

タグプロファイルに対するユーザーフィードバックを用いた推薦システム

徐 銘蔓[†] 宮崎 純[†] 常 穹[†]

[†] 東京工業大学情報理工学院情報工学系 152-8550 東京都目黒区大岡山二丁目 12 番 1 号

E-mail: [†]xu.m.af@m.titech.ac.jp, ^{††}{miyazaki,q.chang}@c.titech.ac.jp

あらまし 本研究では、評価履歴から計算したユーザーが各タグへの嗜好度をユーザー特徴量、アイテムとタグとの関連性情報をアイテム特徴量として抽出し、ニューラルネットワークを用いて、推薦を行う方法を提案する。近年、情報推薦システムにおけるユーザー満足度を向上させるために、推薦精度だけでなく、システムの透明性、新規性やセレンディピティを評価指標とする研究も注目を集めている。さらに、推薦過程へのユーザーの関与度もユーザー満足度に大きく影響することが証明されている。そこで、本研究は、ユーザーの評価履歴情報から生成したタグクラウドをタグプロファイルとしてユーザーに提示し、ユーザーからのフィードバックを用いて推薦を行い、ユーザー満足度を向上させる。

キーワード 推薦システム, タグ, ニューラルネットワーク

1 はじめに

情報化技術の進展とインターネットの急速な普及により、情報の蓄積や処理にかかるコストが大幅に低下し、膨大な量の情報がウェブに流通している。誰でも様々な情報が容易に得られる一方で、この膨大な量の中から有用な情報を特定することはますます困難になっている。このような背景から、利用者の意思決定を支援する推薦システムに注目が集まり、様々な領域で活用されている。

従来多くの推薦システムでは、ユーザーの嗜好に合ったアイテムを推薦することが、ユーザ満足度の向上に繋がるという考えのもとで、「どれくらい正確にユーザーの嗜好に合致するアイテムを提示できるか」という推薦精度で推薦システムを評価することが重要視されてきた。しかし、近年では推薦精度だけを重視した推薦ではユーザー満足度が向上しないという指摘があり、推薦されたアイテムに対して多様性 (Diversity)、新規性 (Novelty) や意外性 (Serendipity) などの評価指標が新たに要求されるようになってきている。[1]

推薦システムにおける満足度についての研究では、推薦の透明性—ユーザーに推薦過程や理由などを理解させる能力—が推薦アイテムの受け入れられやすさやシステムへの信頼性、利用満足度などに寄与することが示唆されている [2][3]。システムの透明性を向上させるために、説明可能な推薦システムが多く提案されたが、「このアイテムはおすすめで、なぜなら…」などといった推薦するアイテムになぜ推薦されたかという理由を付与し、ユーザーに提供する推薦モデルは多いが、ユーザー嗜好の全体像を把握できるような説明可能な推薦システムに関する研究はほとんど見られない。

さらに、ユーザーの満足度を決定する重要な要因の一つとして、推薦システムにおいてのユーザー関与度が挙げられる [4]。従来の推薦システムではいかにユーザー負荷を軽減させるかという観点で研究されることが多いが、ユーザー負荷を多少増大

したとしても、ユーザーが推薦過程に関与すればするほど満足度が向上することが実験で確認されている [4]。しかし、現在多くの Web サービスでは、ユーザーの個人情報や行動履歴を利用して推薦を行うため、ユーザーが自分の嗜好に合わせて、推薦結果に影響を与える手段が非常に限られている。例えば、動画共有アプリの一つである Tiktok では、推薦された動画に「いいね」や「興味ありません」と言ったようなフィードバックに応じて、推薦アルゴリズムを更新する手段をユーザーに提供しているが、ユーザーの興味が大きく変化した場合、過去全ての関連評価を変更することが非常に難しい。また、ユーザが過去の行動履歴にない興味がある場合、その興味に関連するアイテムを検索、評価しない限り、推薦結果に反映されないことが多い。このような事情を踏まえて、ユーザーが知らず知らずのうちにフィルターバブルに陥る可能性が高くなり [5]、興味外の情報や新しい情報に触れる機会が少なくなる。

そこで本研究は、推薦システムの透明性を高めながら、より高い関与度をユーザーに提供することで、ユーザー満足度を向上させることを目的とする。具体的には、ユーザー嗜好の全体像を把握できるために、ユーザーの評価履歴情報から生成した個々ユーザーのタグ嗜好度をタグクラウドという形でユーザーに提示し、ユーザーがそのタグプロファイルを自分の嗜好に合わせて、タグへの嗜好度を調整することにより、また新たな推薦結果が得られるシステムを提案する。

従来の推薦手法では対象ユーザーに対して推薦を行う際に、推薦結果が毎回独立に決定されることが多いが、現実世界において、ユーザー嗜好が様々な要因によって過去に利用した時点から変化する場合がある。ユーザーの最新の嗜好に合うアイテムの推薦を実現するためにはユーザーの行動履歴の前後関係を考慮することが必要である。そこで、本研究のシステムの推薦アルゴリズムとして、アイテムとタグとの関連性情報をアイテム特徴量、ユーザーが各タグへの嗜好度と評価履歴の時系列データをユーザー特徴量として抽出し、ニューラルネットワークを用いて、推薦を行う。

