイテム間の価値のある組合せを抽出し, あるユーザが過去に評価したアイテムに関係する未評価のアイテムを推薦する [12].

本研究で対象とする日本酒銘柄の推薦では, 式(3)をもとに, ユーザの日本酒銘柄の評価履歴を使用し, ユーザごとに未評価な日本酒銘柄の評価の予測値を算出する. その結果, ユーザごとに, 評価の予測値が高い日本酒銘柄の順に推薦される.

$$score(u, j) = \sum_{i \in H_u} conf(i \Rightarrow j) \quad (3)$$

ここで $score(u, j)$, H_u , および $conf(i \Rightarrow j)$ はそれぞれ, あるユーザ u の未評価な日本酒銘柄 j の評価の予測値, あるユーザが評価した日本酒銘柄のリスト, および日本酒銘柄 i を気に入ったときに日本酒銘柄 j を気に入る割合 (信頼度) を表す. その信頼度は, 式(4)をもとに算出される.

$$conf(i \Rightarrow j) = \frac{R_i^T R_j}{\sum_{j=1}^n |R_j|} \quad (4)$$

ここで, R_i は i 列目の評価行列を表し, 日本酒銘柄 i に対するすべてのユーザの評価値を示す.

(4) Matrix Factorization

Matrix factorization を用いた情報推薦システムでは, ユーザを表す行列とアイテムの評価を表す行列に分解し, 再度乗算することで, 未評価アイテムの予測値を算出し, 高評価なアイテムを推薦する[13].

本研究で対象とする日本酒銘柄の推薦では, 式(5)をもとに, ユーザを表す行列と日本酒銘柄の評価を表す行列を乗算し, その結果の評価値が高い日本酒銘柄を推薦する.

$$\hat{p}_{ui} = x_u^T y_i \quad (5)$$

ここで, $\hat{p}_{ui}$, x_u^T , および y_i はそれぞれ, ユーザごとの日本酒銘柄の評価, ユーザの転置行列成分, および日本酒銘柄評価の行列成分を示す. 日本酒銘柄評価に対しては各評価の異常値の抑制のため, 交互最小二乗法を用いることで正則化を行う.

4. 評価結果

本調査のために, 日本酒口コミオンラインサービスのサケアイを 2021 年 11 月 11 日から 2021 年 12 月 30 日まで運用しながら, データを収集した. その間, 各ユーザは, 「コンテンツベースフィルタリング, アイテムベース協調フィルタリング, アソシエーション分析, Matrix factorization」の推薦アルゴリズムのひとつがランダムに割り当てられ, システム上に日本酒銘柄が推薦された. 評価期間における推薦アルゴリズムごとのユーザ数, およびユーザアクセス数の推移について, それぞれ表 2 および図 2 に示す.

表 2 本実験で各推薦アルゴリズムに割り当てられたユーザ数

Table 2 Number of Users Assigned to Each Recommendation

Algorithm in the Experiment.	
アルゴリズム名	ユーザ数
コンテンツベースフィルタリング	190
アイテムベース協調フィルタリング	228
アソシエーション分析	200
Matrix factorization	240

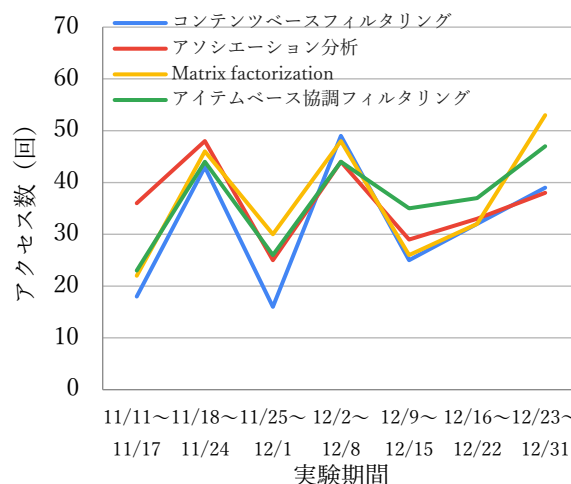


図 2 実験期間における各推薦アルゴリズムに割り当てられたユーザのアクセス数の推移

Figure 2 Number of System Accesses of Users Assigned to Each Recommendation Algorithm in the Experiment.

