

# 控制与决策

Control and Decision

## 基于序贯三支神经网络的个性化推荐

成淑慧, 武优西

引用本文:

成淑慧,武优西. 基于序贯三支神经网络的个性化推荐[J]. *控制与决策*, 2024, 39(3): 1012–1020.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1227>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [基于地标特征和元学习方法推荐最适用优化算法](#)

Recommending best suitable metaheuristic based on landmarking feature and meta-learning approach

*控制与决策*. 2021, 36(5): 1223–1231 <https://doi.org/10.13195/j.kzyjc.2019.0993>

#### [基于MCPDDPG的智能车辆路径规划方法及应用](#)

The method and application of intelligent vehicle path planning based on MCPDDPG

*控制与决策*. 2021, 36(4): 835–846 <https://doi.org/10.13195/j.kzyjc.2019.0460>

#### [基于改进卷积神经网络的动力下肢假肢运动意图识别](#)

Intent recognition of power lower-limb prosthesis based on improved convolutional neural network

*控制与决策*. 2021, 36(12): 3031–3038 <https://doi.org/10.13195/j.kzyjc.2020.0326>

#### [基于生物启发神经网络和DMPC的多机器人协同搜索算法](#)

Multi-robot cooperative search algorithm based on bio-inspired neural network and DMPC

*控制与决策*. 2021, 36(11): 2699–2706 <https://doi.org/10.13195/j.kzyjc.2020.0959>

#### [基于卷积长短时记忆神经网络的城市轨道交通短时客流预测](#)

Metro short-term traffic flow prediction with ConvLSTM

*控制与决策*. 2021, 36(11): 2760–2770 <https://doi.org/10.13195/j.kzyjc.2020.0501>

## 基于序贯三支神经网络的个性化推荐

成淑慧<sup>1</sup>, 武优西<sup>1,2†</sup>

(1. 河北工业大学 经济管理学院, 天津 300401; 2. 河北工业大学 人工智能与数据科学学院, 天津 300401)

**摘要:** 虽然协同过滤可以实现用户的个性化推荐, 但是大多数协同过滤及其改进模型未考虑用户和项目等特征, 因而不能发掘样本间的非线性关系. 与协同过滤相比, 深度学习能挖掘丰富的用户兴趣模式, 但网络拓扑结构是基于二支决策的方式, 忽略了推荐样本的难易程度. 为了增强模型的非线性表达, 同时区分推荐样本的难易, 受序贯三支决策的启发, 提出序贯三支决策神经网络个性化推荐模型 (personalized recommendation model based on sequential three-way decision with single feedforward neural network, STWD-SFNN-PR). 首先, 为了将高维稀疏特征向量映射为低维稠密的特征向量, STWD-SFNN-PR 采用嵌入进行特征处理. 其次, 在增量式的网络结构中学习推荐样本, 使用 Adam 优化网络参数, 并返回难以推荐的样本. 再次, 利用序贯三支决策增加延迟决策的策略, 并在不同的粒度层采用序贯的阈值, 从而动态地实现难以推荐样本的划分. 最后, 为了验证模型的可行性和有效性, 选择多种电影推荐数据集进行研究, 并选择经典的神经网络推荐、经典的深度学习推荐和最新的三支协同过滤推荐进行对比. 实验结果表明, STWD-SFNN-PR 具有更优的推荐质量.

**关键词:** 序贯三支决策; 神经网络; 网络拓扑结构; 推荐系统; 协同过滤; 个性化推荐

中图分类号: TP391.3; TP183

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1227

引用格式: 成淑慧, 武优西. 基于序贯三支神经网络的个性化推荐[J]. 控制与决策, 2024, 39(3): 1012-1020.

## Personalized recommendation based on sequential three-way decisions with single feedforward neural network

CHENG Shu-hui<sup>1</sup>, WU You-xi<sup>1,2†</sup>

(1. College of Economics and Management, Hebei University of Technology, Tianjin 300401, China; 2. College of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China)

**Abstract:** Although collaborative filtering (CF) can realize the personalized recommendation of users, most of the CF and their improved models do not consider the characteristics of users and items, hence they cannot explore the nonlinear relationship among samples. Compared with the CF, although deep learning can mine rich patterns of user interests, the network topology is based on two-way decision-making, ignoring the difficulty of recommended samples. To enhance the nonlinear expression of the model and distinguish the difficulty of the recommended samples, a personalized recommendation model based on sequential three-way decision with single feedforward neural network (STWD-SFNN-PR) is proposed. Firstly, to map high-dimensional sparse feature vectors to low-dimensional dense feature vectors, the STWD-SFNN-PR adopts embedding technology for feature preprocessing. Secondly, the characteristics of users and items are learned in the incremental network structure, Adam is employed to optimize network parameters, and samples that are difficult to recommend are returned. Thirdly, sequential three-way decision-making is utilized to increase the strategy of delayed decision-making, and the sequential threshold parameters are adopted at different granularity levels, thus realizing the classification of the recommended samples dynamically. Finally, to verify the feasibility and effectiveness of the model, a variety of movie recommendation datasets are selected for research, and the classic neural network recommendation model, the classic deep learning recommendation model, and the latest CF recommendation models with three-way are compared. Experimental results show that the STWD-SFNN-PR has better recommendation quality.

**Keywords:** sequential three-way decision; neural network; network topology; recommendation system; collaborative filtering; personalized recommendation

收稿日期: 2022-07-10; 录用日期: 2022-11-22.

基金项目: 国家社会科学基金项目 (18BGL191).

†通讯作者. E-mail: wuc567@163.com.

