

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

藤井 流華^a 岡本 一志^b

電気通信大学 大学院情報理工学研究科 情報学専攻

a) r.fujii@uec.ac.jp b) kazushi@uec.ac.jp

概要 本研究では、モデルベース協調フィルタリングにおいて、スコア予測だけでなく予測に影響を与えている他のユーザやアイテムの特定も行うことを目指す。具体的には、線形回帰に着目し、得られた偏回帰係数を推薦の透明性の尺度として活用する。偏回帰係数は各説明変数の重みであり、予測に影響を与える他のユーザやアイテムの分析に用いることができると考える。線形回帰に基づくモデルとして、線形回帰モデル、Factorization Machines モデル、ニューラルネットワークによる予測モデルの3種類のモデルを検討する。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実験により、3種類のモデルのスコアの予測精度の検証と予測に影響を与えている他のユーザやアイテムを分析する。実験の結果より、Factorization Machines モデルの予測精度が最も良く、線形回帰モデルとニューラルネットワークによる予測モデルについて、正則化を適用した場合にはスコアの予測精度と推薦の透明性に関する解釈可能性がトレード・オフになっていることを確認している。

1 はじめに

推薦システムとは、アイテム（映画や商品、記事など）の内容や特徴、そのアイテムに対する好みの度合いなどからユーザの嗜好を予測し、登録ユーザやアイテムを推薦するシステムである。協調フィルタリングは推薦システムの実現手法のひとつであり、アイテムに付与されたスコアやユーザの購入・閲覧履歴から、類似ユーザの発見や未評価のスコアの予測などを行う [1]。

推薦システムにおいて推薦理由を提供することを推薦の透明性という [2]。先行研究では、推薦の透明性が推薦の受け入れられやすさやシステムへの信頼性、迅速な意思決定やシステムの利用満足度などに寄与することが示唆されている [3][4][5]。協調フィルタリングのアプローチにはメモリベース法とモデルベース法があり、また、協調フィルタリングにおける推薦の透明性の研究はメモリベース法が主流となっている。メモリベース法は推薦のたびに全ての登録ユーザやアイテムを対象とした近傍探索を行う必要があるため、事前に構築した予測モデルを用いるモデルベース法よりも推薦処理の計算コストが高くなる傾向がある。

本研究では、モデルベース協調フィルタリングによる推薦の透明性の実現を目指す。具体的には、線形回帰に着目し、得られた偏回帰係数を推薦の透明性の尺度として活用する。偏回帰係数は各説明変数の重みであり、予測に影響を与える他のユーザやアイテムの分析に用いることができると考える。線形回帰に基づいたモデルとして、線形回帰モデル、Factorization Machines モデル、ニューラルネットワークによる予測モデルの3種類の

モデルを検討する。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て3種類のモデルの予測精度を検証し、予測に影響を与えている他のユーザやアイテムについて分析する。

2 関連研究

推薦システムにおいて、「この商品はおすすめです。なぜなら…」や「この薬は…のような効果があるのでおすすめです」というような、推薦されたユーザやアイテムの他に提示される情報を推薦の説明という [6]。特に、推薦理由を提示することを推薦の透明性という。これまでに、推薦の透明性が推薦システムのユーザに対してどのような影響を与えるのかについて、検証が行われている [3][4][5]。

Sinha ら [3] は、12人の被験者に5つの音楽推薦システムを利用して、「推薦されたアイテムを好むか」「推薦は信頼できるか」「推薦に透明性があったか」の3つの質問に答えてもらう実験を行っている。この実験では、推薦に透明性があるほうがユーザは推薦されたアイテムを好み、推薦システムを信頼することが示されている。また、Gedikli ら [4] は、推薦アイテムの予測と同時に、類似ユーザの評価のヒストグラムや全ユーザのスコアの平均などをランダムにひとつ提示する推薦システムをユーザに使用してもらい、ユーザのシステムに対する利用満足度を評価する実験を行っている。実験より、これらの推薦の透明性に関連した情報を提示したほうが、そうでない場合に比べてユーザの満足度が高くなることが明らかになっている。Herlocker ら [5] は、MovieLens システムにおいて21種類の説明インターフェースを実装し、そのインターフェースがユーザに与える影響につ

