

令和 4 年 6 月 7 日現在

機関番号：12612

研究種目：若手研究

研究期間：2018～2021

課題番号：18K18159

研究課題名（和文）回帰分析に基づき推薦理由を説明する情報推薦システムの開発

研究課題名（英文）Development of Recommender System with Explanations Based on Linear Regression

研究代表者

岡本 一志 (Okamoto, Kazushi)

電気通信大学・大学院情報理工学研究科・准教授

研究者番号：10615032

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究では、推薦理由の説明機能を有するモデルベース協調フィルタリングに関する技術開発に取り組んでいる。ユーザとアイテムを同時に1つの線形回帰式で扱い回帰係数の提示により推薦理由の説明が可能なモデルを提案しており、Factorization Machinesと比較して提案法が同程度の推薦精度を有し25倍以上高速に学習できることを検証している。さらに、ユーザとアイテムの交互作用を追加したモデルも提案している。これらの結果から、協調フィルタリングにおいて、線形回帰モデルで高速かつ十分な精度の推薦が行えることを明らかにしている。

研究成果の学術的意義や社会的意義

協調フィルタリングの実現手法にはメモリベース法とモデルベース法の2つがあり、推薦理由の説明に関する多くの研究はメモリベース法を対象としている。一方で、モデルベース法は推薦モデルが学習できれば高速な推薦を行えるが、既存手法では推薦がどの変数に影響を受けたかの把握が難しく、推薦理由の説明には直接活用できない課題がある。本研究の意義は、モデルベース協調フィルタリングにおいて、線形回帰モデルという予測の透明性を有し広く知られた手法を用いることでも高速かつ十分な精度の推薦が行えることを明らかにしたところにある。

研究成果の概要（英文）：This study develops model-based collaborative filtering techniques with explanations for recommended items based on linear regression. The proposed model treats users and items simultaneously in a single linear regression equation, and the reason for recommendation can be explained by presenting regression coefficients. We evaluated that the proposed model has achieved 25 times faster and the same level of recommendation accuracy, compared with Factorization Machines. In addition, a linear regression model with interaction between users and items is proposed. According to the results, we validated that the linear regression models have potentials for fast and accurate enough to make recommendations.

研究分野：データサイエンス

キーワード：情報推薦 推薦理由の説明 協調フィルタリング 線形回帰モデル 図書推薦システム

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

情報推薦システムは、膨大なアイテム(商品や店舗、記事など)の中からユーザにとって価値のあるアイテムの発見を支援するツールであり、その重要な研究テーマのひとつとして推薦理由の説明がある。これまでの研究では、例えば、推薦の受け入れられやすさやシステムへの信頼、迅速な意思決定やシステムの利用満足度などに推薦理由の説明が寄与することが知られている。協調フィルタリングは人の口コミによる推薦をモデル化した推薦技術であり、意外性のある推薦ができる長所に加え、推薦にあたりアイテムの内容情報の収集や推薦ルールの設定などが不要という大規模な数のアイテムを扱う上で有利な特徴がある。

協調フィルタリングの実現手法にはメモリベース法とモデルベース法の2つがあり、推薦理由の説明に関する多くの研究はメモリベース法を対象としている。しかし、メモリベース法は、ユーザやアイテムの近傍探索が必要であり推薦(予測)の計算コストが高く、大規模な数のアイテムを扱にくい。一方で、モデルベース法は推薦モデルが構築できれば高速な推薦を行えるが、Matrix Factorization や非負値行列因子分解といった既存の代表的な手法では推薦がどの変数に影響を受けたかの把握が難しく、推薦理由の説明には直接活用できない課題がある。

### 2. 研究の目的

データ分析一般に、予測に対する説明が必要な場合、回帰分析の適用が第一の選択肢である。本研究では、推薦理由の説明機能を有するモデルベース協調フィルタリングを回帰分析により実現することを目的とする。

