

Neural Collaborative Filtering に基づく推薦モデルの汎化性能に関する一考察

1X20C071-4 長命祥吾
指導教員 後藤正幸

1. 研究背景・目的

近年 EC サイト等では膨大なアイテムからユーザの嗜好に沿った提案を行う推薦モデルの需要が高まっている。特にユーザのインタラクション（購入、閲覧、クリック等）の有無から導かれる評価データ（以下、暗黙的評価）を用いた推薦モデルである協調フィルタリングは現在の推薦システムを支える重要な技術である。暗黙的評価を用いた推薦モデルの1つに Neural Matrix Factorization（以下、NeuMF）[1]がある。NeuMF は代表的な協調フィルタリングである Matrix Factorization（以下、MF）[2]を一般化した線形モデルである Generalized Matrix Factorization（以下、GMF）[1]と、ニューラルネットワークを用いた非線形モデルである Multi-Layer Perceptron（以下、MLP）[1]を統合することで、複雑な構造の表現を可能にしたモデルである。しかし、暗黙的評価はユーザが直接評価を行っていないためノイズデータが混入しやすく、線形モデルに比べて非線形モデルの NeuMF はこういったノイズを過学習しやすい。そこで本研究では、線形の機構と非線形の機構を1つずつ持つ NeuMF に対して、さらに線形の機構を1つ追加した Twin-G NeuMF を提案する。本稿では、実データを用いた分析により、提案モデルが NeuMF の精度を上回ること、特にノイズが多いときに提案手法が有用であることを示す。

2. Neural Collaborative Filtering (NCF)

Neural Collaborative Filtering (NCF) [1] は協調フィルタリングの1つである。まず NCF ではユーザ $u \in \mathcal{U}$ とアイテム $i \in \mathcal{I}$ の ID を表す one-hot ベクトルの入力をそれぞれ埋め込みベクトル $\mathbf{p}_u, \mathbf{q}_i$ に変換する。これらをニューラル CF 層において変換関数 f に入力することで、購入や閲覧などのインタラクションの予測値 \hat{y}_{ui} が得られる。NCF は式 (1) で表される。

$$\hat{y}_{ui} = f(\mathbf{p}_u, \mathbf{q}_i) \quad (1)$$

このとき、 f を $[0, 1]$ の値域をとる関数とすれば、 \hat{y}_{ui} はユーザ u がアイテム i にインタラクションする確率と捉えることができる。また、NCF の損失関数は式 (2) で表される。

$$p(\mathcal{Y}, \mathcal{Y}^- | \mathbf{P}, \mathbf{Q}, \Theta_f) = \prod_{(u,i) \in \mathcal{Y}} \hat{y}_{ui} \prod_{(u,j) \in \mathcal{Y}^-} (1 - \hat{y}_{uj}) \quad (2)$$

ただし、 \mathbf{P}, \mathbf{Q} はそれぞれ埋め込み層でのユーザとアイテムの埋め込み行列、 Θ_f はニューラル CF 層のパラメータである。また、 \mathcal{Y} はインタラクションが観測されたユーザとアイテムの組の集合、 \mathcal{Y}^- はインタラクションが未観測な組のうち、負例としてランダムにサンプリング（以下、ネガティブサンプリング）した組の集合である。ネガティブサンプリングを行う理由は、インタラクションが観測されたアイテムのみで学習を行うとデータがスパースになり、多様性も欠いて

しまうためである。

NCF には GMF, MLP, NeuMF の3種類のモデルが存在する。GMF は、ニューラル CF 層にアダマール積を用いた線形モデルであるため、単純な構造しか表現できないが過学習は起こりづらい。GMF は式 (3) で表される。以降、 $\sigma(\cdot)$ は活性化関数、 $\mathbf{h}^{G \top}, \mathbf{h}^{M \top}$ はそれぞれ GMF, MLP の出力層において、ベクトルをスカラーに変換する際の各次元の結合の重みを示す。

$$\hat{y}_{ui} = \sigma(\mathbf{h}^{G \top} (\mathbf{p}_u^G \odot \mathbf{q}_i^G)) \quad (3)$$

