

積層自己符号化器による 類似性の変化に着目した協調フィルタリングの提案

Stacked Denoising Autoencoder-based Deep Collaborative Filtering using the change of similarity

鈴木陽介 尾崎知伸*
Yosuke Suzuki Tomonobu Ozaki

日本大学大学院 総合基礎科学研究科
Graduate School of Intergrated Basic Sciences, Nihon University

Abstract: Recommendation systems based on deep learning technology pay attention recently. In this paper, we propose a collaborative filtering based recommendation algorithm that effectively utilize the difference of similarities among users derived from different layers in stacked denoising autoencoders. Since different layers in stacked autoencoder represent relationships among items with rating at different levels of abstraction, we can expect to make recommendations more novel, various and serendipitous, compared with ordinal collaborative filtering using single similarity.

1 はじめに

機械学習界隈における深層学習の隆盛は周知の事実である。深層学習はデータの高次の表現を自動的に学習することが可能であり [4], さらに画像認識のタスクにおいて, 従来の手法に比べ認識率が大幅に向上したことなどが注目を集めた要因に挙げられる。現在は学習アルゴリズムの理論研究のみならず, 応用研究も盛んに行われている。

深層学習の推薦システムへの応用もその一つである [1, 2, 3, 5, 7, 8]. 推薦システムとは, 大量のアイテムから被推薦者の嗜好や興味に合致したアイテムを提示・推薦するシステムのことであり, ユーザが有用なアイテムを探し出す行動を支援する効果がある。推薦システム研究における代表的な国際会議の一つである RecSys において, 深層学習を用いた推薦システムを扱うワークショップ "DLRS" (workshop on Deep Learning for Recommender Systems)¹ が今年から開催されていることから, 深層学習への期待の高さが窺える。

推薦システムへの深層学習の適用例として自己符号化器を用いた研究がある [2, 3, 5, 7, 8]. 自己符号化器とは, 教師無しの深層学習モデルの一つであり, 入力データの高次の表現を得る事ができると言われている。

例として, ユーザが各アイテムに対して評価値を付与している嗜好情報を入力とした自己符号化器を構築

した場合を考える。入力データ (評価値ベクトル) がアイテム個々に対する絶対的な評価を表しているが, 自己符号化器を用いるとその高次の表現, 例えば, どのような特徴を持ったアイテムを好むか, といった高次の嗜好情報を得ていると考えられる。

このような高次の嗜好情報は, ユーザ間の類似度を計算する際に利点が存在する。基本的な推薦アルゴリズムの一つである協調フィルタリング [6] では, 嗜好に関するユーザ間類似度を計算する必要がある。その際, 従来の類似度計算では, お互いのユーザが共に評価値を付けているアイテムだけを考慮したピアソン相関によって嗜好類似度を計算していた。従って, アイテム個々に対する嗜好の類似度を計算しており, 好みのジャンルや傾向等を考慮していない。以降, この類似度を“表面的な類似度”とよぶ。一方, 中間層の値は高次の嗜好を表現するため, 共通に評価をしているアイテムが少ない場合でも類似度計算が高精度に行える可能性がある。ひいては, 協調フィルタリングの推薦精度向上に繋がるものと考えられる。以降, 中間層の値を利用して計算された類似度を“高次の類似度”とよぶ。

また, ピアソン相関による類似度と中間層における類似度の変化を利用した推薦アルゴリズムも考えられる。本研究では, ピアソン相関では類似度が低いが, 中間層における類似度が高いユーザを利用した協調フィルタリングを提案する。このようなユーザは, 直感的には, 被推薦者と同じアイテムを評価していないためピアソン相関が低いが, 実際には似たアイテムを評価しているために, 中間層における類似度が向上してい

*連絡先: 日本大学 文理学部 情報科学科
〒156-8550 東京都世田谷区桜上水 3-25-40
tozaki@chs.nihon-u.ac.jp

¹<http://dlrs-workshop.org/>

ると考えることができる。つまり、被推薦者と普段の活動コミュニティは異なるが、潜在的な嗜好が類似したユーザ（以降、潜在的類似ユーザ）であると解釈できる。このようなユーザを利用した結果、嗜好は合致しているながらも普段利用するアイテムとは非類似したアイテムが推薦される可能性がある。

