

レビューテキストを対象としたハイブリッド型推薦手法における トピックモデルの有効性について

Effectiveness of Topic Model for Review Texts in Hybrid Model Recommendation Method

北原将平*¹
Shouhei Kitahara

ジェプカ ラファウ*¹
Rafal Rzepka

荒木健治*¹
Kenji Araki

*¹ 北海道大学大学院情報科学研究科

Graduate School of Information Science and Technology, Hokkaido University

In recent years, it is more and more important for recommendation system to show information considered to be useful for a user of growing electronic commerce on the Internet. In our paper, we propose a new method using topic model in hybrid recommendation method that combines the content-based filtering with the collaborative filtering for review text. In evaluation experiment, we compared rating prediction accuracies of our proposed method and conventional methods and confirmed effectiveness of the former.

1. はじめに

情報通信機器の発達により、個人が容易に情報を閲覧、蓄積、発信できる時代になった。しかし、一方で情報過多[神嶋 07]という問題も発生している。情報過多とは、大量の情報によって必要な情報が埋もれてしまい、問題の理解や意思決定が困難になる状況を指す。たとえば、類似した資料が大量にあり、目的の文献を探し出して理解することに予想以上に時間が必要となってしまう状況がある。情報通信機器の普及にともない、今後も情報過多の問題はさらに深刻化していくと考えられる。

このような状況を解決するために、推薦システムの必要性はますます重要となっている。推薦システムとは、特定の情報を必要としている利用者に対して、有用と思われる情報を提供するシステムのことであり、現在、Amazon や楽天といった e-commerce の企業では実用化されており、推薦システムによる個人の嗜好を正確に反映した情報提示が望まれている。

推薦システムに関する手法としては、既存のユーザが過去に行動した履歴を用いて情報を予測する協調型推薦[Resnick 94]がある。協調型推薦は大きくユーザベースとアイテムベースの2種類の手法に分類される。また、[Wang 06]らの手法では、これらを組み合わせた推薦手法を提案しており、予測精度の向上を実現している。

一方で、推薦されるアイテムの情報を用いる内容ベース型推薦の研究としては、家電製品を対象にユーザのこだわりを反映させた手法を提案した[服部 12]の研究がある。また、[奥 13]では書籍のドメインにおいてビブリオバトルと推薦システムの比較を行っている。奥らは書籍に関する表層的な情報を用いるだけでは推薦のユーザ満足度が下がるということ報告している。

書籍という商品は、家電のような実用品ではなく、嗜好品の一種であるためユーザの評価基準が曖昧である。さらに、書籍を分類する役目でもあるジャンルは書籍の内容に応じて整備されていない。たとえば、Amazon では、夏目漱石の著書である「こころ」は日本文学のカテゴリになっており、「こころ」の「悲しい恋」の内容を反映した分類とはなっていない。

そこで、商品の内容について多くの情報が書かれているレビューテキストを利用する手法が考えられる。レビューテキストを利用

した推薦問題の評価値予測に取り組む研究はいくつかある。そのなかでも、[Ganu 09]はレビューテキストの極性情報を利用した推薦手法を提案している。また、[岡田 12]はレビューテキストに tf-idf 法を適用した推薦手法を提案している。両者の手法の共通点は、レビューテキストの表層語を用いて評価値を予測している点である。一方で、本稿ではレビューテキストの表層に出現しない潜在的な意味に着目する。

本研究の目的は、書籍を対象とした推薦システムの評価値予測タスクにおいて、推薦精度を向上させることである。そこで、本稿では、「内容の似ている書籍は類似した評価値を付与される」という仮説のもと、評価値予測のタスクにおいて、書籍のレビューテキストにトピックモデル[Blei 03]を適用することで書籍の内容を反映させたアイテム間の類似度を用いるハイブリッド型推薦手法を提案し、評価値予測の精度向上を試みる。トピックモデルは潜在的意味(トピック)を扱うことが可能なため、書籍のレビューテキストにおいて表層には出現しない情報を利用することができる。

本稿の構成は以下の通りである。まず 2 章では、関連研究について述べる。3 章では、提案手法について述べる。4 章では評価実験について述べる。5 章では評価実験における評価値予測タスクの結果と考察について述べる。最後に 6 章で結論と今後の課題を述べる。