### 3. 研究の方法

推薦アイテムの他に提示される情報を推薦の説明という。推薦の説明の役割は、透明性、精密性、信頼性、有効性、説得力、効率性、満足度の7つに分類されており、本研究では、システムの仕組みを説明する透明性について着目する。具体的には、予測スコアがどのように算出されたかをユーザに説明することを透明性とし、線形回帰モデルで得られる回帰係数を活用した協調フィルタリング技術を開発する。その推薦精度、適切な次元削減技術や正則化法、データの欠損への対処法などを複数のデータセットを用いて明らかにする。

### 4. 研究成果

本研究では、ユーザの one-hot ベクトルとアイテムの one-hot ベクトルを結合したベクトルを入力としスコアを予測する線形回帰モデルを構築し、学習により求めた回帰係数を提示する協調フィルタリング技術を提案している。このモデルは、Factorization Machines (FM) と同様にユーザとアイテムを同時に1つの回帰式で扱うものであり、観測したスコアのみを用いるため、未観測のデータの補間といった処理が不要な利点がある。推薦の説明処理では、推定した回帰係数をユーザやアイテムのスコアとして提示することになる。

(1) まず、ユーザとアイテムの交互作用を考慮しないモデルについて検討を行った。提案する線形回帰モデルの学習法として、表1に示す正則化と次元圧縮の適用の組み合わせ9種類を、表2に示す5種類の協調フィルタリング用ベンチマークデータセットを用いたスコア予測実験により、予測誤差(RMSE: Root Mean Squared Error)と計算時間、求めた回帰係数の分布の観点で検証した。また、比較対象のベースラインモデルとして、ユーザが付与したスコアの平均値を出力するモデル(user-avg)とFMを用いた。

表 1: 検討する線形回帰モデルの学習法

推定法	正則化		次元圧縮	
	L1 正則化	L2 正則化	Word2Vec	ランダム射影
LR				
LR-L1				
LR-L2				
LR-W2V				
LR-W2V-L1				
LR-W2V-L2				
LR-RP				
LR-RP-L1				
LR-RP-L2				

表 2：使用したベンチマークデータセットの概要

データセット	ユーザ数 [人]	アイテム数 [件]	総スコア数 [件]	付与可能なスコア
ml-100k	943	1,682	100,000	{1, 2, 3, 4, 5}
ml-1m	6,040	3,706	1,000,209	{1, 2, 3, 4, 5}
webscope	7,642	11,916	221,367	{1, 2, 3, 4, 5}
movietweetings	53,600	30,985	704,211	{0, 1, 2, ..., 10}
epinions	40,164	139,739	664,824	{1, 2, 3, 4, 5}

表 3：データセット・学習法毎の予測誤差 (RMSE) 一覧

	ml-100k	ml-1m	webscope	movietweetings	epinions
LR	1.053	0.956	1.030	1.738	1.100
LR-L1	1.096	1.093	1.151	1.844	1.193
LR-L2	<b>1.025</b>	0.950	1.014	1.662	1.043
LR-W2V	1.061	1.199	1.156	1.763	1.195
LR-W2V-L1	1.060	1.189	1.159	1.769	1.194
LR-W2V-L2	1.059	1.197	1.152	1.764	1.193
LR-RP	1.141	1.120	1.264	1.897	1.202
LR-RP-L1	1.135	1.117	1.257	1.885	1.198
LR-RP-L2	1.139	1.119	1.263	1.898	1.200
ベースライン(user-avg)	1.218	1.140	1.250	1.850	1.328
ベースライン (FM)	1.206	<b>0.934</b>	<b>0.984</b>	<b>1.660</b>	<b>1.036</b>