## 0 引言

随着互联网的发展,推荐系统<sup>[1]</sup>从网络购物和新闻资讯等方面深远影响着人们的生活,它不仅可以帮助用户从海量的信息中找到自己感兴趣的商品或服务,还可以使生产者生产的信息受到目标用户的关注.因此,推荐领域受到了很多研究者的关注.

推荐模型是推荐系统中最核心的一个环节,可以归纳为传统推荐模型和深度学习推荐模型.传统推荐模型具有较强的模型可解释性和较快速的训练过程,例如协同过滤<sup>[2]</sup>、逻辑回归<sup>[3]</sup>和因子分解机<sup>[4]</sup>等.深度学习推荐模型具有较强的表达能力,可以挖掘更多有价值的信息,例如自编码推荐<sup>[5]</sup>、基于因子分解机支持的神经网络<sup>[6]</sup>和神经协同过滤<sup>[7]</sup>等.然而,当前推荐算法在应用中存在两点不足:1)以协同过滤<sup>[2,8]</sup>为代表的传统推荐模型主要根据用户对项目的评分来推测用户的偏好,并没有考虑项目信息等特征对用户评分的影响,导致用户偏好测度不精确;2)以图神经网络<sup>[9]</sup>为代表的深度学习推荐模型的本质是基于二支决策的形式,即在推荐过程中只考虑推荐和不推荐的决策方案,忽略了推荐样本的难易程度.

为了考虑有关用户信息和项目信息等特征,同时有效地挖掘推荐样本间的非线性关系,本文选择单隐藏层的前馈型神经网络<sup>[10]</sup>(single feedforward neural network, SFNN)为主推荐模型.原因是SFNN具有较强的非线性表达能力,更重要的是SFNN网络拓扑结构简单,灵活性较好.此外,为有效地区分推荐样本的难易程度,且面向难以推荐的样本进行有针对性地学习,本文在推荐模型中引入序贯三支决策思想<sup>[11-12]</sup>,在推荐过程中考虑延迟推荐的决策方案,提出序贯三支决策神经网络的个性化推荐模型(personalized recommendation model based on sequential three-way decisions with single feedforward neural network, STWD-SFNN-PR).STWD-SFNN-PR是将个性化推荐问题转化为三支决策的分类问题,在向目标用户推荐其可能喜欢的项目时,针对无法立即做出决策的项目,通过将其划分至延迟决策域的方式来降低错误推荐的损失,即利用序贯三支决策思想动态给出目标用户的推荐列表.本文的主要工作如下:

1)为了综合考虑有关用户信息和项目信息等特征,并避免二支决策的缺陷,受序贯三支决策的启发,采用序贯三支决策优化SFNN的拓扑结构,提出高效的模型STWD-SFNN-PR.

2)在推荐过程中,STWD-SFNN-PR使用嵌入处

理高维稀疏的特征向量,并将SFNN难以决策的推荐样本划分至延迟决策域,不断深入挖掘这类样本的特征,从而动态地实现用户的个性化推荐.

3)选用多种电影推荐数据集验证模型的推荐质量,实验表明,STWD-SFNN-PR的 $F_1$ 和MAP高于经典的神经网络推荐模型、经典的深度学习推荐模型和最新的三支指导下的协同过滤推荐模型.

## 1 相关理论

### 1.1 神经网络

SFNN<sup>[13]</sup>是由输入层、隐藏层和输出层组成,通过激活函数的作用,将输入数据进行非线性组合,产生非线性决策边界,从而增强网络的表达能力.

假设SFNN有 $m$ 个输入层结点, $K$ 个隐藏层结点和 $n$ 个输出层结点,则在二分类的数据集中,输入层到隐藏层之间的权重和偏置分别为 $\mathbf{w}^{(1)} = [\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, \dots, \mathbf{w}_K^{(1)}]^T \in \mathbf{R}^K$ 和 $\mathbf{b}^{(1)} = [\mathbf{b}_1^{(1)}, \mathbf{b}_2^{(1)}, \dots, \mathbf{b}_K^{(1)}]^T \in \mathbf{R}^K$ ,隐藏层到输出层之间的权重和偏置分别为 $\mathbf{w}^{(2)} = [\mathbf{w}_1^{(2)}, \mathbf{w}_2^{(2)}]^T \in \mathbf{R}^2$ 和 $\mathbf{b}^{(2)} = [\mathbf{b}_1^{(2)}, \mathbf{b}_2^{(2)}]^T \in \mathbf{R}^2$ .在激活函数 $g$ 的作用下,隐藏层的输出矩阵表示为

$$\mathbf{G}(\mathbf{w}^{(1)} \times \mathbf{X} + \mathbf{b}^{(1)}) = \begin{bmatrix} g_1(\mathbf{w}_1^{(1)} \times \mathbf{x}_1 + \mathbf{b}_1^{(1)}) & \dots & g_1(\mathbf{w}_1^{(1)} \times \mathbf{x}_d + \mathbf{b}_1^{(1)}) \\ \vdots & \ddots & \vdots \\ g_K(\mathbf{w}_K^{(1)} \times \mathbf{x}_1 + \mathbf{b}_K^{(1)}) & \dots & g_K(\mathbf{w}_K^{(1)} \times \mathbf{x}_d + \mathbf{b}_K^{(1)}) \end{bmatrix},$$

其中 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_d]$ .SFNN输出层的输出矩阵为

$$\bar{\mathbf{y}} = \mathbf{w}^{(2)} \times \mathbf{G}(\mathbf{w}^{(1)} \times \mathbf{X} + \mathbf{b}^{(1)}) + \mathbf{b}^{(2)} = \begin{bmatrix} \bar{y}_{11} & \dots & \bar{y}_{1d} \\ \bar{y}_{21} & \dots & \bar{y}_{2d} \end{bmatrix}.$$