表 2 が示すように, 対象ユーザ 858 のち, コンテンツベースフィルタリングは 190, アイテムベース協調フィルタリングは 228, アソシエーション分析は 200, Matrix factorization は 240 と, それぞれの推薦アルゴリズムに割り当てられた. 各推薦アルゴリズムに割り当てられたユーザ数の違いは, 対象期間にログインされなかったユーザ数を除いた結果である. また各推薦アルゴリズムに割り振られたユーザは図 2 が示すように, 実験期間中にいずれも同程度, 同様な傾向のアクセス数が確認された.

次に, 各推薦アルゴリズムにおける飲みたいリストの日本酒銘柄の登録総数, ページ遷移数の総数, および口コミ投稿の総数について, それぞれ図 3-5 に示す. さらに各推薦アルゴリズムにおける飲みたいリストの日本酒銘柄登録, ページ遷移数, および口コミ投稿の日本酒銘柄の人気順位ごとの総数について, それぞれ図 6-8 に示す. また飲みたいリストの登録とページ遷移を行ったユーザは期間中にシステムにアクセスした総数の 10%程度であった. 日本酒銘

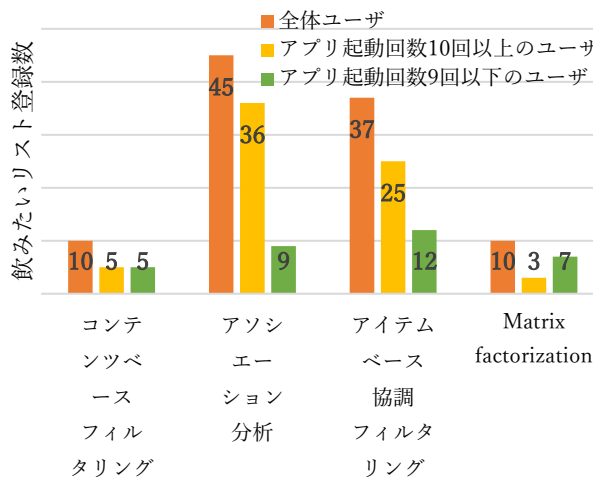


図3 本実験期間における各推薦アルゴリズムに割り当てられたユーザの飲みたいリストの日本酒銘柄の登録総数

Figure 3 Total number of sake brands on the "to drink" list of Users Assigned to Each Recommendation Algorithm in the Experiment.

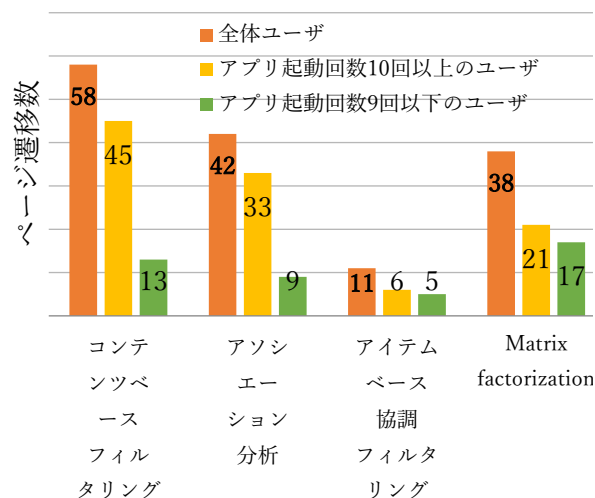


図4 本実験期間における各推薦アルゴリズムに割り当てられたユーザのページ遷移数の総数

Figure 4 Total number of page transitions of Users Assigned to Each Recommendation Algorithm in the Experiment.

柄の人気順位については、本システム内でのユーザ評価を参照し作成された。

飲みたいリストに登録された日本酒銘柄ではアソシエーション分析およびアイテムベース協調フィルタリングが45および37と、コンテンツベースフィルタリングおよびMatrix factorizationの各10に比べ、登録数が多かった。そのなかで、アイテムベース協調フィルタリングの主な特

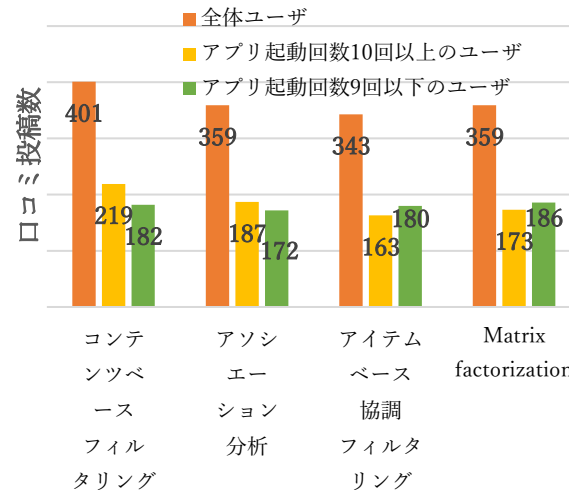


図5 本実験期間における各推薦アルゴリズムに割り当てられたユーザのロコミ投稿数の総数

Figure 5 Total Number of Sake Reviews of Users Assigned to Each Recommendation Algorithm in the Experiment.