2. 関連研究

本実験でベースラインとして提案手法と比較するユーザベースとアイテムベースの協調型推薦について説明を行う。さらに、それらの手法を組み合わせた Wang らの手法の説明を行う。

協調型推薦では、与えられたユーザの評価値行列を入力として用い、新規に追加されたユーザに対する予測評価値を出力として生成する。このとき、ユーザ数を K 、アイテム数を M とすると、評価値行列 Q は K 行 M 列で表現できる。また、図 1 で示すように、評価値行列の要素は評価値を表しており、ユーザ k のアイテム m に対する評価値は $x_{k,m}$ と表現することができる。

評価値行列の行ベクトルはユーザの特徴ベクトルであるので、ユーザの特徴ベクトルを \mathbf{u} とすると、式(1)で表現できる。

$$Q = [\mathbf{u}_1, \dots, \mathbf{u}_k]^T, \mathbf{u}_k = [x_{k,1}, \dots, x_{k,M}], k = 1, \dots, K \quad (1)$$

同様に、評価値行列の列ベクトルはアイテムの特徴ベクトルであるので、アイテムの特徴ベクトルを \mathbf{i} とすると、式(2)で表現することができる。

$$Q = [\mathbf{i}_1, \dots, \mathbf{i}_M], \mathbf{i}_m = [x_{i,m}, \dots, x_{K,m}]^T, m = 1, \dots, M \quad (2)$$

2.1 ユーザーベースの協調型推薦

ユーザーベースの手法では、図 2 で示すように、評価傾向の類似するユーザー(以下、最近傍と呼ぶ)の評価値を利用して、対象ユーザーの評価値予測を行う。最近傍は調整コサイン類似度を用いて決定される。ユーザー \mathbf{u}_k とユーザー \mathbf{u}_a の調整コサイン類似度を $S_u(\mathbf{u}_k, \mathbf{u}_a)$ とすると、ユーザー k のアイテム m に対する予測評価値 X_{ui} は式(3)となる。

$$X_{ui} = \bar{u}_k + \frac{\sum_a S_u(\mathbf{u}_k, \mathbf{u}_a)(x_{a,m} - \bar{u}_a)}{\sum_a |S_u(\mathbf{u}_k, \mathbf{u}_a)|} \quad (3)$$

ただし、ユーザーベクトルの類似度を算出する際には、共通に評価されているアイテムの評価値のみを用いて算出する。また類似度が 0 の場合は、ユーザー平均値である \bar{u} を返す。

2.2 アイテムベースの協調型推薦

アイテムベースの手法では、2.1 のユーザーベースの手法と同様に、図 3 で示すように、評価傾向の類似するアイテムの評価値を利用する。アイテム $\mathbf{i}_m, \mathbf{i}_b$ の調整コサイン類似度を $S_i(\mathbf{i}_m, \mathbf{i}_b)$ とすると、予測評価値 X_i は式(4)となる。

$$X_i = \bar{u}_k + \frac{\sum_b S_i(\mathbf{i}_m, \mathbf{i}_b)(x_{k,b} - \bar{u}_k)}{\sum_b |S_i(\mathbf{i}_m, \mathbf{i}_b)|} \quad (4)$$