为了以概率的形式输出SFNN的预测值,选择Sigmoid函数.于是,SFNN的最终预测结果为

$$\hat{\mathbf{y}} = \begin{bmatrix} \frac{\exp(\bar{y}_{11})}{\exp(\bar{y}_{11}) + \exp(\bar{y}_{21})} & \dots & \frac{\exp(\bar{y}_{1d})}{\exp(\bar{y}_{1d}) + \exp(\bar{y}_{2d})} \\ \frac{\exp(\bar{y}_{21})}{\exp(\bar{y}_{11}) + \exp(\bar{y}_{21})} & \dots & \frac{\exp(\bar{y}_{2d})}{\exp(\bar{y}_{1d}) + \exp(\bar{y}_{2d})} \end{bmatrix} = \begin{bmatrix} \hat{y}_{11} & \dots & \hat{y}_{1d} \\ \hat{y}_{21} & \dots & \hat{y}_{2d} \end{bmatrix}.$$

### 1.2 序贯三支决策

STWD<sup>[14]</sup>体现了多粒度<sup>[15]</sup>分层次处理问题的思想,随着信息的更新渐进地给出完整的决策方案.

假设论域 $U$ 由 $n$ 层粒度构成,即 $\{L_1, L_2, \dots, L_n\}$ ,且粒度层从第1层逐级遍历至第 $n$ 层.此外,STWD在粒度层中存在全序关系<sup>[16]</sup>,即 $[x]_{L_1} \preceq [x]_{L_2} \preceq \dots \preceq [x]_{L_n}$ ,其中 $[x]_{L_1}$ 是样本 $x$ 在信息量相对较少时对应

的划分,而  $[x]_{L_n}$  则是样本  $x$  在信息量相对丰富时对应的划分.同时,STWD从第1粒度层遍历至第  $n-1$  粒度层存在条件概率  $P^{(i)}(1 \leq i \leq n-1)$  及其相应的一组阈值  $(\alpha_i, \beta_i)$ ,而在第  $n$  粒度层存在条件概率  $P^{(n)}$  和其相应的一个阈值  $\gamma$ ,且这些阈值满足如下的关系式:

$$0 < \beta_1 \leq \dots \leq \beta_{n-1} < \gamma < \alpha_{n-1} \leq \dots \leq \alpha_1 < 1.$$

给定第  $i(1 \leq i \leq n-1)$  粒度层  $L_i$ ,则  $L_i$  可分为3个互不相交的域,即正域  $\text{POS}_{(\alpha_i, \beta_i)}(Y)$ 、边界域  $\text{BND}_{(\alpha_i, \beta_i)}(Y)$  和负域  $\text{NEG}_{(\alpha_i, \beta_i)}(Y)$ ,且它们的定义如下所示:

$$\begin{aligned} \text{POS}_{(\alpha_i, \beta_i)}(Y) &= \{x \in U_i \mid P^{(i)} \succeq_i \alpha_i\}, \\ \text{BND}_{(\alpha_i, \beta_i)}(Y) &= \{x \in U_i \mid \beta_i \prec_i P^{(i)} \prec_i \alpha_i\}, \\ \text{NEG}_{(\alpha_i, \beta_i)}(Y) &= \{x \in U_i \mid P^{(i)} \preceq_i \beta_i\}, \end{aligned}$$

其中  $Y$  是由正类标签构成的集合.随着粒度层的增加,边界域将越来越小.当  $\text{BND}_{(\alpha_{n-1}, \beta_{n-1})}(Y) \neq \emptyset$  时,STWD在第  $n$  层采取二支决策的方案.给定第  $n$  粒度层  $L_n$ ,则  $L_n$  可分为如下两个互不相交的区域:

$$\begin{aligned} \text{POS}_\gamma(Y) &= \{x \in U_n \mid P^{(n)} \succeq \gamma\}, \\ \text{NEG}_\gamma(Y) &= \{x \in U_n \mid P^{(n)} \prec \gamma\}. \end{aligned}$$

## 2 模型构建

为了提高推荐质量,提出STWD-SFNN-PR,通过对比用户对项目的偏好概率值和序贯阈值参数的大小,从而决策是否向目标用户推荐某项目.

### 2.1 粒度层的构建

给定一个决策集  $\{a_P, a_B, a_N\}$ ,其中  $a_P, a_B$  和  $a_N$  分别表示将样本划分至 POS、BND 和 NEG 的决策方案.假设论域  $U$  有  $n(n \geq 1)$  层粒度,由于不同的决策方案造成的决策风险是不同的,则在第  $i(1 \leq i \leq n)$  粒度层的决策损失矩阵有

$$\mathbf{A}^{(i)} = \begin{bmatrix} \lambda_{PP}^{(i)} & \lambda_{BP}^{(i)} & \lambda_{NP}^{(i)} \\ \lambda_{PN}^{(i)} & \lambda_{BN}^{(i)} & \lambda_{NN}^{(i)} \end{bmatrix}. \quad (1)$$

其中:  $\lambda_{PP}^{(i)}$ 、 $\lambda_{BP}^{(i)}$  和  $\lambda_{NP}^{(i)}$  表示针对正样本分别采取  $a_P$ 、 $a_B$  和  $a_N$  的结果损失;  $\lambda_{PN}^{(i)}$ 、 $\lambda_{BN}^{(i)}$  和  $\lambda_{NN}^{(i)}$  表示针对负样本分别采取  $a_P$ 、 $a_B$  和  $a_N$  的结果损失.重要的是,为了满足序贯性,决策损失矩阵的各个元素需要满足如下条件:

- 1)  $0 \leq \lambda_{PP}^{(i)} < \lambda_{BP}^{(i)} < \lambda_{NP}^{(i)} < 1$ ;
- 2)  $0 \leq \lambda_{NN}^{(i)} < \lambda_{BN}^{(i)} < \lambda_{PN}^{(i)} < 1$ ;
- 3)  $(\lambda_{BN}^{(i)} - \lambda_{NN}^{(i)}) \times (\lambda_{BP}^{(i)} - \lambda_{PP}^{(i)}) < (\lambda_{PN}^{(i)} - \lambda_{BN}^{(i)}) \times (\lambda_{NP}^{(i)} - \lambda_{BP}^{(i)})$ .

由于 STWD-SFNN-PR 在各个粒度层上是采取

三支决策的策略,为了防止模型的过拟合,STWD-SFNN-PR 设置了防止过拟合的机制:当第  $n$  粒度层的 BND 域非空时,STWD-SFNN-PR 将采取二支决策方案.因此,STWD-SFNN-PR 在前  $n-1$  粒度层和第  $n$  粒度层具有不同的序贯阈值参数.前  $n-1$  粒度层的三支决策的阈值  $(\alpha_i, \beta_i)(1 \leq i \leq n-1)$  定义为

$$\alpha_i = \frac{\lambda_{PN}^{(i)} - \lambda_{BN}^{(i)}}{(\lambda_{PN}^{(i)} - \lambda_{BN}^{(i)}) + (\lambda_{BP}^{(i)} - \lambda_{PP}^{(i)})}, \quad (2)$$

$$\beta_i = \frac{\lambda_{BN}^{(i)} - \lambda_{NN}^{(i)}}{(\lambda_{BN}^{(i)} - \lambda_{NN}^{(i)}) + (\lambda_{NP}^{(i)} - \lambda_{BP}^{(i)})}. \quad (3)$$

同时,第  $n$  粒度层的二支决策的阈值  $\gamma$  定义为

$$\gamma = \frac{\lambda_{PN}^{(n)} - \lambda_{NN}^{(n)}}{(\lambda_{PN}^{(n)} - \lambda_{NN}^{(n)}) + (\lambda_{NP}^{(n)} - \lambda_{PP}^{(n)})}. \quad (4)$$

假设论域  $U$  从第1粒度层遍历至第  $i$  粒度层,此时 STWD-SFNN-PR 的正域、边界域和负域分别表示为  $\text{POS}^{(i)}$ 、 $\text{BND}^{(i)}$  和  $\text{NEG}^{(i)}$ ,则

$$\text{POS}^{(i)} = \bigcup_{k=1}^i \text{POS}(L_k), \quad (5)$$

$$\text{BND}^{(i)} = \text{BND}(L_i), \quad (6)$$

$$\text{NEG}^{(i)} = \bigcup_{k=1}^i \text{NEG}(L_k). \quad (7)$$

### 2.2 STWD-SFNN-PR

STWD-SFNN-PR 包含输入层、嵌入层、拼接层、序贯三支决策神经网络层和输出层,如图1所示.

1) 输入层.若第  $s$  位用户对第  $t$  个项目存在交互行为,则 STWD-SFNN-PR 的输入层数据可以形式化为  $\{(x_s^t, y_s^t) \mid x_s^t \in \mathbf{R}^m, y_s^t \in \mathbf{R}^n, 1 \leq s \leq S, 1 \leq t \leq T\}$ .其中:  $x_s^t$  表示由第  $s$  位用户的用户信息和第  $t$  个项目的项目信息构成的  $m$  维特征向量;  $y_s^t$  表示第  $s$  位用户对第  $t$  个项目的交互行为构成的  $n$  维向量. STWD-SFNN-PR 是将个性化推荐问题转化为三支决策的分类问题,因此交互行为向量进一步表示为  $y_s^t \in \mathbf{R}^1$ .例如,在5分值的评分表中,当用户  $s$  对项目  $t$  的评分大于3.5分时,定义为用户  $s$  喜欢项目  $t$ ,即  $y_s^t \stackrel{\text{def}}{=} +1$ ;反之,则用户  $s$  不喜欢项目  $t$ ,即  $y_s^t \stackrel{\text{def}}{=} -1$ .

2) 嵌入层.第  $s$  位用户对第  $t$  个项目的  $m$  维特征向量  $x_s^t$ ,它一般包含用户年龄和项目派别等类别型特征,用户 ID 和项目 ID 等 ID 型特征以及数值型特征.由于类别型特征和 ID 型特征在转换成数值型数据时存在高维度的极度稀疏现象,一定程度上影响模型的运行效率和推荐质量.