いて実験を行っている。実験の結果より、類似ユーザが推薦アイテムに付与したスコアのヒストグラムや他に高く評価されている映画との類似度、好きな俳優などの映画の内容に関する根拠を述べた説明インターフェースを用いた場合に、ユーザのシステムに対する評価が高くなることがわかっている。

3 線形回帰に基づいた予測モデル

本研究ではモデルベース法により推薦の透明性を有した協調フィルタリングの実現を目指す。モデルベース法による協調フィルタリングのアプローチとして、行列分解や教師あり学習などが挙げられる。本研究では線形回帰に着目する。線形回帰で推定する偏回帰係数は説明変数に対する重みであると考えられるため、目的変数にどの説明変数が影響を与えているのかの分析に活用することができる。これによりモデルベース法での推薦の透明性の実現が可能になると考える。

m 人のユーザが n 個のアイテムについて評価しているデータセットを考える。本研究では、予測したいスコア（目的変数） y と、ユーザの one-hot ベクトルとアイテムの one-hot ベクトルを連結したユーザ・アイテムベクトル（説明変数） $\mathbf{x} \in \mathbb{R}^{m+n}$ を用いて、線形回帰による予測モデルを

$$\hat{y} = \alpha_0 + \boldsymbol{\alpha}^T \mathbf{x} = \alpha_0 + \sum_{i=1}^{m+n} \alpha_i x_i \quad (1)$$

で定義する。ここで、 $\boldsymbol{\alpha} \in \mathbb{R}^{m+n}$ は偏回帰係数を意味している。式 (1) より、目的変数 y は偏回帰係数 $\boldsymbol{\alpha}$ を重みとしたユーザ・アイテムベクトル \mathbf{x} の内積となっているため、偏回帰係数 $\boldsymbol{\alpha}$ の値から予測に影響を与えているアイテムを分析することができる。本研究では、偏回帰係数を推薦の透明性の尺度として用いる。

3.1 正則化を適用した線形回帰モデル

上記の定義では、高次元の入力データを扱うこととなるため、次元数に対して学習データ数が不足し過学習が発生する可能性がある。そこで、偏回帰係数の推定に正則化の適用を検討する。正則化には、L1 正則化 [7] と L2 正則化 [8] の 2 種類を適用する。L1 正則化の損失関数 $E_L(\boldsymbol{\alpha})$ と L2 正則化の損失関数 $E_R(\boldsymbol{\alpha})$ は、式 (1) のスコアの予測値 \hat{y} を用いて

$$E_L(\boldsymbol{\alpha}) = (y - \hat{y})^2 + \lambda \sum_{i=0}^{m+n} |\alpha_i| \quad (2)$$

$$E_R(\boldsymbol{\alpha}) = (y - \hat{y})^2 + \lambda \sum_{i=0}^{m+n} |\alpha_i|^2 \quad (3)$$

と定義される。このとき、 y はスコアの真値を、 λ は正則化項の寄与率を調節するパラメータを意味する。

これら 2 つの正則化の手法では、得られる偏回帰係数の傾向が異なる。L1 正則化は一部の偏回帰係数に 0 が

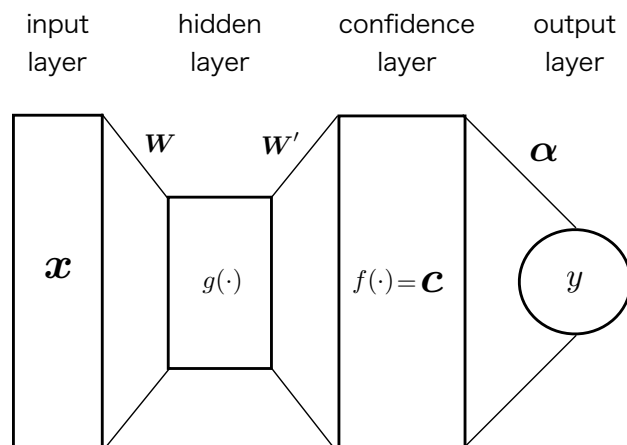


図 1 ニューラルネットワークによる予測モデル

含まれるような推定となるため、変数選択を行うことに相当する。L2 正則化は偏回帰係数に 0 ができるだけ含まれないような推定となる。本研究では偏回帰係数を推薦の透明性の尺度としており、変数選択を行うことのできる L1 正則化は考慮すべき偏回帰係数の数を減らすことができるため、L1 正則化が適用できることが望ましいと考える。