2 関連研究

ユーザの評価値データに対して深層学習を適用した推薦システムの研究はいくつか提案されている。その多くは本研究で利用する自己符号化器を用いたものである。それら研究で問題としていることは、推薦システム特有の未評価値を含むデータに対して、如何に自己符号化器を適用するかという点である。

Sedhain ら [2] や, Strub ら [3] の提案手法は次章で詳しく解説をするが, 簡単には目的の構造の自己符号化器を学習する際に, データ毎に未評価値と対応するユニットが存在しないかの様に学習を行い, 予測時には完全な構造の自己符号化器を利用することで, 未評価値への対応をしている。本研究ではこれらの提案手法を参考にしたモデルを利用する。その他に, 川上ら [8] や, 田中ら [5], 鈴木ら [7] は未評価値を何らかの値(平均評価値など)に変換することで未評価値に対応している。

3 積層自己符号化器を用いた協調フィルタリング

本研究では, 自己符号化器の中間層を類似度計算に利用した推薦システムと, その類似度の変化に着目した推薦システムを提案し, 提案手法の実験的評価を行う。

まず, 本研究で利用するデータを以下のように与える。ユーザ集合を $U = \{u_1, \dots, u_{|U|}\}$, アイテム集合を $I = \{i_1, \dots, i_{|I|}\}$ とする。評価値の定義域を $\mathcal{R} \in \mathbb{R} \cup \{NA\}$ とし, ユーザ u の評価値ベクトルを $\mathbf{r}_u = (r_{u1}, \dots, r_{u|I|}) \in \mathcal{R}^{|I|}$ で表す。ここで, NA は未評価値を表す。

3.1 積層自己符号化器

自己符号化器とは, 深層学習の教師無し学習モデルの一つである。単層の自己符号化器は図1のような入力層, 中間層, 出力層からなる構造をもつ。まずは簡単のため, あるユーザの未評価値を“含まない”評価値ベクトル $\hat{\mathbf{r}}_u \in \mathbb{R}$ を入力とする自己符号化器を考えると, その出力は,

$$h(\hat{\mathbf{r}}_u; \theta) = f(\mathbf{W} \cdot g(\mathbf{V} \hat{\mathbf{r}}_u + \boldsymbol{\mu}) + \mathbf{b})$$

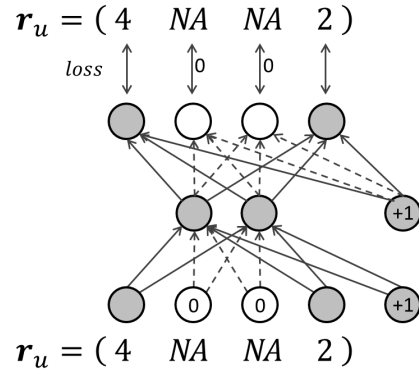


図 1: 単層自己符号化器

のように表せる。ここで, f と g は活性化関数と呼ばれる任意の非線形関数であり, \mathbf{V} と \mathbf{W} は重み, \mathbf{b} と $\boldsymbol{\mu}$ はバイアス項で, $\theta = \{\mathbf{W}, \mathbf{V}, \mathbf{b}, \boldsymbol{\mu}\}$ は重みとバイアス項をまとめた変数である。自己符号化器の学習は, 入力データ $\hat{\mathbf{r}}_u$ に対して誤差関数

$$E(\theta) = \sum_{u \in U} \|\hat{\mathbf{r}}_u - h(\hat{\mathbf{r}}_u; \theta)\|^2$$

が小さくなるように確率的勾配降下法を用いて θ を更新していく。

さらに, モデルの頑健性を高めるために入力データに何らかのノイズを加えて自己符号化器に入力し, 出力値とノイズを入れる前の入力データとで誤差計算を行い学習をする, デノイズ自己符号化器を採用した。具体的には, 入力データ $\hat{\mathbf{r}}_u$ に何かしらのノイズを加えたデータを $\tilde{\mathbf{r}}_u$ とすると, 誤差関数は以下のようになる。

$$E(\theta) = \sum_{u \in U} \|\tilde{\mathbf{r}}_u - h(\tilde{\mathbf{r}}_u; \theta)\|^2$$