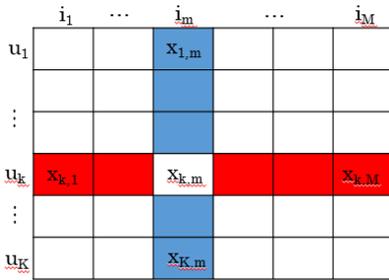


図 1 評価値行列

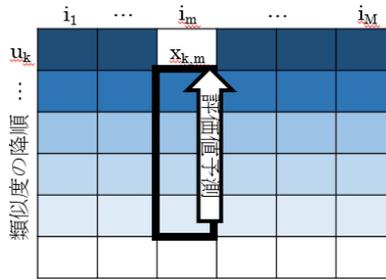


図 2 ユーザーベースの評価値予測

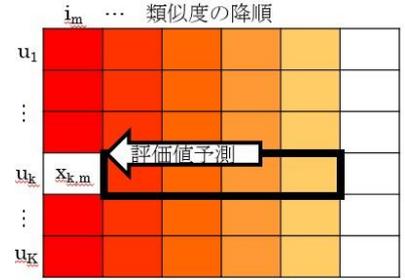


図 3 アイテムベースの評価値予測

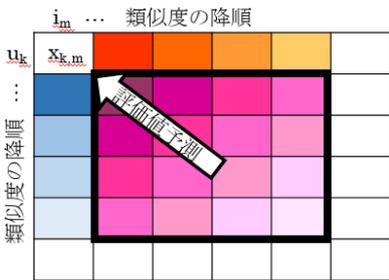


図 4 ユーザー・アイテムベースを用いた評価値予測

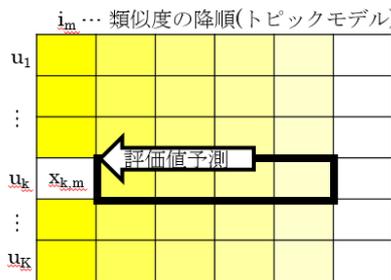


図 5 トピックモデルをアイテムベースに用いた評価値予測



図 6 トピックモデルをユーザー・アイテムベースに用いた評価値予測

2.3 Wang らの手法

Wang らの手法は 2.1 のユーザーベースと 2.2 のアイテムベースと 2.3 の(1)ユーザー・アイテムベースを組み合わせた手法である。

(1) ユーザー・アイテムベースの協調型推薦

ユーザー・アイテムベースを定義することで、図 4 で示すように、類似するユーザーの類似するアイテムに対する評価値を利用することが可能になる。このとき、対象評価値 $x_{k,m}$ との類似度 $S_{ui}(x_{k,m}, x_{a,b})$ は式(5)で定義される。また、類似度 $S_{ui}(x_{k,m}, x_{a,b})$ を用いたユーザー・アイテムベースの予測評価値 X_{ui} は式(6)によって算出される。

$$S_{ui}(x_{k,m}, x_{a,b}) = \frac{S_u(\mathbf{u}_k, \mathbf{u}_a) S_i(\mathbf{i}_m, \mathbf{i}_b)}{\sqrt{S_u(\mathbf{u}_k, \mathbf{u}_a)^2 + S_i(\mathbf{i}_m, \mathbf{i}_b)^2}} \quad (5)$$

$$X_{ui} = \bar{u}_k + \frac{\sum_a \sum_b S_{ui}(x_{k,m}, x_{a,b})(x_{a,b} - \bar{u}_a)}{\sum_a \sum_b S_{ui}(x_{k,m}, x_{a,b})} \quad (6)$$

(2) Wang らの手法の予測評価値

そのうえで、式(3)、式(4)、式(6)のそれぞれの予測評価値 X_{ui}, X_i, X_{ui} に重みを付けることで、最終的な予測評価値 X_{Wang} を式(7)より算出する。

$$X_{Wang} = \lambda(1 - \delta)X_u + (1 - \lambda)(1 - \delta)X_i + \delta X_{ui} \quad (7)$$

なお、 λ と δ は重み付けのパラメータであり、訓練データに対する予測評価値が最適となるように調整される。

3. 提案手法

本稿では、2.3 で述べた Wang らの手法におけるアイテム間の類似度を算出する際に、トピックモデルを利用する手法を提案する。潜在的意味を扱うトピックモデルを利用することで「内容の類似する」アイテムを評価値予測に利用できるようになる。

3.1 トピックモデル

トピックモデルとは、文書を確率的生成モデルによってモデリングする手法である。文書や単語には表層的に現れない潜在的な意味(トピック)を仮定し、文章や単語間の関連性を推定することを実現する。本稿では、代表的なトピックモデルであるLDA(Latent Dirichlet Allocation)を使用した。以下にLDAの定義について説明を行う。