3.2 Factorization Machines モデル

Factorization Machines [9] は、組み合わせ特徴量を扱う教師あり学習モデルである。Factorization Machines の予測モデルは

$$\hat{y} = \alpha_0 + \sum_{i=1}^{m+n} \alpha_i x_i + \sum_{i=1}^{m+n} \sum_{j=i+1}^{m+n} \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j \quad (4)$$

で定義される。ここで、 $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$ は $\mathbf{v}_i, \mathbf{v}_j \in \mathbb{R}^d$ とし、組み合わせ特徴量に対する重みの内積を表す。

式 (4) は、

$$\begin{aligned} \hat{y} &= \alpha_0 + \sum_{i=1}^{m+n} \alpha_i x_i + \sum_{i=1}^{m+n} \sum_{j=i+1}^{m+n} \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j \\ &= \alpha_0 + \sum_{i=1}^{m+n} \left(\alpha_i x_i + \sum_{j=i+1}^{m+n} \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j \right) \\ &= \alpha_0 + \sum_{i=1}^{m+n} \left(\alpha_i + \sum_{j=i+1}^{m+n} \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_j \right) x_i \end{aligned} \quad (5)$$

と変形することができ、 $\left(\alpha_i + \sum_{j=i+1}^{m+n} \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_j \right) = \alpha'_i$ とすることで式 (1) と同等とみなすことができる。よって、Factorization Machines モデルでも偏回帰係数を用いた推薦の透明性についての分析を行うことができると考える。

3.3 ニューラルネットワークによる予測モデル

本研究では、ユーザ・アイテムベクトルに非線形変換を適用することによってスコアの予測精度が向上するこ

表 1 正則化適用時の予測精度 (MAE)

		λ					
モデル	正則化	0.001	0.01	0.1	1	10	100
LR	L1	0.856	0.948	0.950	0.950	0.950	0.998
	L2	0.810	0.908	0.945	0.950	0.950	4.75×10^{30}
NN	L1	0.766	0.772	0.950	0.949	0.949	0.915
	L2	0.771	0.774	0.803	0.929	0.952	39.4

とを期待し、ニューラルネットワークを用いた予測モデルを提案する。モデルの概要を図 1 に示す。

3.3.1 予測モデルと学習処理

提案するニューラルネットワークによる予測モデルは、入力層 (input layer)、隠れ層 (hidden layer)、confidence 層、出力層 (output layer) の 4 層で構成される。図 1 において、 W, W' は隠れ層と confidence 層の線形結合のパラメータを、 α は偏回帰係数を、 $g(\cdot), f(\cdot)$ は活性化関数を意味している。本研究では、 $g(\cdot)$ に relu を、 $f(\cdot)$ に softmax を適用する。出力層には活性化関数を適用しないため、confidence 層の出力を c とすると、出力層は

$$\hat{y} = \alpha_0 + \alpha c \quad (6)$$

となり、単純な線形回帰に相当する。

式 (1) の線形回帰モデルを基準とすると、式 (5) の Factorization Machines モデルは偏回帰係数を工夫して推定する問題に、本研究で提案するニューラルネットワークによる予測モデルは入力データを工夫して偏回帰係数を推定する問題であると考えられることができる。