上記の自己符号化器は未評価値を含まないデータ $\hat{\mathbf{r}}_u$ を入力とした自己符号化器であるが, 実際のデータは未評価値を含む。本研究では, Sedhain ら [2] や Strub ら [3] が提案した, 未評価値を含む評価値ベクトル \mathbf{r}_u に対応した自己符号化器を参考にしたモデルを利用した。アイデアは, 各データに対して, 入力層, 出力層共に未評価値に対応するユニットが存在しないかの様に学習をする。テスト時は, 出力層の全てのユニットを利用して評価値を予測する。このようにすることで, 学習時にはユーザ毎に不完全な構造のモデルを学習し, 予測時には全てのユニットを利用する完全な構造のモデルにより出力することとなり, 他のユーザの嗜好を反映した評価値ベクトルを出力することができる。具体的には, 学習時に未評価値を 0 に変換して順伝播させ, 未評価値と対応する出力層のユニットの誤差を 0 にしてパラメータの更新を行うことで, 上

記のアイデアを実現する。また、予測時も同じく、未評価値を0に変換して順伝播する。

以上の手順を応用すると簡単に積層自己符号化器を構築できる。図2のように、まず単層の自己符号化器を学習し、その中間層の出力を入力とした、単層の自己符号化器を学習することで多層化することができる。

各層の出力を表すため、以下に定式化する。入力層を1層目、それ以降を2, 3, ... 層目と表す。ユーザ u の評価値を入力としたときの1層目の出力を z_u^1 とする。つまり、 $z_u^1 = r_u$ であり、 $z_u^{l+1} = f^l(\mathbf{W}^l \cdot z_u^l + \mathbf{b}^l)$ である。ただし、 f^l , \mathbf{W}^l , \mathbf{b}^l はそれぞれ1層目の活性化関数、重み、バイアス項を表す。

3.2 潜在的類似ユーザの抽出

3.2.1 ユーザ間類似度計算

表面的な類似度は低いが、高次の類似度が高いユーザ、すなわち潜在的類似ユーザの抽出手法について説明する。まずは、表面的な類似度計算を説明する。表面的な類似度とは、お互いが共通に評価しているアイテムだけを考慮したユーザ間の嗜好類似度のことであり、ユーザが好むジャンルや傾向を考慮していない類似度である。本研究では、ユーザ x と y の表面的な類似度として、次式で計算されるピアソン相関を採用する。

$$\rho_{xy} = \frac{\sum_{j \in \mathcal{J}_{xy}} (r_{xj} - \bar{r}_x)(r_{yj} - \bar{r}_y)}{\sqrt{\sum_{j \in \mathcal{J}_{xy}} (r_{xj} - \bar{r}_x)^2} \sqrt{\sum_{j \in \mathcal{J}_{xy}} (r_{yj} - \bar{r}_y)^2}}$$

ただし、 \mathcal{J}_{xy} はユーザ x , y が共に評価値を付与しているアイテム集合であり、 \bar{r}_u はユーザ u の平均評価値とする。

次に、潜在的な類似度計算手法を説明する。中間層における出力 z_u^l は、ユーザの嗜好の高次の表現、例えば、アイテムのジャンルやドメインに対する嗜好を表していると捉える事ができる。従って、この中間層の出力ベクトル間の類似度を計算することで、好みの傾向に関する類似度を計算できる可能性がある。ここで、本研究ではさらにもう一つの仮定を設けた。ユーザがアイテムを選ぶ際に、自身が興味を持ちそうなアイテムを優先的に選択するため、各ユーザの評価値データから得られる嗜好情報には偏りが存在すると仮定できる。つまり、中間層で得られる特徴量は非常にスパースな可能性があり、そのため2ユーザ間の中間層の大抵の成分は似た値になると考えられる。したがって、単純に類似度計算を行うと、多くのユーザが著しく類似してしまう。そこで、お互いに嗜好情報が得られていない成分は無視したい。本研究では、以下の理由から、ベクトルの各成分を平均0、分散1に正規化した上でコサイン類似度による類似度計算を行うことで実現する。