MLP は、ニューラル CF 層に多層パーセプトロンを用いた非線形モデルであり、複雑な構造を表現可能であるが、過学習が起こりやすい。MLP は式 (4) で表される。ただし、 f_{NN} はニューラルネットワークを表す関数である。

$$\hat{y}_{ui} = \sigma(\mathbf{h}^{M \top} (f_{NN}(\mathbf{p}_u^M, \mathbf{q}_i^M))) \quad (4)$$

NeuMF は、GMF と MLP を統合したモデルであり、ある程度過学習を抑制しつつ複雑な構造を表現可能である。NeuMF は式 (5) で表される。

$$\hat{y}_{ui} = \sigma \left(\begin{bmatrix} \mathbf{h}^G \\ \mathbf{h}^M \end{bmatrix}^\top \begin{bmatrix} \mathbf{p}_u^G \odot \mathbf{q}_i^G \\ f_{NN}(\mathbf{p}_u^M, \mathbf{q}_i^M) \end{bmatrix} \right) \quad (5)$$

NeuMF のモデル概要を図 1 に示す。

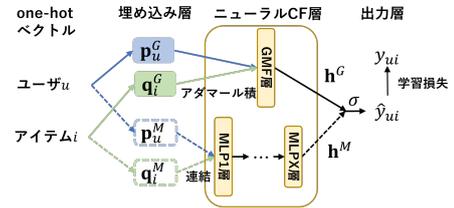


図 1: NeuMF のモデル概要

NeuMF の GMF と MLP に対応する部分のパラメータは、それぞれ事前学習した GMF と MLP のパラメータを初期値として設定し、その後の学習で全体として最適化される。この点において、事前学習後の GMF と MLP の予測値を平均化する通常のアンサンブルとは異なる。

3. 提案モデル

3.1. 着想

暗黙的評価は、インタラクションが容易であるため、様々なノイズの影響を受けやすい。例えば、動画・音楽ストリーミングサービスでは、普段再生しないジャンルでも流行のアイテムだからという理由で再生したり、単に誤って再生したりすることもある。協調フィルタリングは個々のユーザの嗜好に基づいて異なる推薦を行うモデルであるが、これらのインタラクションはユーザ固有の嗜好を反映しているとは言えないため、モデルにとってはノイズとなる。

実際にデータにこれらのノイズを加えた実験では、GMF, MLP, NeuMF の精度が低下することが確認されている (図 3)。比較的ノイズに頑健な線形モデルである GMF の精度

も低下しているため、ノイズが多い状況下で GMF の出力比率を大きくする等では対応が不十分である。そのため、複数の異なるモデルを統合する等でモデルのノイズに対する頑健性を備える必要がある。

3.2. 提案手法

本研究では、NeuMF に対して比較的ノイズに頑健な線形モデルであり、学習コストも低い GMF をさらに追加した Twin-G NeuMF を提案する。すなわち、提案手法では事前学習を行った 2 つの GMF と 1 つの MLP を統合する。ここで、2 つの GMF はネガティブサンプリングによって負例の学習データに差異があるため、学習結果のパラメータも異なる。NeuMF と同様に、GMF と MLP は事前学習時のパラメータを初期値として設定しており、統合後全体として再度学習を行う。提案モデルの概要を図 2 に示す。