提案するニューラルネットワークにおける予測モデルの学習では、まず、autoencoder[10] を用いて入力層と confidence 層の一対一対応をとる。推薦の透明性の分析において、予測に影響を与えているユーザやアイテムを明らかにするため、入力層と confidence 層の対応をとる必要がある。autoencoder とは、ニューラルネットワークにおいて、入力信号と出力信号が同じになるように学習させる手法である。本研究では、confidence 層の出力ベクトルが入力ベクトルと同じになるように学習させることで、入力層と confidence 層の一対一対応をとる事前学習を行う。autoencoder による事前学習の後、図 1 のように出力層を追加し、ファインチューニングを行う。また、線形回帰モデルとの比較を行うため、confidence 層から出力層にかけて L1 正則化および L2 正則化を適用した場合についても検証する。

3.3.2 結合パラメータの初期値

結合パラメータ W, W' と、偏回帰係数 α の初期値は超一様分布列 (low-discrepancy sequence) [12] のひとつである Halton 数列 [13] を用いて決定する。Halton 数列は他の乱数生成手法と比較すると均等にばらつく性質が

あり、予備的実験から、ニューラルネットワークの安定した学習が期待できることを確認している。

4 モデルの予測精度と偏回帰係数の分析に関する評価実験

3 種類の予測モデルのスコアの予測精度を明らかにするため、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評価実験を行う。

4.1 使用データセットと予測精度の評価方法

本研究では、協調フィルタリング用ベンチマークデータセットである Movie Lens 100k を使用する。このデータセットは、ユーザが映画に 1 から 5 の 5 段階評価で付与したスコアを集計したものであり、ユーザ数 943、アイテム数 1,682、総スコア数 100,000 である。

予測モデルの汎化性能を検証するために、交差検証法を用いる。交差検証法は、データを k 分割し、そのうちの 1 個をテスト用データに、残りの $k-1$ 個を学習用データに使用する手法である。 k 回検証を行った結果を平均して予測精度を算出する。本研究では $k = 10$ とする。

予測精度の評価として平均絶対誤差 (MAE: Mean Absolute Error)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

を使用する。ここで、 n はデータ数を、 \hat{y}_i は予測値を、 y_i は真値を意味している。MAE は予測値が真値からどれだけ離れているかを表す指標であるため、MAE の値が小さいほど予測精度は良いといえる。

4.2 実験環境

本研究では、3 種類の予測モデルとして、L1 正則化および L2 正則化を適用した線形回帰モデルと Factorization Machines モデル、ニューラルネットワークを用いた予測モデルを対象とする。

Factorization Machines モデルは libFM[15] を用いて実装し、ハイパーパラメータに関しては、予備的実験により得られた知見から、圧縮次元数 $d = 200$ 、学習の反復回数 1,000 を適用する。その他のハイパーパラメータはデフォルトの値を用いる。L1 正則化および L2 正則化を適用した線形回帰モデルとニューラルネットワークによる予測モデルの実装には Python の Keras ライブラリ