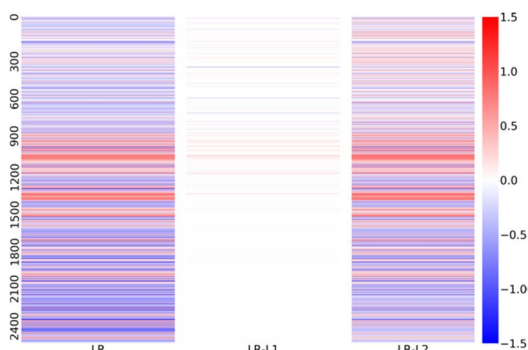


図 1：ml-100k における回帰係数の分布

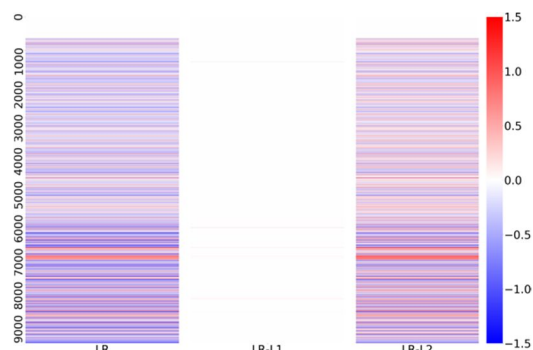


図 2：ml-1m における回帰係数の分布

データセット・学習法毎の RMSE 一覧を表 3 に示す。実験の結果，ベースラインモデルである FM が ml-100k 以外のデータセットで最も RMSE が小さく，L2 正則化を適用した線形回帰モデル (LR-L2) が，検討する表 1 の学習法の中で最も RMSE が小さいことを確認した。ml-100k 以外のデータセットについて，LR-L2 と FM の RMSE の差は 0.007~0.03，user-avg と FM の RMSE の差 0.19~0.292 と比べると 10 分の 1 以下であるため，LR-L2 と FM の RMSE の差は小さく同程度の予測精度を有しているものと考えられる。また，LR-L2 は FM と比べて学習時間が 203~1712 倍高速であることを確認した。図 1 と図 2 はそれぞれ ml-100k と ml-1m データセットにおける LR，LR-L1，LR-L2 の回帰係数の分布を表したものである。このような回帰係数の分析により，L1 正則化を適用しないほうが回帰係数に 0 が出現しづらく，その値にばらつきが出やすくなることから，LR-L1 以外の提案モデルはパーソナライズ化された予測モデルになることを確認している。

(2) 次に，ユーザとアイテムの交互作用を追加したモデルについて検討を行った。このモデルを図書推薦システムに適用し，電気通信大学附属図書館の 3 年分の貸出履歴データを用いた評価実験を実施した。実験結果から，交互作用なしモデル LR-L2 の RMSE が 0.9122，交互作用ありモデル LR-L2 の RMSE が 0.8825，FM の RMSE が 0.9748 であったため，交互作用の追加が推薦精度の向上に寄与することが確認できた。さらに，図 1 や図 2 と同様に交互作用項の可視化も行っており，その結果，回帰係数に 0 が出現しづらく，その値にばらつきが出やすくなる傾向も確認している。このことから，交互作用を追加したモデルもパーソナライズ化された予測モデルになるものと考えられる。