本論文の構成は以下の通りである。2 節では、関連研究について説明する。3 節では、提案手法について説明し、4 節では実験および考察について述べる。5 節で本論文をまとめ、今後の課題について述べる。

2 関連研究

本研究は、推薦システムの透明性を高めながら、より高い関与度をユーザーに提供することで、ユーザー満足度を向上させることを目的とする。

透明性を考慮した説明可能な推薦に関する研究について 2.1 節で、推薦システムにおけるユーザー関与度を考慮したインタラクティブな推薦に関する研究について 2.2 節でそれぞれ説明し、本研究のベースとなるタグベース推薦に関する研究について 2.3 節で説明する。

2.1 説明可能な推薦

説明可能な推薦とは、ユーザーに推薦アイテムを提供することと同時に、なぜそのアイテムが推薦されるのかをユーザーに認識させる推薦アルゴリズムを指す。例えば、オンラインショッピングサービスサイト Amazon では「あなたと似ている人が買ったアイテム」や「今閲覧しているアイテムと似ているアイテム」といった説明をするアルゴリズムが説明可能な推薦の一種である。推薦システムにおける推薦説明はシステムに対する信頼性や満足度の向上、ユーザーの意思決定の支援など、様々な効果が期待されている。

説明の焦点は「入力: ユーザモデル」、「プロセス: アルゴリズム」、「出力: 推薦アイテム」の三つに分類されている。プロセスに説明の焦点を当てた推薦システムでは、アルゴリズムがどのように機能するかを理解することを目的としている。この点に着目し、様々な研究が行われてきた。例えば、Ribeiro ら [7] の任意の学習済み推薦モデルが与えられたとき、当該推薦モデルに対する入力-出力のペアを線形回帰モデルで学習し、結果得られた各特徴の重みを重要度として推薦モデルを解釈する手法を提案している。一方で、出力に説明の焦点を当てた推薦システムは推薦プロセスをブラックボックスとして扱い、推薦結果を正当化しようとする。例えば、Suzuki ら [8] のリカレントニューラルネットワーク (RNN) を用いて、多基準評価データとレビュー情報を使用して説明文を生成する手法や Wang ら [9] のナレッジグラフとレビューのアスペクト情報から抽出される感情を融合して推薦理由を生成する手法が挙げられる。それらに対して、入力に説明の焦点を当てた推薦システムが近年ますます重要視されている。これはユーザーがシステムに使われる個人データの収集と使用に不安感を持ち、インプットとしてのユーザーモデルの透明性への関心が高まるためである。入力に説明の焦点を当てた推薦システムとして、Hernan ら [10] が開発した「System U」が挙げられる。「System U」では、心理言語学的の分析により、ソーシャルメディアの情報から人々の性格特性を自動的に導き出すことができる。

2.2 インタラクティブな推薦

近年、Amazon や Netflix のようなウェブサービスでの推薦システムは、ユーザーの意思決定にかかる労力を減らすために、ユーザー嗜好の全般によく合う推薦アイテムを提示することを目的としている。しかしながら、ユーザーはシステムに支配されすぎて、より多様で斬新な情報に触れる機会が減ると感じるかもしれない。特に、まだ行動履歴が十分でない対象に対して推薦を行うことが難しいコールドスタート状況において、適切な推薦を得る前にユーザー側でかなりの労力が求められる。これらの問題点を踏まえて、近年推薦プロセスに対するユーザー関与度を高め、システムの透明性を向上させることに主眼を置いたインタラクティブな推薦に関する研究が注目されている。

例えば、J.Vig ら [11] が提案した Movie tuner モデルでは、ユーザーがタグへの嗜好度を調整することにより、現在推薦された映画と特有の特徴が異なり、全体的に似ている映画の推薦が得られる。その他にも、関与度に着眼した研究は多くある。例えば、Muhammad ら [12] が提案する TEA システムでは、ユーザーが年齢、性別、職業などの特徴の重要度を調整することができる。Chen ら [13] はクリティークング (Critiquing) と呼ばれる技術を開発し、ユーザーが推薦されたアイテムの属性に対して評価することにより、推薦結果を操作することができる手法を提案している。Bostandjie ら [14] が提案するモデル Taste Weights では、ユーザーのフィードバックにより、異なるソーシャルウェブ上の情報ソースを元に推薦を行うことが可能である。Denis ら [15] はアルゴリズムの結合の仕方をコントロールできるハイブリッド型の推薦システムを提案している。これらのアプローチは、推薦システムの透明性を向上させ、より満足度の高い推薦結果をユーザーに提供することを目的としている。

