

協調フィルタリングにおけるメタラーニングの適用による疎なデータからの学習と不確か性の推論

平井 遥斗

富山県立大学 情報システム工学科

2023 年 5 月 19 日

はじめに

近年、あらゆる方面での推薦システムの研究が広く行われている，中でも代表的なアプローチとして協調フィルタリングが用いられている．

しかし，これらの手法には二つの欠点が大きくあげられる．

第一に過去の履歴のデータが豊富に利用できる場合には有効だが，まったく新しいアイテムやまったくの新規ユーザといった高度に疎なデータの場合に性能が極端に減少する．

第二に不確実性のモデリング．現実世界への応用を考える際，予測結果の出力のみではなく，その予測の確信度も含めてモデリングすることが重要である．

本研究の目的

本研究では、メタラーニングの派生手法である Nerural Processes(NP) の学習の枠組みを協調フィルタリングに応用し，上記の問題の解決を図る．

背景知識

先行研究に基づき、M 人のユーザと N 個のアイテムからインタラクション行列 $Y \in \mathbb{R}^{M \times N}$ をユーザーのフィードバック、例えばクリックや視聴、収集、購買の履歴から構築する。

$$y_{ui} = \begin{cases} 1, & \text{フィードバックが観測された場合} \\ 0, & \text{それ以外} \end{cases} \quad (1)$$

y_{ui} にベルヌーイ分布を仮定し、確率的にモデリングすることでランキングを構築する。

$$P(y_{ui} = k | p_{ui}) = \begin{cases} 1 - p_{ui}, & k = 0 \\ p_{ui}, & k = 1 \end{cases} \quad (2)$$

p_{ui} は $y_{ui} = 1$ となる確率を意味する。どれだけ u が i とマッチするかの確信度としても解釈することができる。 p_{ui} が 1 に近いということは高い確信度でマッチすると予測している。

Neural Processes

NP は few-shot 学習などを行うメタラーニングの手法の一つである。未知の確率過程 $f: \mathcal{X} \rightarrow \mathcal{Y}$ 由来のいくつかの観測が与えられた際、NP は条件付き表現 r に基づき f の観測値を推論するように学習を行う。

n 個データが観測された際、 m 個の文脈データ $C = (x_j, y_j)_{j=1}^m$ と $n - m$ 個の目標データ $T = (x_j, y_j)_{j=m+1}^n$ に分けられる。

$$r_i = h_{\theta}(x_i, y_i), \quad \forall (x_i, y_i) \in C \quad (3)$$

$$r = r_1 \oplus r_2 \oplus \dots \oplus r_{m-1} \oplus r_m \quad (4)$$

$$\phi = g_{\theta}(x_j, r), \quad \forall (x_j) \in T \quad (5)$$

最初にエンコーダー h_{θ} を持ち文脈データの集合を条件付き表現 r_i に変換する。続いて r_i を順列非依存な集約計算などを用いて固定長な表現である r へ集約を行う。最終的にデコーダー g_{θ} を用いて予測分布の推論を目標点に対して行う。

全体の流れ

行列における未観測の要素の予測というタスクは画像補完タスクの事例に近い部分がある．本研究では，推論を行う目的 y_{ui} に対して，対応する同行，同列における観測されている点のみを文脈データの候補として用いる．

y_{u*} ユーザー u と対応する行ベクトル

y_{*i} ユーザー i と対応する列ベクトル

目的とする y_{ui} に対し， y_{u*} と y_{*i} のうち観測された点を確率的に任意の数サンプリングし，文脈点として用いる．

本研究ではこの仮定のことを MetaCF サンプリングと呼ぶ．

$$p_u \sim \text{Uniform}(0, 1) \quad (6) \qquad r_u = h_{\theta}^U(y'_{u*}) \quad (9)$$

$$m_{u*} \sim \text{Bernoulli}(p_u) \quad (7) \qquad r_i = h_{\theta}^I(y'_{*i}) \quad (10)$$

$$y'_{u*} = m_{u*} \odot y_{u*} \quad (8) \qquad \phi_{ui} = g_{\theta}(r_u, r_i) \quad (11)$$

ネットワーク機構

手法の有効性の検証として従来手法の CFNet に適応させることで有効性を確認する.

従来の深層学習を用いた手法は大きく 2 種類に大別される.

一つ目は表現学習

表現学習は DNN を用いて直接同空間内で類似度を図ることで推論を行うことを目的とする. 本研究では提案手法 MetaCF と組み合わせた表現学習ネットワークを **MetaCF-fl** と呼ぶ.