## 5. 主な発表論文等

〔雑誌論文〕 計14件（うち査読付論文 13件 / うち国際共著 0件 / うちオープンアクセス 6件）

1. 著者名 Ryota Yamamoto, Kazushi Okamoto	4. 巻 93
2. 論文標題 Worker Organization System for Collaborative Crowdsourcing	5. 発行年 2021年
3. 雑誌名 Proceedings of 26th International Conference on Production Research (ICPR26)	6. 最初と最後の頁 1-6
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Yoshihito Kosaka, Kazushi Okamoto	4. 巻 M2-3-4
2. 論文標題 Personal Values Modeling with Rough Sets for Collaborative Filtering	5. 発行年 2021年
3. 雑誌名 Proceedings of 7th International Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII2021)	6. 最初と最後の頁 1-5
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 服部 凌典, 柴田 淳司, 岡本 一志	4. 巻 33
2. 論文標題 賃料予測モデルにおける間取り図の影響分析	5. 発行年 2021年
3. 雑誌名 知能と情報	6. 最初と最後の頁 640-650
掲載論文のDOI (デジタルオブジェクト識別子) 10.3156/jssoft.33.2_640	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 中村 幹太, 岡本 一志	4. 巻 33
2. 論文標題 有向研究者ネットワークによる 研究の役割を考慮した共同研究者推薦	5. 発行年 2021年
3. 雑誌名 知能と情報	6. 最初と最後の頁 686-696
掲載論文のDOI (デジタルオブジェクト識別子) 10.3156/jssoft.33.3_686	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kanta Nakamura, Kazushi Okamoto	4. 巻 -
2. 論文標題 Directed Graph-based Researcher Recommendation by Random Walk with Restart and Cosine Similarity	5. 発行年 2020年
3. 雑誌名 Proceedings of Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS2020)	6. 最初と最後の頁 24-28
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SCISISIS50064.2020.9322700	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ryosuke Hattori, Kazushi Okamoto, Atsushi Shibata	4. 巻 -
2. 論文標題 Visualizing the Importance of Floor-Plan Image Features in Rent-Prediction Models	5. 発行年 2020年
3. 雑誌名 Proceedings of Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS2020)	6. 最初と最後の頁 172-174
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SCISISIS50064.2020.9322769	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 藤井 流華, 岡本 一志	4. 巻 35
2. 論文標題 線形回帰による推薦の透明性を有したモデルベース協調フィルタリング	5. 発行年 2020年
3. 雑誌名 人工知能学会論文誌	6. 最初と最後の頁 D-J61_1-10
掲載論文のDOI (デジタルオブジェクト識別子) 10.1527/tjsai.D-J61	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kanta Nakamura, Kazushi Okamoto	4. 巻 F1-4
2. 論文標題 Development of a Collaborator Recommender System Based on Directed Graph Model	5. 発行年 2019年
3. 雑誌名 Proceedings of 20th International Symposium on Advanced Intelligent Systems and International Conference on Biometrics and Kansei Engineering (ISIS2019&ICBAKE2019)	6. 最初と最後の頁 338-345
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ruka Fujii, Kazushi Okamoto	4. 巻 SUN4-B5
2. 論文標題 Loan Record Based Book Recommendation Model with Transparency	5. 発行年 2019年
3. 雑誌名 Proceedings of 6th International Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII2019)	6. 最初と最後の頁 1-6
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ryosuke Hattori, Kazushi Okamoto, Atsushi Shibata	4. 巻 -
2. 論文標題 Rent Prediction Models with Floor Plan Images	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE 8th Global Conference on Consumer Electronics (GCCE2019)	6. 最初と最後の頁 451-452
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE46687.2019.9015208	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Yuta Segawa, Kazuhiko Kawamoto, Kazushi Okamoto	4. 巻 2018:33
2. 論文標題 First-Person Reading Activity Recognition by Deep Learning with Synthetically Generated Images	5. 発行年 2018年
3. 雑誌名 EURASIP Journal on Image and Video Processing	6. 最初と最後の頁 1-13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s13640-018-0272-z	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kazushi Okamoto	4. 巻 4M2-3-5
2. 論文標題 Feature Map Approximation of Triangular Norms with Siamese Neural Network	5. 発行年 2018年
3. 雑誌名 Proceedings of Joint 8th Int. Symp. on Computational Intelligence and Industrial Applications and 12th China-Japan Int. Workshop on Information Technology and Control Applications (ISCIIA-ITCA2018)	6. 最初と最後の頁 1-6
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kazushi Okamoto	4. 巻 22
2. 論文標題 Analysis of Influence Factors for Learning Outcomes with Bayesian Network	5. 発行年 2018年
3. 雑誌名 Journal of Advanced Computational Intelligence and Intelligent Informatics	6. 最初と最後の頁 943-955
掲載論文のDOI (デジタルオブジェクト識別子) 10.20965/jaciii.2018.p0943	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 岡本 一志, 藤井 流華	4. 巻 31
2. 論文標題 協調フィルタリング入門	5. 発行年 2019年
3. 雑誌名 知能と情報	6. 最初と最後の頁 5-9
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計17件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 柳木 佑真, 岡本 一志
2. 発表標題 分散表現を用いたリアルタイム学習型セッションベース推薦システム
3. 学会等名 第35回人工知能学会全国大会
4. 発表年 2021年

1. 発表者名 山本 亮太, 岡本 一志
2. 発表標題 クラウドソーシングにおける協調的な共同作業に対する組織構成システム
3. 学会等名 第35回人工知能学会全国大会
4. 発表年 2021年