嵌入技术可以将稀疏的高维类别型特征转换成稠密的低维数值型特征.针对  $p$  位用户,假设某一特征有  $q$  个离散值,首先将该特征的原始向量进行独热

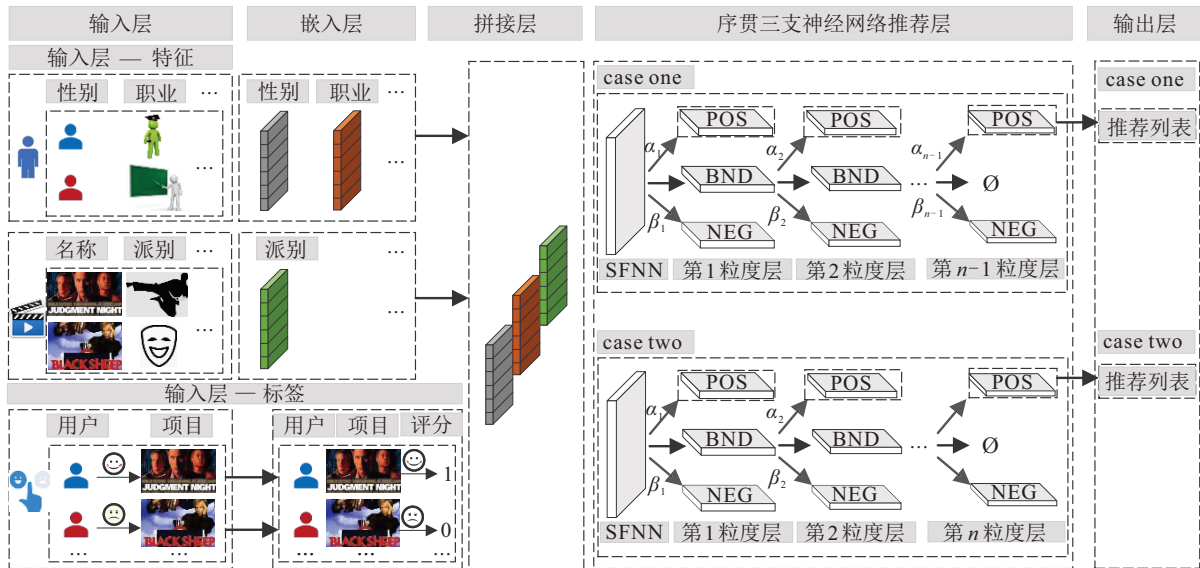


图1 STWD-SFNN-PR结构框架

编码并标记为  $V_{p \times q}$ . 然后, 给定嵌入层的维度  $k$  和嵌入层向量  $E_{q \times k}$ , 其中  $k < q$ , 则嵌入层的输出向量是由  $V_{p \times q}$  与  $E_{q \times k}$  在矩阵乘法作用下产生的, 即  $\bar{V}_{p \times k} = V_{p \times q} \times E_{q \times k}$ , 此时嵌入层的维度从  $q$  维压缩至  $k$  维. 此外,  $E_{q \times k}$  可以随机产生, 也可以经过预训练产生.

相较于矩阵分解等方法产生的特征向量, 嵌入技术可以引入任意信息进行编码, 使其本身可以包含更多有价值的信息, 因此嵌入技术的表达能力更强. 因此, 为了解决输入层的稀疏特征向量问题, STWD-SFNN-PR 在模型构建时考虑增加嵌入层, 即在嵌入层使用嵌入技术, 实现推荐样本的预处理.

3) 拼接层. STWD-SFNN-PR 拼接层是将嵌入层的不同特征横向拼接起来, 从而形成包含全部特征的全新的特征向量. 需要注意的是, 当输入层的特征包含数值型特征时, 将其横向拼接在嵌入层的后方即可. 如图1所示, STWD-SFNN-PR 组合来自嵌入层的用户嵌入向量和项目嵌入向量.

4) 序贯三支神经网络层. 在获得更新后的特征向量后, STWD-SFNN-PR 首先统计每位用户的候选项目, 通过组合所有用户的候选项目构成模型的推荐候选项目集合. 其次, 在给定的 SFNN 拓扑结构下捕获特征的交叉信息, 分别进行神经网络的前向学习和误差反向传播过程.

为了防止模型的过拟合, STWD-SFNN-PR 在前向学习过程考虑带有  $L_2$  正则化的代价函数

$$L_{nn} = L_{fl} + \frac{\lambda_S}{2} \times \sum_{o=1,2, i=1,2} \|w_i^{(o)}\|^2 + \|b_i^{(o)}\|^2. \quad (8)$$

其中:  $\lambda_S$  为 STWD-SFNN-PR 的正则化因子;  $w_i^{(o)}$  和

$b_i^{(o)}$  分别代表第  $i$  ( $1 \leq i \leq n$ ) 个隐藏层结点的网络权重和偏置;  $o$  为 1 和 2 分别代表输入层与隐藏层的连接和隐藏层与输出层的连接. 同时,  $L_{fl}$  是 Focal loss 函数<sup>[17]</sup> 且计算公式如下所示:

$$L_{fl} = \begin{cases} -\delta(1 - \hat{y})^\theta \log \hat{y}, & y = +1; \\ -(1 - \delta)\hat{y}^\theta \log(1 - \hat{y}), & y = -1. \end{cases}$$

其中:  $\hat{y}$  是模型的预测值;  $\delta$  是平衡正负样本比例不均的因子;  $\theta$  是调节易分类样本权重的聚焦参数. 为了增强神经网络的学习能力, 选取 Adam<sup>[18]</sup> 指导网络参数的优化过程. 同时, 模型返回由误分类样本构成的数据集, 在误分类样本数据集上使用  $k$ -means++<sup>[19]</sup> 将数值型数据转化为 STWD 需要的整型数据.