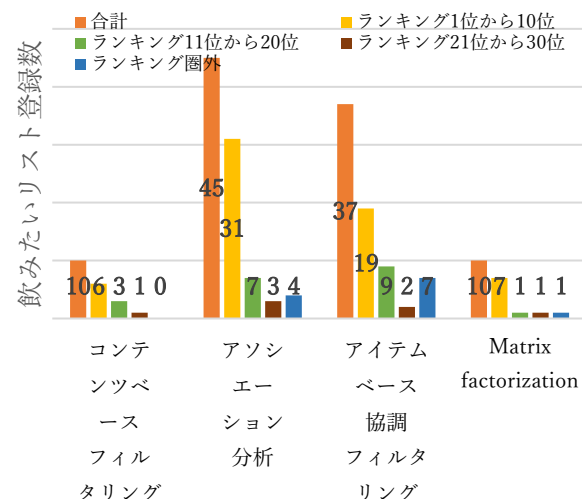


図6 本実験期間における各推薦アルゴリズムに割り当てられたユーザの飲みたい日本酒銘柄登録の人気順位ごとの総数

Figure 6 Total Number of Sake Brands for Popularity on the List of Users Assigned to Each Recommendation Algorithm in the Experiment.

徴として、システムを10回以上起動したユーザが飲みたいリストに登録した日本酒銘柄の総数の割合が0.67を占めていたことが挙げられる。その一方で、ページ遷移数の総数ではコンテンツベースフィルタリング、アソシエーション分析、およびMatrix factorizationが58, 42, および38と、アイテムベース協調フィルタリングの11に比べ、多い結果となった。そのなかで、コンテンツベースフィル

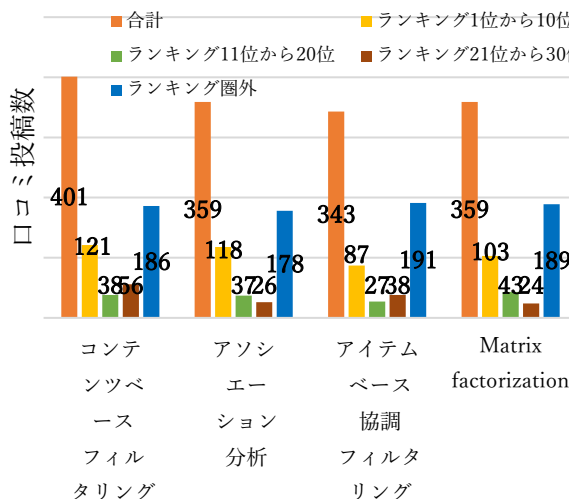


図8 本実験期間における各推薦アルゴリズムに割り当てられたユーザのロコミ投稿数の日本酒銘柄人気順位ごとの総数

Figure 8 Total Number of Sake Reviews for Popularity on the List of Users Assigned to Each Recommendation Algorithm in the Experiment.

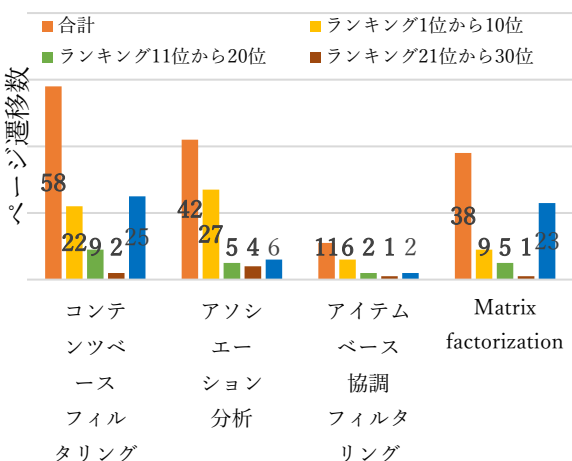


図7 本実験期間における各推薦アルゴリズムに割り当てられたユーザの日本酒銘柄の説明ページ遷移への人気順位ごとの総数

Figure 7 Total Number of Page Transitions for Popularity on the List of Users Assigned to Each Recommendation Algorithm in the Experiment.