2.3 タグベース推薦

近年、Delicious、Flickr、Amazon などのウェブサイトでのタグを用いて検索したり、投稿したりすることが活発になっている。これはユーザーがタグを利用することで、コンテンツの管理や検索を容易にすることができるためである。タグは特定のアイテムが何であるか、何についてであるか、またどのような特徴を持っているかなどを説明することができ、推薦を説明するのに有用であると考えられる。また、タグにより、アイテムの特徴を捉えることができるだけでなく、ユーザーの一連のタグ付けの行動を通して、ユーザーの嗜好を把握することができる。Shirky [16] は、タグはユーザーによって作成されるため、ユーザーにとって意味のある概念を表すと主張している。タグは、ユーザーがアイテムと自分との未知の関係をよりよく理解するための橋渡しの役割を果たす。また、ユーザーのタグに対する態度は、ユーザーの嗜好を反映していると見られる。

これまでのタグベースの推薦システムに関する研究では、タグがユーザーのアイテムの発見、意思決定に役立つことが検証されている。Brooks [17] らはブログのタグとブログのテキストを用いたハイブリッドアルゴリズムを提案し、ブログを正確にクラスタリングすることを目指している。Diederich [18] らは、ユーザーが自分の興味に対応するタグプロファイルを作成し、

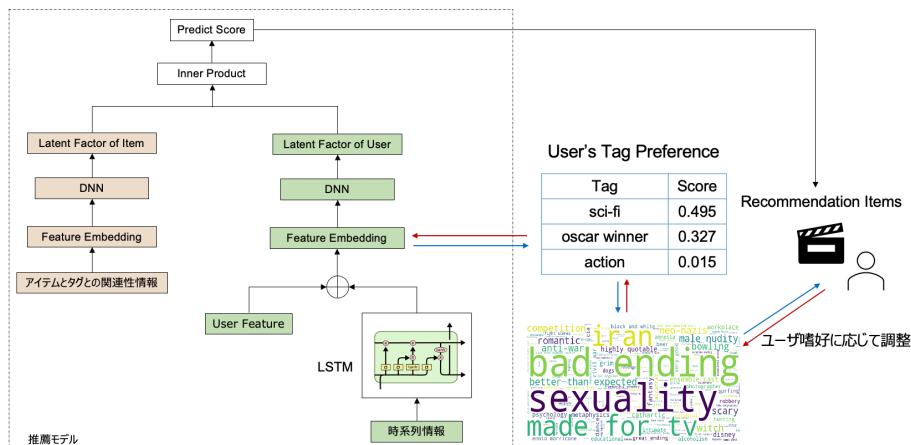


図1 提案する推薦システムのイメージ

そのタグプロフィールに基づいて推薦を受けるモデルを提案している。Niwa[19]らは、ユーザがタグ付けしたページとそのページに付けられたタグに基づいて、ウェブページを推薦するクラスタリングベースのアルゴリズムを提案している。

3 提案手法

本章では、提案するタグプロフィールに対するユーザーフィードバックを用いた推薦システムの仕組みについて説明する。システム全体のイメージ像を図1に示す。図の左側では提案する推薦モデルの構造を表している。具体的には、アイテム側とユーザー側がそれぞれアイテムとタグとの関連性情報、ユーザーの評価履歴をEmbeddingする。生成したItem EmbeddingとUser Embeddingをそれぞれニューラルネットワークを通して学習され、潜在的特徴量を生成する。アイテムの特徴量とユーザーの特徴量の内積を取ることで、評価値が予測される。予測された評価値よりTopKの推薦アイテムを生成し、ユーザーに提示する。それと同時に、ユーザーのタグ嗜好に基づいて、そのユーザーのタグプロフィールとしてタグクラウドを生成し、ユーザーに提示する。その次に、ユーザーは自分の嗜好に合わせてタグ嗜好度を調整し、システムはまたそのフィードバックに応じて推薦を行う。

提案システムをアイテム・モデリング、ユーザー・モデリング、レート予測の3つの構成要素に分割することができる。ユーザーモデリングについては3.1節で、アイテム・モデリングについては3.2節で、レートの予測については3.3節でそれぞれ説明する。また、ユーザーフィードバックの仕組みについては3.4節で説明する。

3.1 アイテム・モデリング

Item EmbeddingはJ. Vigら[20]が提案したタグゲノムからアイテムとタグとの関連性情報をニューラルネットワークを用いて学習する。タグゲノムがMovieLensにおけるユーザー調査から得られたタグの関連性に関する人間の判断に基づいて、ユーザーが作成したレビュー、評価値、タグの適用などのコンテンツの特徴を抽出し、作成されるものである。生物ゲノムは

DNAの全ての遺伝情報を含むように、タグゲノムはベクトル空間におけるアイテムと共通タグとの関係性情報を含む。タグゲノムでは、各アイテムとタグとの関連性がどの程度であるかを0から1の連続値で(0=全く関連性がない, 1=非常に強く関連する)示すことが可能である。例えば、表1より、タグゲノムより映画「Toy Story」とタグ「computer animation」との関連性と比べ、タグ「toys」との関連性がより高いことがわかる。

表1 タグゲノムの例

Toy Story (1995)	
Tag	Relev.
toys	0.999
computer animation	0.998
pixar animation	0.995
kids and family	0.995
animation	0.987
animated	0.979
children	0.973
cartoon	0.947
pixar	0.941
disney	0.929