トピック数を T 、文書数を D 、文書 j の単語数を N_j とする。また、トピック t における単語の出現頻度を表すベクトルを Φ_t 、文書 j におけるトピックの出現確率を θ_j と表し、文書 j における i 番目の出現した単語を $w_{j,i}$ 、文書 j における i 番目に出現した単語を $z_{j,i}$ とすると、文書の確率的生成モデルは式(8),(9),(10)により計算される。

$$\Phi_t \sim \text{Dir}(\beta), t = 1, \dots, T \quad (8)$$

$$\theta_j \sim \text{Dir}(\alpha), j = 1, \dots, D \quad (9)$$

$$z_{j,i} \sim \text{Multi}(\theta_j) \text{ and } w_{j,i} \sim \text{Multi}(\Phi_{z_{j,i}}), i = 1, \dots, N_j \quad (10)$$

α と β はディリクレ分布のパラメータであり、コーパス中の文書と確率的生成モデルが最適となるようにパラメータ調整が行われる。

3.2 LDAを用いた特徴ベクトルの作成方法

トピックモデルを使用することで、文書ごとのトピック分布を生成することが可能になる。提案手法では、このトピック分布を利用して、アイテムの特徴ベクトルを生成する。

そのためにまず、それぞれのアイテムに対する全てのレビューテキストの集合を Bag of Words で表現する。この際、非自立語の単語はストップワードとして取り除き、自立語のみによる単語でレビュー集合の Bag of Words 表現を行う。この単語の集合を入力として、LDA に適用する。出力結果は文書ごとのトピックの分布として表現される。ある文書 d のトピック分布 P_d はトピック数を T とすると式(11)によりベクトル表現が可能となる。

$$P_d = [p_1, \dots, p_T], \sum_l p_l = 1 \quad (11)$$

3.3 予測評価値の算出

アイテム i に対するレビューの集合を $d(i)$ とすると、3.2 で定義したトピック分布 $P_{d(i)}$ をアイテムの特徴ベクトルとみなすことで、図 5 で示すように、アイテム間の類似度を用いて評価値を予測することができる。

これより、アイテム \mathbf{i}_m と \mathbf{i}_b の類似度 $S_t(\mathbf{i}_m, \mathbf{i}_b)$ はトピックの分布ベクトル $P_{d(\mathbf{i}_m)}$ と $P_{d(\mathbf{i}_b)}$ のコサイン類似度を用いて算出できる。よって、トピックモデルを用いたアイテムベース(トピックアイテムベースと呼ぶ)の予測評価値 X_{ti} は式(12)により定義される。

$$X_{ti} = \bar{u}_k + \frac{\sum_b S_t(\mathbf{i}_m, \mathbf{i}_b)(x_{k,b} - \bar{u}_k)}{\sum_b S_t(\mathbf{i}_m, \mathbf{i}_b)} \quad (12)$$

また、提案手法は Wang らの手法と同様に、ユーザ・アイテムベースの予測評価値を利用する。このとき、対象評価値との評価類似度 $S_{tui}(x_{k,m}, x_{a,b})$ は式(13)で定義され、トピックモデルを適用したユーザ・アイテムベース(トピックユーザ・アイテムベースと呼ぶ)の予測評価値 X_{tui} は式(14)より算出される。

$$S_{tui}(x_{k,m}, x_{a,b}) = \frac{S_u(u_k, u_a) S_{ti}(\mathbf{i}_m, \mathbf{i}_b)}{\sqrt{S_u(u_k, u_a)^2 + S_{ti}(\mathbf{i}_m, \mathbf{i}_b)^2}} \quad (13)$$

$$X_{tui} = \bar{u}_k + \frac{\sum_a \sum_b S_{tui}(x_{k,m}, x_{a,b})(x_{a,b} - \bar{u}_a)}{\sum_a \sum_b S_{tui}(x_{k,m}, x_{a,b})} \quad (14)$$

そのうえで、式(3)、式(12)、式(14)のそれぞれの予測評価値 X_u, X_{ti}, X_{tui} に重みを付けることで、最終的な提案手法の予測評価値 X_p を式(15)により算出する。

$$X_p = \lambda(1 - \delta)X_u + (1 - \lambda)(1 - \delta)X_{ti} + \delta X_{tui} \quad (15)$$