- コサイン類似度は、共に0である成分を無視する。
- 嗜好情報がスパースであるため、あるユニットの平均値は嗜好情報が得られていない状態を表すと仮定できる。

コサイン類似度の計算式は以下の通りである。

$$\cos(x, y) = \frac{\sum_{j=1}^{|z^l|} z_{xj}^l \cdot z_{yj}^l}{\sqrt{\sum_{j=1}^{|z^l|} z_{xj}^l{}^2} \cdot \sqrt{\sum_{j=1}^{|z^l|} z_{yj}^l{}^2}}$$

3.2.2 潜在的類似ユーザを用いた推薦

上記の表面的嗜好類似度と高次の嗜好類似度を利用して、潜在的類似ユーザの抽出、及び推薦を行う。アイデアは単純であり、被推薦者との高次の嗜好類似度ランキングと、表面的な嗜好類似度ランキングの差が大きいユーザを利用して、協調フィルタリングにより推薦を行う。以下に定式化する。

被推薦者 u とユーザ x との表面的な類似度における順位を $rank_u^s(x)$ 、高次の嗜好類似度における順位を $rank_u^h(x)$ とする。さらに、この二つの類似度の差を降順に並べ、その順位を $rank_u^d(x)$ とするこのとき、潜在的類似ユーザ集合を

$$\mathcal{U}_u^h = \{x \in U \setminus \{u\} \mid rank_u^d(x) > topN\}$$

と定義する。ここで、 $topN$ は協調フィルタリングで評価値を予測する際に利用するユーザ数である。このとき、被推薦者 u に対する未評価アイテム i の予測評価値は

$$\tilde{r}_{ui} = \bar{r}_u + \frac{\sum_{x \in \mathcal{U}_u^h \cap \mathcal{J}_i} \cos(u, x) \cdot (r_{xi} - \bar{r}_x)}{\sum_{x \in \mathcal{U}_u^h \cap \mathcal{J}_i} \cos(u, x)}$$

である。ここで、 \mathcal{J}_i はアイテム i を評価しているユーザ集合を表す。

4 評価実験と考察

4.1 データセットと実験の概要

提案手法の評価を行うため、MovieLens1M² のデータセットに対して協調フィルタリングとの比較実験を行った。まずは、評価実験で用いたデータセットの解説を行う。MovieLens1Mとは映画に対する評価値データであり、映画数は約3900件、ユーザ数は約6000ユーザ、評価値数は約100万件である。評価値は1から5の5段階で評価が付けられている。今回の実験では、データセットの内、70%を学習用、10%を検証用、20%をテ

²<http://grouplens.org/datasets/movielens/>

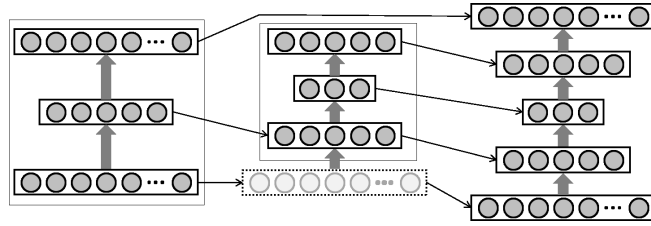


図 2: 積層自己符号化器

スト用に分割した。さらに、それぞれのデータから、評価値の 20% をマスクした。また、自己符号化器のハイパーパラメタは複数のモデルを学習し、検証データの学習誤差が最も低くなったモデルのパラメタを採用した。また、過学習を起こしていることが検証データで確認された場合は、学習を早期打ち切りした。自己符号化器の各種ハイパーパラメタは表 1 にまとめた。また、協調フィルタリングで利用するユーザ数は $topN = 100$ とした。