アイテムのモデリングは全結合ニューラルネットワークを用いて計算される。各映画とタグとの関連性情報を入力とし、入力層のユニット数はタグの数に対応させる。アイテム特徴量 I はReLU活性化関数より計算される。

3.2 ユーザー・モデリング

User Embeddingは推測されたユーザーのタグ嗜好情報のEmbeddingと、ユーザー興味の変化を把握するための時系列データをリカレントニューラルネットワークモデル(RNN)の一種であるLSTM(長・短期記憶)を用いて生成したEmbeddingの組み合わせで学習する。

リカレントニューラルネットワーク(RNN)は連続データから時間的な変化のパターンを学習することに有効なモデルである。通常のニューラルネットワークでは、ある層の入力は次の層の入力に利用されるのみであることに対して、リカレント

ニューラルネットワーク（RNN）では、ある層の出力は、次の層の入力として利用されるだけでなく、出力としても利用される。しかし、リカレントニューラルネットワーク（RNN）は勾配爆発および消失問題より「長期の依存性」のあるデータを取り扱うことが十分ではない。この「長期の依存性」の問題を避けるために、本研究で提案するユーザーモデルでは LSTM（長・短期記憶）ネットワークを使用する。

提案するユーザーモデルでは、LSTM を用いてユーザーの評価履歴にある長期的な依存関係を学習する。具体的に、各ユーザーを「時系列最大の長さ × User Embedding の次元数」の行列で表し、評価履歴にある潜在的な依存関係を学習する。学習されたものを式 (1) のように、ユーザーの特徴の追加情報として用いる。 w_1 , w_2 はそれぞれユーザー Embedding と時系列 Embedding の重みである。そして、ユーザーのモデリングはアイテムのモデリングと同じく、全結合ニューラルネットワークを用いる。しかし、ユーザーのタグ嗜好情報を入力とする。

$$U_{embedding} = w_1 \cdot U_{feature_embedding} + w_2 \cdot U_{seq_embedding} \quad (1)$$

3.3 レートの予測

アイテム特徴量とユーザー特徴量を用いて推測される評価値は次の式 (2) より計算される。

$$\hat{r}_{ui} = \tilde{U} \cdot \tilde{I} \quad (2)$$

\hat{r} は推定される評価値という意味し、 \tilde{U} はユーザー特徴量を表し、 \tilde{I} はアイテム特徴量を表す。損失関数は式 (3) のように、実際の評価値と予測評価値との間の二乗誤差であると定義する。

$$J(\theta) = \sum_{r_{ui} \in R} (r_{ui} - \hat{r}_{ui})^2 \quad (3)$$

さらに、モデルの精度を向上させるために、ユーザー、アイテムごとの平均的な評価を表すバイアス項を導入する。ユーザーバイアスは、ユーザーがシステム全体の平均評価よりも高いまたは低い評価を与える固有の傾向と考えることができる。同様に、アイテムバイアスは、特定のアイテムが平均よりも高い評価や低い評価を受ける傾向を示している。バイアス項を考慮する場合、評価値の予測値は式 (4) のように変更する。

$$\hat{r}_{ui} = \tilde{U} \cdot \tilde{I} + \mu + b_i + b_u \quad (4)$$

ここでの μ は欠損値ではない全ての評価値の平均値である。 b_i はアイテム i の平均評価値から μ を引いて計算され、アイテムの評価バイアス項である。 b_u はユーザー u の平均評価値から μ を引いて計算され、ユーザーの評価バイアス項である。

3.4 ユーザーフィードバック

ユーザーの推薦アイテムの受け入れられやすさやシステムへの信頼性、利用満足度を向上させるため、ユーザーにフィードバックを要求する機能をシステムに導入している。ユーザーは自分の嗜好がシステムにどう捉えているかを直感的に理解でき

るようにするために、表 2 で示した user embedding から図 2 のようなタグクラウドを生成し、ユーザー嗜好のタグプロファイルとしてユーザーに提示する。このユーザーが図 2 のタグプロファイルより、自分がタグ「bad ending」、「sexuality」、「made for tv」と関連性の高い映画に興味を持っているとシステムに認識されていることを把握できる。

表 2 user embedding の例

tag	embedding weight
bad ending	0.049998
sexuality	0.049998
made for tv	0.049923
iran	0.049886
better than expected	0.049879
neo-nazis	0.049811
romantic	0.049804
male nudity	0.049743
witch	0.049578
competition	0.049498

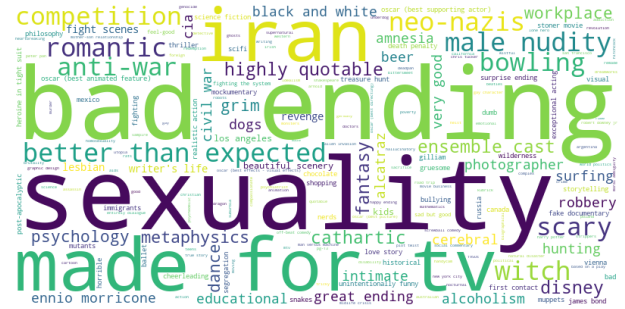


図 2 タグプロファイルの例

このタグプロファイルを用いて、ユーザーが自分の嗜好に合わせて、タグの大きさを調整することが可能である。タグの embedding weight は式 (5) のように更新される。 $weight_{tag}$ と $weight_{tag}$ はそれぞれ調整前と調整後のタグの embedding weight であり、 $weight_{variation}$ は調整されるタグの embedding weight の調整量である。

$$weight_{tag} = \frac{sum(weight_{tag}) \cdot weight_{tag}}{sum(weight_{tag}) + sum(weight_{variation})} \quad (5)$$