二つ目はマッチング学習

マッチング学習では DNN を用いてユーザとアイテムの表現の合成を複雑な関数で近似することで推論を行う. 本研究では提案手法 MetaCF と組み合わせたマッチング学習のネットワークを **MetaCF-ml** と呼ぶ.

MetaCFNet

また MetaCF-rl と MetaCF-ml を組み合わせることで最終的な提案ネットワーク機構 **MetaCFNet** となる.

これらのネットワークに適用することで提案手法がネットワークの構造に依存しないことを確認する.

データセット

一般公開されている 4 つのデータセットを用いる：MovieLens 1M(ml-1m), LastFM(lastfm), Amazon music(AMusic), Amazon toys(AToys)

ml-1m と lastfm は最低 20 個の履歴があるユーザのみが含まれるように加工されて公開している．一方で AMusic では 59% のユーザが，AToy では 49% のユーザが 20 より少ない履歴しかもっていない．後者二つはより疎なデータセットであるといえる．

評価

ランキングの評価には二つの代表的な指標 Hit Ratio(HR) と Normalized Discounted Cumulative Gain(NDCG) を用いる．

$$HR@k = \begin{cases} 1, & \text{if } r_{test}(u, i) \in R_k \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

$$NDCG@k = \frac{DCG@k}{IDCG@k} = \sum_{j=1}^k \frac{2^{r(u,j)} - 1}{\log(i+1)} \quad (13)$$

実験結果

表 1: ベースライン手法との比較

Datasets	Measures	Existing Methods			MetaCF			Improvement from CFnet
		CFNet-rl	CFNet-ml	CFnet	MetaCF-rl	MetaCF-ml	MetaCFNet	
ml-1m	HR	0.7127	0.7073	0.7253	0.5354	0.7142	0.7280	0.37%
	NDCG	0.4336	0.4264	0.4416	0.2960	0.4339	0.4353	-1.4%
lastfm	HR	0.8840	0.8834	0.8995	0.8489	0.8925	0.8949	-0.5 %
	NDCG	0.6001	0.5919	0.6186	0.5585	0.6046	0.6074	-1.8%
AMusic	HR	0.3947	0.4071	0.4116	0.5051	0.5456	0.5557	35.0%
	NDCG	0.2504	0.2420	0.2601	0.2753	0.3027	0.3083	18.5%
AToy	HR	0.3746	0.3931	0.4150	0.5413	0.6025	0.6063	44.9%
	NDCG	0.2271	0.2293	0.2513	0.2988	0.3346	0.3405	35.1%

MetaCFNet は ml-1m と lastfm においては CFNet とほぼ同等の結果を示し、AMusic と AToy では大きな性能の向上を示した。

後者はコールドユーザを含むという点において、疎なデータセットからの学習に強みを示している。

不確実性の評価

ϕ は確信度であり，不確実性の評価を二値エントロピー関数 H であらわす．

$$H(\phi) = -\phi \log_2 \phi - (1 - \phi) \log_2 (1 - \phi) \quad (14)$$

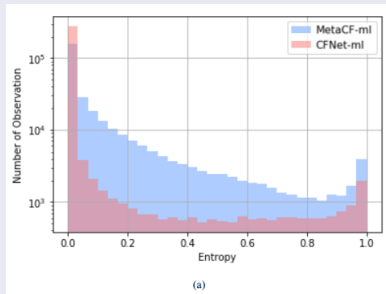


図 1: エントロピーのヒストグラム

図より CFNet はほとんど 0 か 1 に偏っているのに対し，MetaCF はより広範囲に広がっていることが確認できる．提案手法は明示的に不確実性をモデリングできているといえる手法である．

まとめ

- NP の学習の枠組みを推薦システムに拡張し、疎なデータでも効率的にまた不確実性をもって学習を行う MetaCF という学習方法を提案した.
- 近年提案された CFNet に MetaCF を適用した MetaCFNet というネットワークを提案した. top-N 推薦タスクにおいて, 4 つのデータセットにおいてきわめて大きな性能を示した.
- 実験的に過去履歴の少ないユーザやアイテムに対する推論には高い不確実性をもって出力を示していることを確認した.

清水さんの研究において

この研究では, ベルヌーイ分布という観測 (されるか, されないか) の 2 値をとる場合であったが, 清水さんの研究では授業の成績の 5 段階の 5 つの値をとる, そこで解決案として, **カテゴリカル分布**というベルヌーイ分布を多次元に拡張した確率分布に置き換えることで適用できるかもしれない.