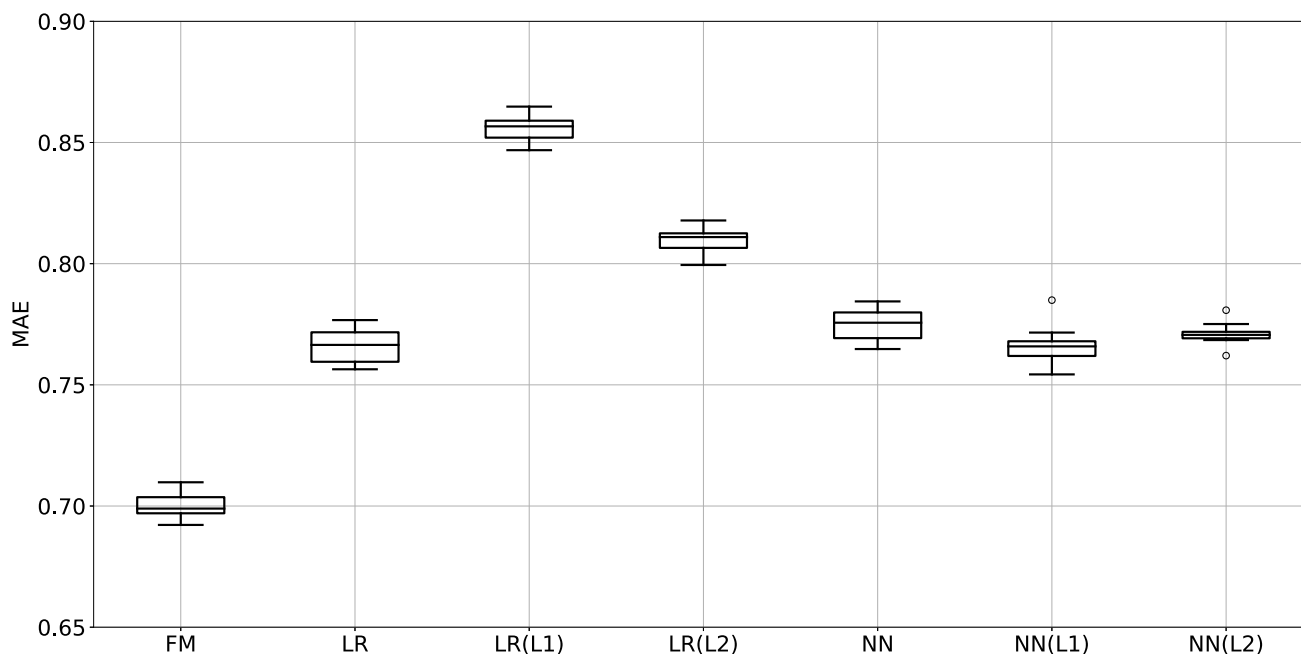


図 2 3 種類のモデルにおける予測精度 (MAE)

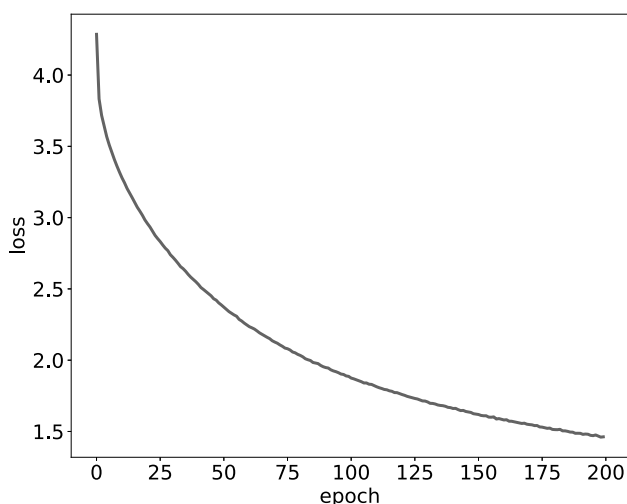


図 3 損失関数の値 (loss) の推移

[16] を用いる。ニューラルネットワークによる予測モデルの隠れ層の次元数は、Factorization Machines モデルと合わせるために $d = 200$ を適用する。また、L1 正則化および L2 正則化について、正則化の寄与率を意味するパラメータ λ は適切な値を実験より明らかにする。線形回帰モデルとニューラルネットワークによる予測モデルの epoch には 200 を適用し、最適化アルゴリズムには確率的勾配降下法を用いる。さらに、過学習の発生を抑えるためにすべての層に dropout を適用し、損失関数には平均二乗誤差を使用する。

5 評価実験の結果と考察

5.1 適切な正則化の寄与率 λ

線形回帰モデルとニューラルネットワークによる予測モデルに適用する L1 正則化および L2 正則化について、正則化項の寄与率を意味するパラメータ λ の適切な値を明らかにする。本研究では、 $\lambda = \{0.001, 0.01, 0.1, 1, 10, 100\}$ の 6 種類について検証を行う。パラメータ λ を変えたときの線形回帰モデル (LR) とニューラルネットワークによる予測モデル (NN) の予測精度を表 1 に示す。

表 1 より、線形回帰モデルとニューラルネットワークによる予測モデルのどちらの場合も、L1 正則化と L2 正則化で $\lambda = 0.001$ のときに最も MAE が小さくなっている。以後の実験では、それぞれの適切な λ の値を用いる。