ここでは、例としてタグ「pixar animation」とタグ「007 (series)」の embedding weight をそれぞれタグ 0.0499、0.0500 にするように調整する。この調整によって、タグ「pixar animation」はタグ「made for tv」とタグ「iran」の間に位置するようになり、タグ「007 (series)」はタグ「romantic」とタグ「male nudity」の間に位置するようになる。更新された user embedding とタグプロファイルの様子をそれぞれ表 3 と図 3 に示す。

システムはこのユーザーからのフィードバックに基づき、図 1 で示した User Feature と、時系列データを LSTM（長・短期記憶）を用いて生成した Embedding の組み合わせで学習された Feature Embedding が変更される。この変更により、ユーザーモ

表3 user embedding の例調整後

tag	embedding weight
bad ending	0.049888
sexuality	0.049888
made for tv	0.049814
pixar animation	0.049790
iran	0.049777
better than expected	0.049770
neo-nazis	0.049702
romantic	0.049694
007 (series)	0.049691
male nudity	0.049634

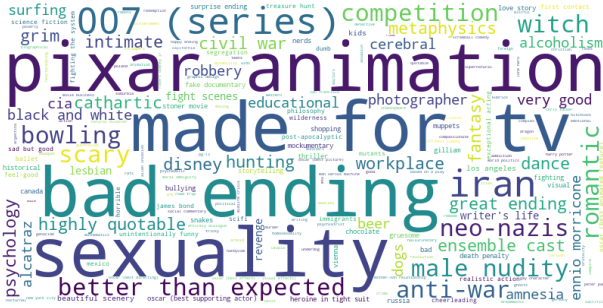


図3 タグプロファイルの例調整後

デリングが再実行され、ユーザー特徴量が更新される。これによって、新しい推薦リストが生成される。

4 評価実験

本研究は説明かつ調整可能な推薦モデルを構築することを目指している。そこで、データセット MovieLens-1M と MovieLens-25M を用いて、提案する推薦モデルの性能を比較し、評価した。また、ユーザーが生成したタグプロファイルにより、自分自身の嗜好を理解し、効率的に自分の嗜好を調整でき、その嗜好似合った推薦結果が得られることを検証するため、ユーザーフィードバックに基づく推薦の確認を行った。

4.1 データセット

提案手法の有効性を検証するために、GroupLens から公開されている MovieLens-1M (ml-1m) と MovieLens-25M (ml-25m) を利用して評価実験を行った。どのデータセットにおいても、各ユーザーは少なくとも 20 個の映画を評価している。実験に利用するデータセットの概要は表 4 の通りである。実験にあたり、使用するデータセットをタイムスタンプに従ってソートし、古いほうから 80 % のデータを訓練用データ、残りの 20 % をテスト用データに分割している。

4.2 評価指標

本実験では RMSE を用いて推薦精度を評価する。RMSE は評価値の予測における精度評価の指標であり、以下の式で定義される。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{r}_i - r_i)^2} \quad (6)$$

表4 データセットの概要

Dataset	ml-1m	ml-25m
Users	6,019	162,541
Movies	2,618	13,816
Ratings	404,273	24,674,113
Tags	1,128	1,128
Rating Scale	{1, 2, 3, 4, 5}	{0.5, 1, ..., 4.5, 5}

n は全予測対象数、 r_i , \hat{r}_i はそれぞれ、実測値と予測値を表す。

4.3 推薦モデル

ネットワーク (LSTM) とバイアス項がモデルの性能に対する影響を評価するために、以下四つのモデルを作成し、評価実験を行った。学習する際、隠れ層の次元数を 512、学習率を 0.001、バッチサイズを 1024 となるようにハイパーパラメータを設定した。

- Rec: アイテムとタグとの関連性情報を Item Embedding, 推測されたユーザーのタグ嗜好情報を User Embedding として学習されるモデルである。
- Rec++: アイテムとタグとの関連性情報を Item Embedding, バイアス項付きの推測されたユーザーのタグ嗜好情報を User Embedding として学習されるモデルである。
- RecLSTM: アイテムとタグとの関連性情報を Item Embedding, 推測されたユーザーのタグ嗜好情報に LSTM で学習された時系列データを加えたものを User Embedding として学習されるモデルである。(学習する時系列の長さの上限を 20 にした。)
- RecLSTM++: アイテムとタグとの関連性情報を Item Embedding, バイアス項付きの推測されたユーザーのタグ嗜好情報に LSTM で学習された時系列データを加えたものを User Embedding として学習されるモデルである。(学習する時系列の長さの上限を 20 にした。)