4. 評価実験

評価値予測のタスクにおいて、3章で述べた提案手法と2章で述べたベースラインとの比較を行う。

4.1 実験データ

本実験では、2つのWebサイトから独自にデータを収集した。まず、評価値行列を構築するために、Amazon から<ユーザ、書籍、評価値>の組でデータを抽出した。このとき、ユーザは、20件以上の書籍を評価しているユーザに限定した。その際、該当するユーザ数は4,010人であった。また、4,010人のユーザが評価した書籍の総数は30,779冊であった。これらのデータを使って、テストデータ、パラメータ調整のためのデータ、そして訓練データを作成した。テストデータは4,010人のユーザからランダムに抽出した100人のユーザによって構成され、 $100 \times 30,779$ の評価値行列となる。また、提案手法のパラメータ調整のために、残りの3,910人のユーザから100人のユーザを別途ランダムに抽出し、 $100 \times 30,779$ の評価値行列を構築する。よって、訓練データは残りの3,810人で構成され、 $3,810 \times 30,779$ の評価値行列で構成される。

一方で、書籍のレビューテキストは読書メーターから集めた。読書メーターとは、レビュー投稿型のWebサイトである。Amazon よりも書籍のレビュー件数が多いため、読書メーターを利用した。

4.2 実験内容

4.1の100人のテストデータに対し、訓練データを用いて評価値予測を行う。ただし、事前にパラメータ調整のための評価値行列を用いて、予測評価値のパラメータ調整を行ったうえで、それぞれの手法における評価値を予測する。

比較する手法は3章で述べた提案手法と2章で述べた3種類のベースライン手法である。3種類のベースライン手法は、ユーザベースとアイテムベースの協調型推薦手法とWangらの手法である。

評価値予測の評価指標はMSE(平均二乗誤差)を用いる。予測対象の評価値を $x_{k,m}$ 、予測評価値を \hat{X} とすると、MSE は式(16)で定義される。

$$\text{MSE} = \frac{(x_{k,m} - \hat{X})^2}{N} \quad (16)$$

N は評価数である。本実験で予測するテストデータの数は100件であるので、今回は $N=100$ となる。

5. 実験結果及び考察

実験結果を表1に示す. 表1より, 全手法の中で提案手法の MSE は 0.507 ポイントで最も小さく, 評価値の予測精度が 1 番高いことが分かった. 次いで Wang らの手法, アイテムベース, そしてユーザベースという実験結果になった. 提案手法がベースラインと比べて, MSE が最大で 0.281 ポイント向上しており, 提案手法の有効性を確認することができた.

さらに, 実験結果を分析するために, 提案手法と Wang らの手法の予測評価値 X_p , X_{Wang} を構成する予測評価値 X_u , X_i , X_{ui} と X_u , X_i , X_{ui} に対する MSE と重みを算出した. その結果を表 2, 3 に示す. 表 2, 3 より, 提案手法におけるアイテムベースの MSE が 0.531 ポイントで最も予測精度が高く, その重みが 0.783 ポイントで提案手法の予測評価値 X_p に大きく影響を与えているので, X_p の予測精度が 1 番高くなったと考えられる.

また, 表 2, 3 より, トピックモデルを適用している X_i , X_{ui} と適用していない X_i , X_{ui} をそれぞれ比較すると, トピックモデルを適用することで, それぞれ 0.098 ポイント, 0.219 ポイントの予測精度の向上を確認することができた. これらの結果より, 評価値予測のタスクにおいて, トピックモデルを用いたアイテム間の類似度は, 評価値のみを用いたアイテム間の類似度よりも有効であることが確認された.

一方で, 表 1 より, Wang ら手法の予測評価値 X_{Wang} の MSE は 0.635 ポイントであり, アイテムベースの評価値 X_i よりも予測精度が 0.203 ポイント低い結果となった. この原因は重み付けの値によるものと考えられる. 表 3 より, 0.832 ポイントで最も MSE の低いユーザ・アイテムベースの予測評価値 X_{ui} が, 0.37 ポイントの重みで X_{Wang} に影響を与えていることがわかる. これが X_{Wang} の予測精度を低下させている原因と考えられる. パラメータ調整用のデータ数を増やすことにより, より適切なパラメータの設定が必要であると考えられる.