1. 発表者名 榑木 佑真, 岡本 一志
2. 発表標題 アイテム分散表現の階層化・集約演算に基づくセッションベース推薦システム
3. 学会等名 第16回Webインテリジェンスとインタラクション研究会
4. 発表年 2020年

1. 発表者名 中村 幹太, 岡本 一志
2. 発表標題 Random Walk with Restartとコサイン類似度に基づく研究者推薦モデル
3. 学会等名 第16回Webインテリジェンスとインタラクション研究会
4. 発表年 2020年

1. 発表者名 小坂 与人, 岡本 一志
2. 発表標題 ラフ集合を用いた価値観モデルベース協調フィルタリング
3. 学会等名 日本知能情報ファジィ学会 ソフトロボティクス研究部会 第18回ポトラック&ワークショップ
4. 発表年 2021年

1. 発表者名 黄 一高, 岡本 一志, 柴田 淳司
2. 発表標題 発売日前のレビューとPU-Learningを用いたスパムレビュー検出
3. 学会等名 統計数理研究所共同研究レポート429: 状態推定法の深化と異分野連携による知の総合 (1)
4. 発表年 2019年



1. 発表者名 藤井 流華, 岡本 一志
2. 発表標題 学生と図書館の交互作用を考慮した透明性を有する図書推薦モデルの検討
3. 学会等名 日本知能情報ファジィ学会 ソフトロボティクス研究部会 第17回ポトラック&ワークショップ
4. 発表年 2019年

1. 発表者名 服部 凌典, 岡本 一志, 柴田 淳司
2. 発表標題 賃料予測モデルにおける間取り図の特徴抽出法の比較
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会 2019
4. 発表年 2019年

1. 発表者名 藤井 流華, 岡本 一志
2. 発表標題 貸出履歴に基づいた透明性を有する図書の推薦モデル
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会 2019
4. 発表年 2019年

1. 発表者名 服部 凌典, 岡本 一志, 柴田 淳司
2. 発表標題 間取り図を用いた賃料予測モデルに関する一検討
3. 学会等名 第33回人工知能学会全国大会
4. 発表年 2019年

1. 発表者名 Richard Pettersson, Kazushi Okamoto
2. 発表標題 Application of Subspace Clustering in Automatic Index Recommendation System
3. 学会等名 第34回ファジィシステムシンポジウム
4. 発表年 2018年

1. 発表者名 藤井 流華, 岡本 一志
2. 発表標題 Autoencoder型協調フィルタリングにおけるモデルパラメータの評価
3. 学会等名 日本知能情報ファジィ学会 ソフトロボティクス研究部会 第16回ポトラック&ワークショップ
4. 発表年 2018年

1. 発表者名 中村 幹太, 岡本 一志
2. 発表標題 科学研究費助成事業データを活用した研究者推薦システムの試作
3. 学会等名 日本知能情報ファジィ学会 ソフトロボティクス研究部会 第16回ポトラック&ワークショップ
4. 発表年 2018年

1. 発表者名 服部 凌典, 岡本 一志, 柴田 淳司
2. 発表標題 家賃予測における間取り画像からの特徴抽出に関する一検討
3. 学会等名 日本知能情報ファジィ学会 ソフトロボティクス研究部会 第16回ポトラック&ワークショップ
4. 発表年 2018年

1. 発表者名 服部 凌典, 岡本 一志, 柴田 淳司
2. 発表標題 間取り画像を考慮した家賃予測モデルの検討
3. 学会等名 IDRユーザフォーラム2018
4. 発表年 2018年

1. 発表者名 藤井 流華, 岡本 一志
2. 発表標題 モデルベース協調フィルタリングにおける推薦の透明性に関する検討
3. 学会等名 第13回Webインテリジェンスとインタラクション研究会
4. 発表年 2018年

1. 発表者名 中村 幹太, 岡本 一志
2. 発表標題 リンク予測に基づく共同研究者推薦システムの試作
3. 学会等名 統計数理研究所共同研究リポート419: 不確実状況下での動的状態推定と知能情報科学の融合 (2)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------