4.4 実験結果

以上四つのモデルで推薦を行う場合の RMSE の結果を表 5 に示す。

表5 推薦モデルの精度 (RMSE)

Model	ml-1m	ml-25m
Rec	3.785	3.737
Rec++	1.154	1.248
RecLSTM	1.094	0.986
RecLSTM++	0.971	0.925

4.5 考察

表 5 から、Rec++ 以外の推薦モデルは大規模なデータセットでより推薦精度が得られることが示された。この結果から、提案する推薦モデルは大規模データセットにおいても、高い精度で推薦を行うことが可能であることが見られる。

どのデータセットにおいても、モデル RecLSTM++ の RMSE が最も優れた推薦精度を示した。また、ユーザーとアイテムの

評価バイアスを考慮したモデル Rec++ と RecLSTM++ がそれぞれ評価バイアスを考慮しないモデル Rec と RecLSTM より良い推薦精度が得られることが示された。一方で、ユーザーの過去履歴の前後関係を考慮したモデル RecLSTM と RecLSTM++ がそれぞれ時系列データを考慮しないモデル Rec と Rec++ より良い推薦精度が得られることが示された。以上により、評価バイアスを考慮することと、時系列データを学習することで、モデルの性能を向上させ、高い精度の推薦結果をユーザーに提供することが可能であることが確かめられた。

4.6 Case Study

ここでは、提案するシステムに基づいた推薦を具体的に説明するため、データセット MovieLens-25m からランダムにユーザーを選び、そのユーザーに対して推薦を行う試行をする。また、ユーザーフィードバックを試行するため、対象ユーザーのタグプロファイルを適当に調整する。その調整に応じてまた新たな推薦を提示する。ただし、推薦モデルを評価実験で一番優れた性能を発揮した RecLSTM++ にする。

表 6 は対象ユーザーの嗜好情報が含まれる user embedding の一部であり、図 4 は user embedding から生成したタグプロフィールである。このタグプロフィールより、対象ユーザーがタグ「independent film」、「action」、「talky」、「black comedy」と「weird」などと関連性の高い映画に興味を持っているとシステムに認識されていることを把握することができる。

表6 対象ユーザーの user embedding (Case1)

tag	embedding weight
independent film	0.090688
action	0.070832
talky	0.070338
black comedy	0.067924
weird	0.067227
enigmatic	0.065787
off-beat comedy	0.063943
low budget	0.060274
funny	0.060227
creepy	0.060054



図4 対象ユーザーのタグプロフィール (Case1)

システムはタグプロファイルをユーザーに提示すると同時に、推薦アルゴリズムを通して推薦リストを生成し、ユーザーに提示する。推薦モデル RecLSTM++を用いて生成した推薦リスト

を表 8 に示す。ここでは、ユーザーフィードバックをシミュレーションし、タグ「sci-fi」、「adventure」と「drama」の大きさを変更し、weight を調整する。調整後の user embedding とタグプロフィールをそれぞれ表 7 と図 5 に示す。

表7 対象ユーザーの user embedding 調整後 (Case1)

tag	embedding weight
independent film	0.090290
sci-fi	0.079649
adventure	0.074671
action	0.070521
talky	0.070029
drama	0.069693
black comedy	0.067626
weird	0.066932
enigmatic	0.065499
off-beat comedy	0.063662



図5 対象ユーザーのタグプロフィール調整後 (Case1)

このユーザーフィードバックに応じて、user embedding が更新され、予測評価値が再計算されるようになる。新しく生成された推薦リストを表 8 に示す。

表8 対象ユーザーのフィードバック前後の推薦アイテム (Case1)

Top10	調整前	調整後
1	Heavenly Creatures	City of Lost Children, The
2	City of Lost Children, The	Heavenly Creatures
3	Space Jam	Space Jam
4	That Thing You Do!	That Thing You Do!
5	Artemisia	Yes, Madam
6	Glengarry Glen Ross	Artemisia
7	Miami Rhapsody	Glengarry Glen Ross
8	White Balloon, The	Miami Rhapsody
9	Loaded	What's Eating Gilbert Grape
10	Second Jungle Book	White Balloon, The

表9では、調整前後の推薦アイテムと、対象ユーザーの興味が高いTop5のタグ(「independent」、「action」、「talky」、「black comedy」、「weird」と調整した三つのタグ(「sci-fi」、「adventure」、「drama」との関連性情報を示している。表9より、ユーザー興味の強いタグと関連性の高い映画が上位に入ることが読み取れる。そして、調整前と調整後の推薦リストを比較すると、調整前に2位に位置した映画「City of Lost Children, The

表9 推薦アイテムとタグとの関連性 (Case1)