タリングおよびアソシエーション分析を用いた推薦の場合には、システムを10回以上起動していたユーザのページ遷移数の割合が0.78および0.79と多く、アイテムベース協調フィルタリングおよびMatrix factorizationを用いた場合のベース遷移数は同等か差が4であるため、ユーザのシステム起動回数の多さに影響がみられなかった。

またロコミ投稿の総数は、コンテンツベースフィルタリ

ングが401と他のアルゴリズムよりも42以上多く、ユーザあたり平均2.11回投稿していた。その他の3種類のアルゴリズムでは、ロコミ投稿の総数が343から359の間であり、対象ユーザ数の違いを考慮しても、違いがほとんどみられなかった。またロコミ投稿数はすべての推薦アルゴリズムともに、システム起動回数10回以上と9回以下のユーザ数の割合が0.45から0.55の範囲内であるため、ユーザのシステム起動回数の多さに影響がないといえる。

また各推薦アルゴリズムが機能のなかで、人気のある日本酒銘柄がどの程度含まれていたかを確認した。飲みたいリストの日本酒銘柄の登録数においては、いずれの推薦アルゴリズムにおいても、人気順位10位以内の日本酒銘柄が半数以上含まれていることがわかった。日本酒銘柄の説明ページへの遷移数では、いずれの推薦アルゴリズムにおいても人気の高い日本酒銘柄が含まれているものの、コンテンツベースフィルタリングおよびMatrix Factorizationでは、20位以下の人気順位が低い日本酒銘柄も、0.43および0.61と高い割合で含まれていることがわかった。また日本酒銘柄へのロコミ投稿数においては、いずれの推薦アルゴリズムも順位が10位以上の人気が高い日本酒銘柄の割合が0.25以上含まれ、また順位が20位以下の人気の低い日本酒銘柄に対しても、コンテンツベースフィルタリング、アソシエーション分析、アイテムベース協調フィルタリング、およびMatrix Factorizationの推薦アルゴリズムにおいて、0.46以上の割合で含まれていることがわかった。

5. 考察

4章の評価結果が示すように、各推薦アルゴリズムにおいて、一定数以上のユーザがロコミ投稿をしながらアクティブにシステムを利用し、飲みたいリストへの日本酒銘柄の登録や日本酒銘柄の説明ページへの遷移の数に差があることを確認できたことから、推薦アルゴリズムの影響もみられた。そのなかで、アソシエーション分析およびアイテムベース協調フィルタリングによる推薦システムでは推薦リストからの登録が多くなされていた。その一方で、アソシエーション分析に加え、コンテンツベースフィルタリングおよびMatrix factorizationによる推薦システムでは推薦リストから日本酒銘柄の説明ページへの遷移が多くなされていた。

今回の調査では、推薦アルゴリズムに用いられた学習データ、および対象アイテムの違いが起因していたと推察する。

今回の推薦システムに用いられた学習データには、対象サービスの運用期間1年6か月と調査期間1か月に徐々に収集された、比較的新しいデータであった。アソシエーション分析、およびアイテムベース協調フィルタリングによる推薦システムでは、比較的近い時期に収集されたデータを用いることが有効とされている[13]。また、これら2種

の推薦アルゴリズムは調査期間中に利用頻度が高いユーザが多く、飲みたいリストへの登録数も多かったことから、学習データに直近のデータが含まれていたため、推薦が有効に機能していたと考える。

さらに今回の調査した対象アイテムの日本酒は個人の嗜好性だけでなく、人気度も、ユーザの興味に影響を与えやすいことが考えられる。推薦システムの事例調査に関する先行研究では、アイテムの嗜好性が強い対象（例えば、ゲーム[10]）、嗜好性と人気度が影響しやすい対象（例えば、映画や本[9]）、必要性が影響しやすい対象（例えば、日用品[14]やカー用品[13]）に分けられ、日本酒は嗜好性と人気度が影響しやすい対象に分類される。そのため、飲みたいリストの日本酒銘柄登録においては、アイテムベース協調フィルタリングおよびアソシエーション分析による推薦が有効に機能し、コンテンツベースフィルタリングによる推薦があまり機能しなかったと推察する。その一方で、日本酒銘柄の説明ページ遷移については、コンテンツベースフィルタリング、アソシエーション分析、および Matrix Factorization による推薦システムが説明ページへの遷移数が多く、そのなかでもコンテンツベースフィルタリングおよび Matrix Factorization による推薦システムでは人気が低い日本酒銘柄の説明ページに遷移する傾向がみられた。この傾向は、コンテンツベースフィルタリングおよび Matrix Factorization による推薦システムでは、飲みたいリストへの日本酒銘柄登録数が少ないことから、有効な日本酒銘柄推薦システムには人気の高い日本酒銘柄も含めることが必要であるといえる。