5.2 予測精度の検証

線形回帰モデル (LR)、Factorization Machines モデル (FM)、ニューラルネットワークによる予測モデル (NN) の 3 種類のモデルでの予測精度を箱ひげ図にしたものを図 2 に示す。図 2 の縦軸は MAE を表している。

図 2 より、最も予測精度が良いのは Factorization Machines モデルであることがわかる。線形回帰モデルは、L1 正則化および L2 正則化を適用しない場合のほうが正則化を適用した場合よりも MAE が小さくなっている。ニューラルネットワークによる予測モデルは、L1 正則化を適用した場合に最も予測精度が良くなっている。正則化を適用しない線形回帰モデルと L1 正則化を適用したニューラルネットワークによる予測モデルは、MAE の中央値で 0.001 の差であった。5 段階評価に対して 0.001 の差であるため、予測性能に大きな違いはないと考える。推

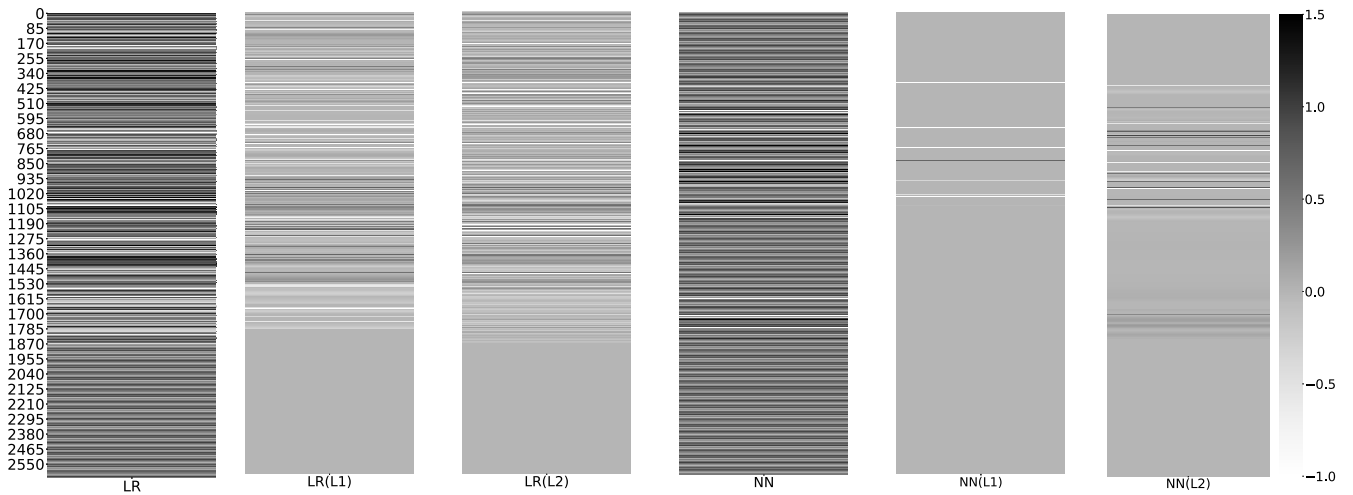


図4 偏回帰係数のヒートマップ

薦の透明性に関する解釈を行う際、L1 正則化を適用することで考慮すべき偏回帰係数の数を少なくすることができるため、線形回帰モデルよりも L1 正則化を適用したニューラルネットワークによる予測モデルのほうが適していると考ええる。

5.3 ニューラルネットワークモデルにおける

入力層と confidence 層の対応関係

本研究では、カルバック・ライブラー情報量 (Kullback-Leibler divergence) [11] により、事前学習で confidence 層の出力ベクトルが入力ベクトルをどの程度再現しているか、また、ファインチューニングによってどの程度 confidence 層の出力が変化しているかの検証を行う。