然后, 利用 STWD 的思想, 在 STWD-SFNN-PR 的第  $i$  ( $1 \leq i \leq n$ ) 粒度层  $L_i$  上计算第  $s$  位用户对第  $t$  个项目存在交互行为的条件概率, 且计算公式如下:

$$P^{(i,s,t)} = P(Y | [x]_{L_i}^{(s,t)}) = \frac{|Y \cap [x]_{L_i}^{(s,t)}|}{|[x]_{L_i}^{(s,t)}|}. \quad (9)$$

其中:  $[x]_{L_i}^{(s,t)}$  是由第  $L_i$  粒度层具有相同特征的样本构成的集合,  $Y$  是第  $s$  位用户喜欢的项目集合. 在第  $i$  ( $1 \leq i \leq n-1$ ) 粒度层, 模型考虑的总风险损失为

$$R^{(i)} = \sum_{\substack{1 \leq s \leq S, 1 \leq t \leq T, \\ P^{(i,s,t)} \geq \alpha_i}} \lambda_{PP}^{(i)} \times P^{(i,s,t)} + \lambda_{PN}^{(i)} \times (1 - P^{(i,s,t)}) + \varepsilon \left( \sum_{\substack{1 \leq s \leq S, 1 \leq t \leq T, \\ \beta_i < P^{(i,s,t)} < \alpha_i}} \lambda_{BP}^{(i)} \times P^{(i,s,t)} + \lambda_{BN}^{(i)} \times (1 - P^{(i,s,t)}) \right) + \sum_{\substack{1 \leq s \leq S, 1 \leq t \leq T, \\ P^{(i,s,t)} \leq \beta_i}} \lambda_{NP}^{(i)} \times P^{(i,s,t)} + \lambda_{NN}^{(i)} \times (1 - P^{(i,s,t)}). \quad (10)$$

其中:矩阵 $\mathbf{A}^{(i)}$ 和阈值 $(\alpha_i, \beta_i)$ 由式(1)~(3)可知; $\varepsilon$ 为惩罚因子,避免过多的样本被划分至BND域.因此,模型在第 $i$  ( $1 \leq i \leq n-1$ )粒度层的推荐准则如下:

- (P1) 若 $P^{(i,s,t)} \geq \alpha_i$ ,则 $[x]_{L_i}^{(s,t)} \in \text{POS}_{(L_i)}$ ;  
 (B1) 若 $\beta_i < P^{(i,s,t)} < \alpha_i$ ,则 $[x]_{L_i}^{(s,t)} \in \text{BND}_{(L_i)}$ ;  
 (N1) 若 $P^{(i,s,t)} \leq \beta_i$ ,则 $[x]_{L_i}^{(s,t)} \in \text{NEG}_{(L_i)}$ .

最后,当STWD-SFNN-PR遍历至第 $n-1$ 粒度层的BND域非空时,由式(1)和(4)可知第 $n$ 粒度层的矩阵 $\mathbf{A}^{(n)}$ 和阈值 $\gamma$ ,并考虑以下的总风险损失:

$$R^{(n)} = \sum_{\substack{1 \leq s \leq S, 1 \leq t \leq T, \\ P^{(n,s,t)} \geq \gamma}} \lambda_{PP}^{(n)} \times P^{(n,s,t)} + \lambda_{PN}^{(n)} \times (1 - P^{(n,s,t)}) + \sum_{\substack{1 \leq s \leq S, 1 \leq t \leq T, \\ P^{(n,s,t)} < \gamma}} \lambda_{NP}^{(n)} \times P^{(n,s,t)} + \lambda_{NN}^{(n)} \times (1 - P^{(n,s,t)}). \quad (11)$$

此时,STWD-SFNN-PR的推荐准则如下:

- (P2) 若 $P^{(n,s,t)} \geq \gamma$ ,则 $[x]_{L_n}^{(s,t)} \in \text{POS}_{(L_n)}$ ;  
 (N2) 若 $P^{(n,s,t)} < \gamma$ ,则 $[x]_{L_n}^{(s,t)} \in \text{NEG}_{(L_n)}$ .

5) 输出层. 根据式(5)~(7)计算各个粒度层的正域、延迟决策域和负域,并返回正域的推荐项目. STWD-SFNN-PR将返回前 $i$  ( $1 \leq i \leq n$ )粒度层的所有项目划分至POS域,同时返回模型的各个评价指标值.

### 2.3 时间复杂度分析

给定含有 $d$ 个样本和 $f$ 个特征的数据集,假设STWD-SFNN-PR有 $m$ 个输入层结点数目、 $K$ 个隐藏层结点和 $n$ 个输出层结点数目,则神经网络的时间复杂度为 $O(d \times K)$ . 使用 $k$ -means++的时间复杂度为 $O(d \times c \times f)$ ,其中 $c$ 是簇类数目. 划分数据获得阈值 $(\alpha_i, \beta_i)$ 的时间复杂度是 $O(n^2 \times m \log m)$ . 在第 $i$ 粒度层计算条件概率的时间复杂度为 $O(n^2 \times m)$ ,计算风险损失 $R^{(i)}$ 的时间复杂度为 $O(n^2 \times m)$ . 假设所有的样本在 $w$ 次循环完成学习过程,则在STWD-SFNN-PR的序贯三支决策部分的时间复杂度为 $O(n^2 \times m \times w)$ . 因此,STWD-SFNN-PR的时间复杂度为 $O(d \times (K + c \times f) + n^2 \times m \times w)$ .

## 3 实验

### 3.1 数据集和评价指标

本文使用推荐领域经典的电影推荐数据集,即MovieLens-100k、MovieLens-1m和MovieLens-10m. 为了方便起见,分别将其简记为ML-100k、ML-1m和ML-10m. 由于不同用户评分数目的差异性较大,为了统一实验标准,选取评分数目在[20, 150]范围之间

的全部用户作为研究的标准用户,并在标准用户的样本集上进行个性化推荐. 数据集的统计信息如表1所示. 选取[20,150]范围的原因如下: 1) 根据6:2:2的数据集划分比例,每位标准用户的测试集至多包含30条待推荐样本,从而可以满足top5~top30推荐列表的实验设置; 2) 由表1可知,由标准用户构成样本集的稀疏度比原数据集具有更高的稀疏度. 这里稀疏度的计算方法如下: 以ML-100k为例,本文选取了717个标准用户对1428部电影的共42529个评分,因此稀疏度为 $1 - 42529 / (717 \times 1428) = 95.85\%$ .