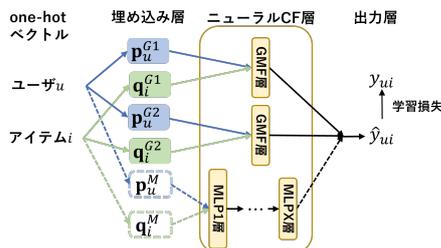


図 2: 提案モデルの概要

4. 実験

本提案モデルの有効性、及び頑健性を検証するため、実データにノイズを加えた人工データを用いて実験を行う。

4.1. 実験条件

本実験では、一般的な推薦システムのデータセットである MovieLens 1M を用いる。このデータセットは映画に関する 5 段階の評価 (明示的評価) であるため、評価の有るときに 1、ないときに 0 として評価の有無に関する暗黙的評価に変換する。さらに、このデータセットにノイズの強さを表すパラメータ $k \in \{0, 1, \dots, 4\}$ で制御される 2 種類のノイズを追加した。1 つ目のノイズは、誤クリック等によるノイズを想定し、あるユーザの評価 1 つにつき $\frac{k}{40}$ の確率で、そのユーザとランダムにサンプリングされた評価未観測アイテムとの間に評価を表す 1 を追加する。2 つ目のノイズは、流行等によるユーザ本来の嗜好と関連の薄いアイテムに対する評価を想定し、ランダムにサンプリングされた $10 \times k$ 個のアイテムについて、それぞれランダムにサンプリングされた全ユーザの 20% に対して評価を表す 1 を追加する。

評価指標には Hit Ratio (HR) を用いる。HR はユーザ 1 人につき、評価されていないアイテム 99 個と評価されたテストアイテム 1 個をランキングしたときに、top10 にテストアイテムが入っている割合を示す。さらに比較手法として、NCF の各モデル、および提案手法に非線形モデルが含まれていない場合として、2 つの GMF のみを統合したモデル (Twin-GMF) を挙げ、提案手法と比較する。

4.2. 実験結果

実験結果を図 3 に示す。

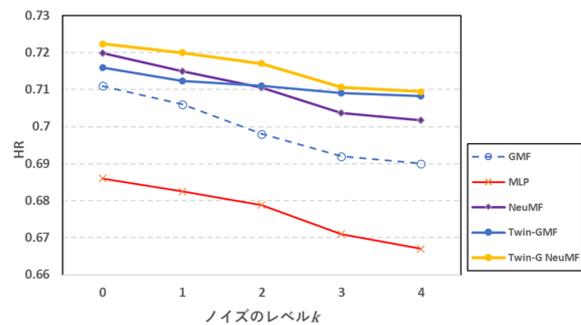


図 3: ノイズがあるデータにおける各モデルの HR

図 3 より、提案手法はノイズがない場合でも NeuMF の精度を上回っており、特にノイズのレベルが大きい環境下ではその差が大きくなることからわかる。また、単に GMF を 2 つ統合するだけでも精度が向上しており、ノイズのレベルが大きくなると提案手法と GMF を 2 つ統合したモデルの精度にほとんど差がなくなる。

5. 考察

Twin-G NeuMF は NeuMF の精度を上回り、特にノイズデータが多いときに精度差が大きくなっていることから、高い頑健性を持ったモデルであるといえる。Twin-G NeuMF はインタラクションがしやすく、データにノイズが多いと考えられる動画・音楽ストリーミングサービスの再生データ等において特に有用であると考えられる。また、ノイズデータが多いときに MLP がある Twin-G NeuMF と、MLP が無い Twin-GMF の精度がほとんど変わらなくなることから、MLP がノイズデータを過学習していると考えられる。さらに、2 つ同じ構造をもつ GMF を統合するだけでも精度が向上したことについては、異なるネガティブサンプリングを基にしたモデルを統合することで、ネガティブサンプリングにおける負例の偏りの影響が緩和されたことが要因であると推測される。

6. まとめと今後の課題

本研究では、2 つの GMF と 1 つの MLP を統合した Twin-G NeuMF を提案した。Twin-G NeuMF は従来の NeuMF の精度を上回り、特にノイズが多いデータでは高い頑健性を示した。また、ネガティブサンプリングが異なるモデルを統合することで、精度が向上する可能性が示唆された。

今後の課題として、ネガティブサンプリングが同じモデルを統合した実験を行うことで、ネガティブサンプリングが異なるモデルを統合したことが精度向上の原因となっているかを検証することが挙げられる。

参考文献

- [1] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceeding of the 26th International Conference on World Wide Web, WWW '17*, p. 173–182, 2017.
- [2] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, Vol. 42, No. 8, pp. 30–37, 2009.