2つの確率分布 $P(x)$ と $Q(x)$ のカルバック・ライブラー情報量 D は

$$D(P(x)||Q(x)) = \sum_i p(i) \log \frac{p(i)}{q(i)} \quad (8)$$

で定義される。カルバック・ライブラー情報量の値が小さいほど、2つの確率分布は類似していると考えられることができる。式 (8) より自然対数の計算と除算が発生するため、カルバック・ライブラー情報量の計算においては、 $P(i)$ と $Q(i)$ に 0 が含まれないよう x を ε だけかさ上げした後、 x の総和が 1 となるよう正規化している。本研究では、 $\varepsilon = 0.00001$ を用いる。

事前学習において、学習データに対する epoch と損失関数の値 (loss) の関係を図3に示す。図3より、学習が進むにつれ loss も小さくなっており、学習が終わる頃の loss は 1.5 となっている。また、事前学習に用いた教師信号と事前学習後の confidence 層の出力、ファインチューニング後の confidence 層の出力について、学習データとテストデータにおけるそれぞれのカルバック・ライブラー情報量は次のようになっている。

- 教師信号 $P(x)$ と autoencoder 後の confidence 層の出力 $Q(x)$ のカルバック・ライブラー情報量
 - 学習データを入力: 3.077
 - テストデータを入力: 3.062
- autoencoder 後の confidence 層の出力 $P(x)$ とファインチューニング後の confidence 層の出力 $Q(x)$ のカルバック・ライブラー情報量
 - 学習データを入力: 0.343
 - テストデータを入力: 0.361

事前学習に用いた教師信号と事前学習後の confidence 層の出力は、カルバック・ライブラー情報量が 3 であることから、autoencoder による事前学習はうまくできていない可能性がある。しかし、事前学習後の confidence 層の出力とファインチューニング後の confidence 層の出力はカルバック・ライブラー情報量が 0.3 であるため、ファインチューニングによる confidence 層の変化は少ない傾向にある。

5.4 偏回帰係数の分析

線形回帰モデルとニューラルネットワークによる予測モデルの偏回帰係数の値をヒートマップにしたものを図4に示す。図4について、縦軸はユーザ・アイテムベクトル $x \in \mathbb{R}^{m+n}$ を意味しており、0 から 942 までがユーザを、943 から 2624 までがアイテムを表している。

図4より、どちらのモデルも正則化を適用していないモデルの偏回帰係数は 0 でない値が多く含まれていることがわかる。線形回帰モデルでは、L1 正則化および L2 正則化を適用したモデルの偏回帰係数が 0 になる割合が多く、特に L1 正則化を適用したモデルは考慮すべき偏回帰係数の数が少ないため、推薦の透明性に関する

解釈を行いやすい。ニューラルネットワークによる予測モデルでは、L1 正則化および L2 正則化を適用したモデルの偏回帰係数がほとんど 0 になっている。スコアの予測が入力データに依らず、バイアス項に大きく影響されているためと考える。推薦の透明性に関する解釈が行いやすい L1 正則化では、図 2 より、スコアの予測精度はニューラルネットワークによる予測モデルが高くなっている。その一方で、ニューラルネットワークによる予測モデルの偏回帰係数はほとんどが 0 になっており、推薦の透明性に関する解釈を行う場合は線形回帰モデルのほうが適していると考えられる。線形回帰モデルとニューラルネットワークによる予測モデルでは、スコアの予測精度と推薦の透明性に関する解釈可能性がトレード・オフの関係になっているといえる。

6 おわりに

本研究では、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の実現を目的とし、線形回帰の偏回帰係数を推薦の透明性の尺度として活用する手法について検討している。線形回帰モデル、Factorization Machines モデル、ニューラルネットワークによる予測モデルの 3 種類について、予測精度の検証と予測に影響を与える他のユーザやアイテムの分析を行っている。