表1 数据集的统计信息

数据集	用户数	项目数	记录数	稀疏度/%	原稀疏度/%
ML-100k	717	1428	42529	95.85	93.70
ML-1m	3956	3455	257937	98.11	95.75
ML-10m	51729	9510	3061314	99.38	98.69

本文按照6:2:2的比例将数据集分别划分为训练集、验证集和测试集,其中训练集用于模型的学习过程,验证集是针对训练集上学习的结果,验证算法的泛化能力,测试集是用于输出模型的预测值. 同时,本文使用5次5折交叉验证法选择STWD-SFNN-PR的最优参数,并返回最优参数下的评价指标. 本文使用的评价指标分为两类: 1) 推荐项目准确性的评估,选取基于精确率( $P@N$ )和召回率( $R@N$ )调和平均的 $F_1@N$ ; 2) 针对推荐列表顺序性的评估,选取平均精度均值( $\text{MAP}@N$ ). 计算公式如下所示:

$$F_1@N = \frac{2 \times P@N \times R@N}{P@N + R@N},$$

$$P@N = \sum_{u=1}^{|U|} |R_u \cap T_u| / \sum_{u=1}^{|U|} |R_u|,$$

$$R@N = \sum_{u=1}^{|U|} |R_u \cap T_u| / \sum_{u=1}^{|U|} |T_u|,$$

$$\text{MAP}@N = \frac{1}{|U|} \times \sum_{u=1}^{|U|} \text{AP}_u@N,$$

$$\text{AP}_u = \frac{1}{|T_u|} \times \sum_{i \in T_u} \frac{\sum_{j \in T_u} h(p_{uj} < p_{ui}) + 1}{p_{ui}}.$$

其中: $R_u$ 表示第 $u$ 位用户的推荐项目列表; $T_u$ 表示第 $u$ 位用户喜欢的项目列表; $N$ 表示推荐项目列表的长度; $p_{ui}$ 表示项目 $i$ 在用户 $u$ 的推荐列表中的排序位置; $p_{uj} < p_{ui}$ 表示在用户 $u$ 的推荐列表中,项目 $j$ 的排序位置位于项目 $i$ 的前方.  $F_1@N$ 和 $\text{MAP}@N$ 的值越大,表明模型的推荐质量越好.

### 3.2 对比算法

为了验证 STWD-SFNN-PR 的可行性, 选用 3 大类共 7 种算法进行对比: 经典的神经网络推荐、经典的深度学习推荐和最新的三支协同过滤推荐. 对比方法的基本原理如下所示.

1) SFNN<sup>[13]</sup>: 单隐藏层的前馈型神经网络算法, 利用经验公式法确定隐藏层的结点数目, 在训练过程中学习推荐系统的用户与项目之间的非线性关系, 从而产生非线性决策边界.

2) AutoRec<sup>[5]</sup>: 单隐藏层的神经网络推荐算法, 利用协同过滤中的共现矩阵完成模型对项目的自编码, 再利用自编码的泛化能力得到用户对项目的预测评分, 从而进行推荐排序.

3) FNN<sup>[6]</sup>: 因子分解机支持的神经网络, 使用因子分解机训练好的各特征的隐向量初始化神经网络嵌入层的参数, 相当于在初始化神经网络参数时引入了先验知识.

4) NeuralCF<sup>[7]</sup>: 基于神经网络的协同过滤算法, 利用多层的神经网络替代协同过滤的内积操作, 即利用多层的神经网络学习用户和项目的交互函数, 从而引入更多的非线性特征.

5) DTWGR-SVD<sup>[20]</sup>: 基于三支决策的奇异值分解模型, 利用奇异值分解在每个粒度层补全评分矩阵并计算预测概率, 再将概率大于序贯参数  $\alpha$  的  $N$  个项目作为推荐项目, 从而完成 top- $N$  推荐.

6) DTWGR-SVD++<sup>[20]</sup>: 基于三支决策的奇异值分解改进模型, 在 DTWGR-SVD 推荐算法的基础上,

DTWGR-SVD++ 在向用户推荐项目时进一步考虑了用户隐式反馈对用户评分的影响.

7) DTWGR-NFM<sup>[20]</sup>: 基于三支决策的非负矩阵分解模型, 在每个粒度层的矩阵分解中加入非负的约束条件, 补全评分矩阵并计算预测概率, 再将概率大于序贯参数  $\alpha$  的  $N$  个项目作为推荐项目.

### 3.3 实验结果分析

为了评估 STWD-SFNN-PR 的有效性, 设置 3 个研究问题 (research question, RQ):

RQ<sub>1</sub>: 与经典的神经网络推荐模型相比, STWD-SFNN-PR 的推荐质量如何?

RQ<sub>2</sub>: 与经典的深度网络推荐模型相比, STWD-SFNN-PR 的推荐质量如何?

RQ<sub>3</sub>: 与最新的三支协同过滤模型相比, STWD-SFNN-PR 的推荐质量如何?