これらの結果より, 評価値予測のタスクにおいて, レビューテキストにトピックモデルを適用することで, 書籍の内容を反映させたアイテム間の類似度を用いるハイブリッド型推薦手法の有効性を確認することができた.

表 1 提案手法とベースラインの MSE

手法	MSE
ユーザベース (X_u)	0.788
アイテムベース (X_i)	0.629
Wang らの手法 (X_{Wang})	0.635
提案手法 (X_p)	0.507

表 2 提案手法における各手法の MSE と重み

手法	MSE	重み
ユーザベース (X_u)	0.788	0.097
トピックアイテムベース (X_i)	0.531	0.783
トピックユーザ・アイテムベース (X_{ui})	0.613	0.120

表 3 Wang らの手法における各手法の MSE と重み

手法	MSE	重み
ユーザベース (X_u)	0.788	0.151
アイテムベース (X_i)	0.629	0.479
ユーザ・アイテムベース (X_{ui})	0.832	0.370

6. おわりに

本稿では, レビューテキストを対象にトピックモデルを適用することで書籍の内容を反映させたアイテム間の類似度を用いるハイブリッド型推薦手法を提案した. その結果, 最大で 0.281 ポイントの予測評価値の精度改善につながり, 評価値予測タスクで, レビューテキストを対象としたハイブリッド型推薦手法におけるトピックモデルの有効性を確認することができた.

今後の課題としては, トピックモデルを適用したユーザの特徴ベクトルを構築することが考えられる. 本実験では, ユーザベクトルによる類似度を評価値のみを使って算出していることが予測精度の低下の原因と考えられる. そのため, トピックモデルを用いて, ユーザベクトルを表現することは今後の精度の向上につながると考えられる. また, トピックの相関関係を利用することも考えられる. Apriori [Agrawal 94] のようなルールマイニングアルゴリズムを用いて, トピック間の相関関係を考慮することで予測精度の向上につながると考えられる. 最後に, 他のカテゴリの商品における提案手法の有効性の実証を試みたいと考えている. 提案手法は, 商品のカテゴリに依存する特別な処理をほとんど必要としないので汎用性があると考えられる. そこで, 映画や音楽といった他のカテゴリでも提案手法が有効であるのかを検討する予定である.

参考文献

- [神寫 07] 神寫敏弘: 推薦システムのアルゴリズム(1), 人工知能学会誌, Vol.22, No.6, pp.826-837, 2007.
- [Resnick 94] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews, In Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, ACM, pp.175-186, 1994.
- [Wang 06] Wang, J., De Vries, A. P., and Reinders, M. J.: Unifying user-based and item-based collaborative filtering approaches by similarity fusion, In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, pp.501-508, 2006.
- [服部 12] 服部俊一, 毛中杰, 高間康史: 価値判断に基づくユーザモデリング手法を用いた情報推薦システムの提案とその特性に関する考察, 第 2 回インタラクティブ情報アクセスと可視化マイニング研究会, pp.12-18, 2012.
- [奥 13] 奥健太, 赤池勇磨, 谷口忠大: 推薦システムとしてのビデオバトルの評価, ヒューマンインタフェース学会論文誌, Vol.15, No.1, pp.95-106, 2013.
- [Ganu 09] Ganu, G., Elhadad, N., and Marian, A.: Beyond the Stars Improving Rating Predictions using Review Text Content, WebDB, Vol.9, pp.1-6, 2009.
- [岡田 12] 岡田瑞穂, 藤井敦: レビューテキスト間の類似度を用いた協調フィルタリング, NLP2012 F, Vol.3, pp.1-4, 2012.
- [Blei 03] Blei, D. M., Ng, A. Y., and Jordan, M. I.: Latent Dirichlet Allocation, The Journal of Machine Learning Research, Vol.3, No.5, pp.993-1022, 2003.
- [Agrawal 94] Agrawal, R., Srikant, R.: Fast algorithms for mining association rules, In Proc. 20th Int. Conf. Very Large Data Bases, Vol.1215, pp.487-499, 1994.