実験の結果、スコアの予測精度は Factorization Machines モデルが最もよいことを明らかにしている。線形回帰モデルは、正則化を適用しないモデルのほうが L1 正則化および L2 正則化を適用した場合に比べて予測精度が良くなっており、ニューラルネットワークによる予測モデルは、L1 正則化を適用したモデルが最も予測精度が良くなっている。ニューラルネットワークによる予測モデルでの推薦の透明性に関する分析は、入力層と confidence 層が一对一対応していることが望ましく、カルバック・ライブラー情報量を計算したところ、事前学習後の confidence 層の出力とファインチューニング後の confidence 層の出力は類似している可能性があることを確認している。偏回帰係数の分析では、L1 正則化を適用した線形回帰モデルの偏回帰係数は考慮すべき数が少ないため予測に影響を与えているユーザやアイテムの特定には適していると考えられる。しかし、スコアの予測精度は L1 正則化を適用したニューラルネットワークによる予測モデルのほうが良く、スコアの予測精度と推薦の透明性の解釈可能性はトレード・オフになっていると考えられる。

今後の課題として、データセットを増やすことを考えている。また、Factorization Machines モデルの推薦の透明性の分析に関する具体的な検証も行っていきたい。

参考文献

- [1] Su, X. and Khoshgoftaar, T. M.: A Survey of Collaborative Filtering Techniques, *Advances in artificial intelligence*, 2009.
- [2] Tintarev, N.: Explaining Recommendations, *Proceedings of international conference on user modeling*, pp. 470–474, 2007.
- [3] Sinha, R. and Swearingen, K.: The Role of Transparency in Recommender Systems, *Proceedings of SIGCHI conference on human factors in computing systems*, pp. 830–831, 2002.
- [4] Gedikli, F., Jannach, D., and MouzhiGe: How Should I Explain? A Comparison of Different Explanation Types for Recommender Systems, *International journal of human-computer studies*, Vol. 72, No. 4, pp. 367–382, 2014.
- [5] Herlocker, J.L., Konstan, J.A., and John, R.: Explaining collaborative filtering recommendations, *Proceedings of the 2000 ACM conference on computer supported cooperative work*, pp.241–250, 2000.
- [6] Dietmar, J., Markus, Z., Alexander, F., Gerhard, F., 田中克己 (監修), 角谷和俊 (監修): 情報推薦システム入門 理論と実践, 共立出版, 2012.
- [7] Tibshirani, R.: Regression Shrinkage and Selection Via the Lasso, *Journal of the royal statistical society*, Vol. 58, pp. 267–288, 1994.
- [8] Hoerl, A.E. and Kennard, R.W.: Ridge Regression: Biased Estimation for Nonorthogonal Problems, *American statistical association*, Vol. 12, No. 1, pp. 55–67, 1970.
- [9] Steffen, R.: Factorization Machines, *Proceedings of the 2010 IEEE international conference on data mining*, pp.995–1000, 2010.
- [10] Ouyang, Y., Liu, W., Rong, W., and Xiong, Z.: Autoencoder-Based Collaborative Filtering, *International conference on neural information processing*, pp. 284–291, 2014.
- [11] Kullback, S. and Leibler, R.A.: On information and sufficiency, *Annals of mathematical statistics*, Vol. 22, No. 1, pp. 79–86, 1951.
- [12] 手塚 集: 擬似乱数と超一様分布列: 最近の話題 (< 小特集 > モンテカルロシミュレーション), *日本シミュレーション学会*, Vol. 22, No.4, pp. 227–231, 2003.
- [13] Halton, J.H.: Algorithm 247: Radical-inverse quasi-random point sequence, *Magazine communications of the ACM*, Vol. 7, No. 12, pp. 701–702, 1964.
- [14] Maxwell, F.H. and Joseph A.K.: The MovieLens Datasets: History and Context, *ACM transactions on interactive intelligent systems (TiiS)*, Vol. 5, No. 4, pp. 19, DOI=<http://dx.doi.org/10.1145/2827872>, 2015.
- [15] Steffen R.: Factorization Machines with libFM, *ACM transactions on intelligent systems and technology*, Vol. 3, No. 3, pp. 1–22, 2012.
- [16] Chollet, F. and others: Keras, <https://keras.io>, 2015.

謝辞

本研究は JSPS 科研費 JP18K18159 の助成を受けたものです。また、本研究は東京大学情報基盤センターの Reedbush スーパーコンピュータシステムを利用し実施したものです。