Movie	independent film	action	talky	black comedy	weird	sci-fi	adventure	drama
Heavenly Creatures	0.484	0.070	0.509	0.266	0.805	0.023	0.111	0.631
City of Lost Children, The	0.273	0.105	0.542	0.289	0.984	0.602	0.312	0.330
Space Jam	0.131	0.602	0.209	0.088	0.257	0.087	0.313	0.137
That Thing You Do!	0.170	0.153	0.220	0.103	0.276	0.043	0.212	0.360
Artemisia	0.253	0.116	0.263	0.102	0.428	0.029	0.104	0.420
Glengarry Glen Ross	0.527	0.161	0.828	0.374	0.470	0.019	0.086	0.931
Miami Rhapsody	0.486	0.101	0.376	0.177	0.225	0.043	0.117	0.279
White Balloon, The	0.280	0.067	0.403	0.147	0.466	0.019	0.197	0.196
Loaded	0.625	0.223	0.323	0.198	0.418	0.071	0.147	0.331
Second Jungle Book	0.101	0.184	0.085	0.065	0.357	0.045	0.313	0.150
Yes, Madam	0.074	0.959	0.229	0.090	0.179	0.058	0.200	0.192
What's Eating Gilbert Grape	0.444	0.076	0.503	0.226	0.527	0.020	0.079	0.790

がユーザーフィードバックによって 1 位になる。これは元推薦リスト 1 位に位置する映画「Heavenly Creatures」と比べ、映画「City of Lost Children, The」の方が調整したタグとの関連性が高いためである。また、調整前に Top10 推薦リストには入らなかった映画「Yes, Madam」と映画「What's Eating Gilbert Grape」がユーザーフィードバックによって、Top10 の推薦リストに入るようになる。この二つの映画は他の映画と比べ、調整した三つのタグと高い関連性を持ちながら、それぞれタグ「action」とタグ「weird」との関連性が非常に高いことにより推薦されたと考えられる。



図6 対象ユーザーのタグプロフィール (Case2)



図7 対象ユーザーのタグプロフィール調整後 (Case2)

提案する推薦手法の有効性を検証するため、もう一人のユーザーに対して同じような試行を行う。調整前のユーザータグプロフィールを図6に示し、タグ「oscar winner」と「vampires」

表 10 フィードバック前後の推薦アイテム (Case2)

Top10	調整前	調整後
1	Blackmail	Blackmail
2	Gordy	Black Beauty
3	Burnt by the Sun	Gordy
4	Black Beauty	Burnt by the Sun
5	Nemesis 2: Nebula	Nemesis 2: Nebula
6	Pretty in Pink	Mother Night
7	For Whom the Bell Tolls	Pretty in Pink
8	Philadelphia Story, The	For Whom the Bell Tolls
9	Broken Arrow	Philadelphia Story, The
10	House Party 3	Broken Arrow

を大きく調整後のタグプロファイルを図7に示す. 調整前後の推薦リストを表10に示す. 表11より, タグ「oscar winner」との関連性がより高い映画がユーザーフィードバックによって上位になることが見られる.

5 おわりに

本研究では、推薦システムの透明性とシステムにおけるユーザー関与度に着目し、ユーザーの評価履歴情報から生成したユーザーのタグ嗜好度をタグクラウドという形でユーザーに提示し、ユーザーがそのタグプロフィールを自分の嗜好に合わせて、タグへの嗜好度を調整することにより、また新たな推薦結果が得られるシステムを提案した。推薦モデルとしてアイテムとタグとの関連性情報をアイテム特徴量、ユーザーが各タグへの嗜好度と評価履歴の時系列データをユーザー特徴量として抽出し、ニューラルネットワークを用いた推薦モデルを用いた。時系列データとバイアス項の学習がモデルの性能に対する影響を評価するため、全四つのモデルの性能をそれぞれデータセット MovieLens-1m とデータセット MovieLens-25m を用いて比較し、評価バイアスを考慮することと、時系列データを学習することで、モデルの性能が向上することを示した。また、ユーザーフィードバックに基づくケーススタディを通して、ユーザーが自分の嗜好の全体像を把握でき、また自分の嗜好に

表 11 推薦アイテムとタグとの関連性 (Case2)

Movie	money	dark hero	paranoid	007 (series)	coming-of-age	oscar winner	vampires
Blackmail	0.066	0.818	0.624	0.019	0.278	0.710	0.003
Gordy	0.100	0.608	0.480	0.020	0.189	0.510	0.003
Burnt by the Sun	0.052	0.424	0.302	0.043	0.332	0.832	0.002
Black Beauty	0.067	0.337	0.263	0.019	0.275	0.948	0.003
Nemesis 2: Nebula	0.023	0.160	0.103	0.024	0.280	0.861	0.003
Pretty in Pink	0.081	0.449	0.235	0.019	0.449	0.797	0.003
For Whom the Bell Tolls	0.049	0.394	0.242	0.034	0.301	0.750	0.003
Philadelphia Story, The	0.024	0.426	0.197	0.038	0.256	0.795	0.003
Broken Arrow	0.017	0.401	0.142	0.034	0.143	0.843	0.005
House Party 3	0.632	0.150	0.342	0.015	0.124	0.454	0.003
Mother Night	0.038	0.458	0.616	0.018	0.216	0.962	0.003

合わせてタグへの嗜好度を調整することにより、また新たな推薦結果が得られることを確認した。それによって、より高い関与度をユーザーに提供し、推薦システムの透明性を高めることにより、ユーザー満足度を向上させることが期待できると考えられる。