評価実験で比較するモデルは次の 5 つを用意した。

- model1
単純な協調フィルタリングによる推薦。
- model2
model1 に対して、類似度計算に z_u^2 を利用して計算したモデル。
- model3
model1 に対して、類似度計算に z_u^2 の正規化した値を利用して計算したモデル。
- model4
model1 に対して、類似度計算に z_u^3 の正規化した値を利用して計算したモデル。
- model5
潜在的類似ユーザを用いた推薦。高次の類似度計算には正規化した z_u^2 を利用。
- model6
潜在的類似ユーザを用いた推薦。高次の類似度計算には正規化した z_u^3 を利用。

4.2 推薦精度

まずは、基本的な評価実験として推薦システムの誤差評価である RMSE を用いて、各モデルの推薦精度を評価する。テストデータに対してマスクした評価値を如何に当てられるかで RMSE を計算した。各モデルの RMSE を表 2 に示す。model3 が最も精度が良く、

表 1: 積層自己符号化器のハイパーパラメタ

中間層数	3
中間層のユニット数	50, 20
エポック数	5000
バッチサイズ	30
中間層の活性化関数	シグモイド関数
出力層の活性化関数	恒等写像
ガウスノイズ	$\mu = 0, \sigma^2 = 0.001$
学習手法	Adam

model1 が最も精度が低いことがわかる。これは自己符号化器の中間層を用いた類似度計算の優位性を示している。また、model2 より model3 が高精度であることから、類似度計算に対して、中間層の値を正規化することの有効性を表している。さらに、model1 に比べ、model5 及び model6 の精度が若干向上している。このことから、一般的な協調フィルタリングで類似度が低いと判断されているユーザの中にも、実際には嗜好が類似したユーザが存在し、そのようなユーザを利用して推薦を行っても、精度が下がらないことがわかった。

表 2: RMSE

モデル名	RMSE
model1	1.106
model2	0.962
model3	0.928
model4	0.944
model5	1.043
model6	1.067

4.3 推薦リストの類似度

次に、各手法で推薦リストの内容がどれだけ変化したかを Jaccard 類似度で評価した。推薦リストは被推薦者の平均評価値より高評価と予測したアイテム全て

を対象とした。モデル A とモデル B の推薦リストの類似度を、

$$sim^l(A, B) = \frac{1}{|U|} \sum_{u \in U} jac(list_A(u), list_B(u))$$

として評価する。ここで $list_X(u)$ とは、ユーザ u に対するモデル X の推薦リストを表し、 $jac(X, Y)$ は集合 X, Y の jaccard 類似度である。表 3 に結果をまとめた。結果から、model1 とそれ以外のモデルでは類似度が低いことが判る。つまり、単純に中間層を利用した場合も、潜在的類似ユーザを利用した場合も、基本的な協調フィルタリングとは異なるアイテムが推薦されやすいことが判明した。また、model3 と model4, model5 と model6 の類似度が比較的高いことから、中間層の何層目を利用したかによる影響より、推薦モデルの差が大きいことが明らかになった。

表 3: 推薦リスト間の jaccard 類似度

A \ B	2	3	4	5	6
1	0.22	0.23	0.25	0.25	0.27
2		0.44	0.33	0.28	0.26
3			0.42	0.33	0.29
4				0.32	0.35
5					0.43

4.4 被推薦者と類似ユーザの共評価アイテム数

潜在的類似ユーザを用いた推薦モデルでは、潜在的類似ユーザを利用して推薦を行っているため、被推薦者と潜在的類似ユーザがそれぞれ評価をしているアイテムの重複は少ないと想定される。そこで、被推薦者と、推薦モデルで利用する $topN$ 人の類似ユーザとの間で、高評価をしているアイテム群の jaccard 類似度を計算した。評価の対象とするモデルは model1, 2, 5, 6 である。modelN の実際の計算式は以下の通りである。

$$sim_N^c(u_t, U_t^{sim}) = \frac{1}{|U_t^{sim}|} \sum_{v \in U_t^{sim}} jac(\hat{r}_t, \hat{r}_v)$$

ここで U_t^{sim} は推薦モデルで利用する $topN$ 人の、ユーザ t の類似ユーザ集合であり、 \hat{r}_u はユーザ u が高評価しているアイテムの集合である。ただし、高評価とは評価値が 4, もしくは 5 の場合とした。結果を表 4 にまとめた。