以上から、今回の調査において、日本酒口コミサイトを対象とした推薦アルゴリズムの効果として、アソシエーション分析による推薦がユーザの興味を反映した日本酒銘柄推薦に有効であることが分かった。

6. おわりに

本研究では日本酒を対象としたオンラインサービスを利用し、異なる推薦アルゴリズムがユーザのオンライン上の振舞いにどのように影響を与えるかを調査した。そのために、システム運用のなかで、1 か月間にログインした 858 ユーザに 4 種類の推薦アルゴリズムのうち 1 種類を割り当て、システムログイン回数、日本酒銘柄の説明ページへの遷移数、飲みたいリストの日本酒銘柄の登録数、および各日本酒銘柄へのコメント投稿数をシステム履歴から分析した。その結果、日本酒口コミサイトを対象とした推薦アルゴリズムの効果として、アソシエーション分析による推薦が個人の嗜好性と人気度を考慮した日本酒銘柄推薦システムとして有効であることが分かった。

今後、推薦アルゴリズムが与える影響として、投稿された口コミの内容についても、分析していきたい。

参考文献

- [1] S. Higgs and J. Thomas, Social influences on eating, *Current Opinion in Behavioral Sciences*, Vol. 9, pp.1-6, 2016.
- [2] L. Tiu Wright, C. Nancarrow, and P. M. Kwok, Food taste preferences and cultural influences on consumption, *British Food Journal*, Vol. 103(5), pp.348-357, 2001.
- [3] T. H. Silva, P. O. de Melo, J. Almeida, M. Musolesi, and A. Loureiro, You are what you eat (and drink): Identifying cultural boundaries by analyzing food & drink habits in foursquare, In *Proc. the Eighth International AAAI Conference on Weblogs and Social Media*, 2014.
- [4] Neema Kotonya, Paolo De Cristofaro, and Emiliano De Cristofaro, Of Wines and Reviews: Measuring and Modeling the Vivino Wine Social Network, In *Procs. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp.28-31 2018.
- [5] M. J. Pazzani and D. Billsus, Content-based recommendation systems, In: Brusilovsky P., Kobsa A., Nejdl W. (eds) *The Adaptive Web. Lecture Notes in Computer Science*, vol 4321. Springer, Berlin, Heidelberg, pp. 325–341, 2007.
- [6] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, Collaborative filtering recommender systems, In: P. Brusilovsky, A. Kobsa and W. Nejdl, Eds., *The Adaptive Web*, LNCS 4321, Springer-Verlag, Berlin Heidelberg, pp. 291-324, 2007.
- [7] Jung-Yeon Seo, Sang-Wook Han, and HwaMin Lee, A Wine Recommendation System using Collaborative Filtering, *Advanced Science Letters*, Vol. 23(10), pp. 10394-10398, 2017.
- [8] M. Deshpande, and G. Karypis, Item-Based Top-N Recommendation Algorithms, *ACM Transactions on Information Systems*, Vol. 22(1), pp.143-177, 2004.
- [9] Z. Huang, D. Zeng, and H. Chen, A comparative study of recommendation algorithms in e-commerce applications, *IEEE Intelligent Systems*, vol. 22(5), pp. 68-78, 2007.
- [10] D. Jannach and K. Hegelich, A case study on the effectiveness of recommendations in the mobile internet, In *Procs the third ACM conference on Recommender systems*, pp.205-208, 2009.
- [11] G. Linden, B. Smith, and J. York, Amazon. com recommendations: Item-to-item collaborative filtering, *IEEE Internet Computing*, Vol. 7(1), pp. 76-80, 2003.
- [12] R. Agrawal and R. Srikant, Fast Algorithms for Mining Association Rules in Large Databases, In *Procs the 20th International Conference on Very Large Data Bases (VLDB '94)*, pp.487–499, 1994.
- [13] I. Benouaret and S. Amer-Yahia, A Comparative Evaluation of Top-N Recommendation Algorithms: Case Study with Total Customers, *2020 IEEE International Conference on Big Data (Big Data)*, pp. 4499-4508, 2020.
- [14] B. Pradel, S. Sean, J. Delporte, S. Guerif, C. Rouveirol, N. Usunier, F. Fogelman-Soulie, and F. Dufau-Joel, A case study in a recommender system based on purchase data, In *Procs the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 377–385, 2011.