今回提案した推薦システムは、ユーザー嗜好に合わせて調整することによって、新しい推薦結果をユーザーに提供することが可能であるが、ユーザーの調整量が小さい場合、推薦リストに変化をもたらすことができない可能性がある。また、このシステムはタグの間の関係性を考慮しないため、ユーザーフィードバックに応じて user embedding を更新する際に、調整するタグと類似度が高いタグは、ユーザー嗜好と逆方向に更新されるようになる。それによって、同じ嗜好に対して異なるユーザー嗜好度が学習され、システムがうまく稼働しない可能性がある。今後これらの問題を検討し、より満足度の高い推薦結果が得られるようにシステムを改善する必要がある。また、今後の課題としては、提案するシステムが実世界における効果について、ユーザー満足度の調査により検証していく必要がある。

謝 辞

本研究の一部は、JSPS 科研費 (18H03242, 18H03342, 19H01138) の支援を受けたものである。

文 献

- [1] Herlocker, Jonathan L., et al. "Evaluating collaborative filtering recommender systems." *ACM Transactions on Information Systems (TOIS)* 22.1 (2004): 5-53.
- [2] Sinha, Rashmi, and Kirsten Swearingen. "The role of transparency in recommender systems." *CHI'02 extended abstracts on Human factors in computing systems*. 2002.
- [3] Gedikli, Fatih, D. Jannach and Mouzhi Ge. "How should I explain? A comparison of different explanation types for recommender systems." *Int. J. Hum. Comput. Stud.* 72 (2014): 367-382.
- [4] Hijikata, Yoshinori et al. "The relation between user intervention and user satisfaction for information recommendation." *ACM Symposium on Applied Computing* (2012).
- [5] Pariser, Eli, *The filter bubble: What the Internet is hiding from you*, Penguin UK, 2011.
- [6] N. Tintarev and J. Masthoff, "A Survey of Explanations in Recommender Systems," 2007 IEEE 23rd International Conference on Data Engineering Workshop, Istanbul, Turkey, 2007, pp. 801-810
- [7] J. Ribeiro, J. Carmona, et al. "A Recommender System for Process Discovery. Business Process Management." Sep 2014, Eindhoven, Netherlands. pp.67 - 83.
- [8] Suzuki, Takafumi, et al. "Toward Explainable Recommendations: Generating Review Text from Multicriteria Evaluation Data," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 3549-3551
- [9] Z. Wang, Y. Li, L. Fang and P. Chen, "Joint Knowledge Graph and User Preference for Explainable Recommendation," 2019 IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 2019, pp. 1338-1342.
- [10] Hernan B. , Mateo N. Bengualid, et al. "System U: automatically deriving personality traits from social media for people recommendation." In *Proceedings of the 8th ACM Conference on Recommender systems (RecSys '14)*. Association for Computing Machinery, New York, NY, USA, 373-374.
- [11] J Vig, S Sen, J Riedl. "Navigating the tag genome." In *Proceedings of the 16th international conference on Intelligent user interfaces (IUI '11)*. Association for Computing Machinery, New York, NY, USA, 93-102.
- [12] Butt, Muhammad Hassan Farooq, Jian Ping Li, and Tehreem Saaboor. "A Tunable and Explainable Attributes (TEA) for Recommendation System." 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (IC-CWAMTIP). IEEE, 2020.
- [13] Chen, Li, and Pearl Pu. "Critiquing-based recommenders: survey and emerging trends." *User Modeling and User-Adapted Interaction* 22.1 (2012): 125-150.
- [14] Bostandjiev, Svetlin, John O'Donovan, and Tobias Höllerer. "TasteWeights: a visual interactive hybrid recommender system." *Proceedings of the sixth ACM conference on Recommender systems*. 2012.
- [15] Parra, Denis, Peter Brusilovsky, and Christoph Trattner. "See what you want to see: visual user-driven approach for hybrid recommendation." *Proceedings of the 19th international conference on Intelligent User Interfaces*. 2014.
- [16] C. Shirky. *Ontology is overrated*. [http://www.shirky.com/writings/ontology overrated.html](http://www.shirky.com/writings/ontology%20overrated.html), 2005. Retrieved on May 26, 2007.
- [17] C. Brooks and N. Montanez. "Improved Annotation of the Blogosphere Via Autotagging and Hierarchical Clustering." In *Proceedings of the 15th International Conference on World Wide Web*, 2006.
- [18] J. Diederich and T. Iofciu. "Finding Communities of Practice from User Profiles Based On Folksonomies." In *Proceedings of the 1st International Workshop on Building Technology Learning Solutions for Communities of Practice*, 2006.
- [19] S. Niwa, T. Doi, and S. Hon'iden. "Web Page Recommender System Based on Folksonomy Mining." *Transactions of Information Processing Society of Japan*, 47(5):1382-1392, 2006.
- [20] Vig, Jesse, Shilad Sen, and John Riedl. "Navigating the tag genome." *Proceedings of the 16th international conference on Intelligent user interfaces*. 2011.