結果から、予想とは反し、model1 が最も類似度が低いことが確認することができた。ピアソン相関は、共

に評価を付与しているアイテムだけを考慮して計算している。従って、嗜好が似ていないために、極少数のアイテムだけで類似度を計算しているユーザが存在する。その際、嗜好は似ていないにも関わらず共評価しているアイテムとは、大衆性の高いアイテムである可能性が高い。さらに、大衆性の高いアイテムは一般的に高評価に分布が偏ることから、偶然類似度が高くなってしまふ。このようにして、model1 には実際に嗜好が似ていないユーザが $topN$ 人の類似ユーザに混入してしまったと考えられる。一方、model3 と比べると、model5, 6 の類似度は減少している。つまり少なくとも、高次の類似ユーザを利用した場合と比べると、潜在的類似ユーザを利用した方が、普段の利用アイテムと非類似したアイテムの推薦が可能である。

表 4: 被推薦者と類似ユーザ間の評価アイテムに関する jaccard 類似度

モデル名	sim_N^c
model1	0.028
model3	0.071
model5	0.048
model6	0.049

5 まとめと今後の課題

本研究では、積層自己符号化器によるユーザ間類似度の変化に着目した協調フィルタリングの提案と、そのシステムの評価のために、MovieLens1M のデータセットを用いて種々のオフライン評価を行った。その結果、自己符号化器の中間層を利用したユーザ間類似度計算の優位性や、その際の値の正規化が精度向上に寄与することを確認した。また、潜在的類似ユーザを利用することで、協調フィルタリング以上の推薦精度を保ちながら、被推薦者と異なるコミュニティのユーザ嗜好を利用した推薦を実現する可能性を示した。

今後の課題としては、まずは被験者実験による評価実験が挙げられる。推薦リストの差異によるユーザ満足度の向上が実際に観測できるのか、ユーザによる主観的評価を検討しなければならない。また、潜在的類似ユーザの抽出方法にも一考の余地が存在する。本研究では、単純に自己符号化器の入力層と中間層での類似度ランキングの差により、潜在的類似ユーザを定義した。しかしながら、類似度ランキングの下位で差が大きい場合に比べ、ランキング上位で差が大きくなった場合を重要視するべきである。さらに、推薦リストの多様性のみならず、セレンディピティなどの評価を向上するための仕組みを入れることも課題である。

参考文献

- [1] R. Salakhutdinov, A. Mnih, and G. Hinton. Restricted boltzmann machines for collaborative filtering. In *Proceedings of the 24th International Conference on Machine Learning, ICML '07*, pp. 791–798, New York, NY, USA, 2007. ACM.
- [2] S. Sedhain, A. K. Menon, S. Sanner, and L. Xie. AutoRec: Autoencoders Meet Collaborative Filtering. In *Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion*, pp. 111–112. ACM, 2015.
- [3] F. Strub and J. Mary. Collaborative Filtering with Stacked Denoising AutoEncoders and Sparse Inputs. In *NIPS Workshop on Machine Learning for eCommerce*, 2015.
- [4] 麻生英樹. 多層ニューラルネットワークによる深層表現の学習 (<連載解説> deep learning(深層学習) [第 2 回]). 人工知能学会誌, Vol. 28, No. 4, pp. 649–659, 2013.
- [5] 田中恒平, 小林亜樹. 深層学習を用いた情報推薦のための欠損値補完手法. 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
- [6] 神畷敏弘. 推薦システムのアルゴリズム (2). 人工知能学会誌, Vol. 23, No. 1, pp. 89–103, 2008.
- [7] 鈴木陽介, 尾崎知伸. 自己符号化器による推薦システムを目的とした入力値ベクトル補完手法の実験的評価. 知識ベースシステム研究会, Vol. 108, pp. 6–11, jun 2016.
- [8] 川上和也, 松尾豊. Deep collaborative filtering deep learning 技術の推薦システムへの応用. 人工知能学会全国大会論文集, Vol. 28, pp. 1–4, 2014.