#### 3.3.1 与经典的神经网络推荐模型的对比 (RQ<sub>1</sub>)

为了验证 STWD-SFNN-PR 在个性化推荐时考虑延迟推荐的可行性, 选择二支决策的 SFNN<sup>[13]</sup> 作为对比模型. 对比模型的参数设置如下所示: 激活函数的类型选择 Sigmoid 类或 ReLU 类; 神经网络的连接权重和偏置服从均匀分布或正态分布. 同时, 利用 3 种经验公式法确定对比模型的隐藏层结点数目, 并分别命名为  $m_1$ -SFNN、 $m_2$ -SFNN 和  $m_3$ -SFNN, 且  $m_1 = \sqrt{m+n} + \alpha$  ( $\alpha \in (1, 10)$ ),  $m_2 = \log_2 m$ ,  $m_3 = \sqrt{mn}$ , 其中  $m$ 、 $n$  和  $m_i$  ( $i = 1, 2, 3$ ) 分别为输入层、输出层和隐藏层的结点数目. 图 2 展示了 STWD-SFNN-PR 与 SFNN 的推荐质量关于推荐列表长度的曲线图.

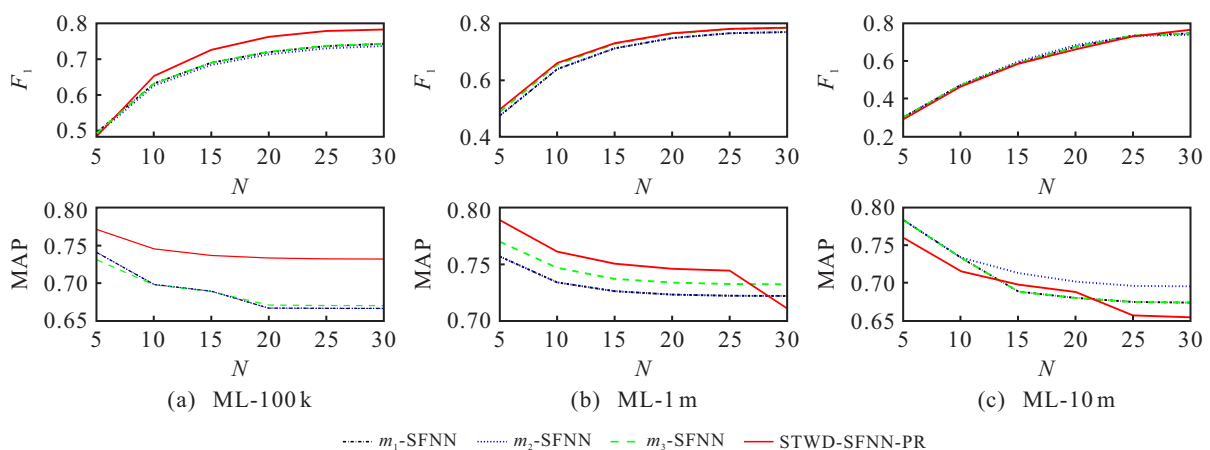


图 2 与经典的神经网络推荐模型的对比

通过上述实验结果可以发现: 在不同的推荐列表长度上, STWD-SFNN-PR 具有较好的推荐质量. 图 2 显示 STWD-SFNN-PR 的  $F_1$  曲线位于对比模型的上方. 例如, 在 ML-100k 数据集上, STWD-SFNN-

PR 的  $F_1@20$  为  $0.763 \pm 0.002$ , 对比模型  $m_1$ -SFNN、 $m_2$ -SFNN 和  $m_3$ -SFNN 的  $F_1$  分别为  $0.720 \pm 0.005$ 、 $0.720 \pm 0.005$  和  $0.721 \pm 0.005$ , 且 STWD-SFNN-PR 比表现最优的  $m_3$ -SFNN 提高了约 5.82%. 此外, 在多

数情况下, STWD-SFNN-PR的MAP曲线位于对比模型的上方. 例如, 在ML-1m数据集上, STWD-SFNN-PR的MAP@10为 $0.761 \pm 0.003$ , 对比模型MAP表现最优的是 $m_3$ -SFNN且值为 $0.747 \pm 0.002$ , STWD-SFNN-PR比 $m_3$ -SFNN的推荐质量提高了约1.87%. 因此, STWD-SFNN-PR在不同数据集的 $F_1$ 和MAP优于经典SFNN模型.

造成这一现象的原因分析如下: STWD-SFNN-PR在推荐过程中采用了三支决策的策略, 可以对推荐样本进行更有针对性地学习, 而SFNN模型没有. 对于易于推荐的实例, 由于网络的自学习能力, STWD-SFNN-PR能够快速获得较好的 $F_1$ 和MAP; 对于难以推荐的实例, STWD-SFNN-PR是将其划分至延迟决策域, 通过不断地补充额外的信息, 逐步实现所有难分样本的划分. 然而, SFNN模型不考虑样本划分的难易程度, 在每轮次需要学习数据集中的

所有样本的特征, 不利于提高模型的推荐质量. 因此, STWD-SFNN-PR的推荐质量优于经典的SFNN模型.

### 3.3.2 与经典的深度学习推荐模型的对比(RQ<sub>2</sub>)

为了验证STWD-SFNN-PR在个性化推荐时考虑延迟推荐的有效性, 选择AutoRec<sup>[5]</sup>、FNN<sup>[6]</sup>和NeuralCF<sup>[7]</sup>作为对比模型. 对比模型的主要超参数设置如下所示: 优化器为Adam, 优化步长为1024, 优化学习率为 $10^{-3}$ ,  $L_2$ 正则化系数为 $10^{-4}$ . 此外, AutoRec的隐藏层结点数目为100. FNN的嵌入层向量维度为8, 两个全连接层的隐藏层结点数目分别设置为[128, 128]. NeuralCF的双线性池化层的失活率为0.5, 嵌入层的向量维度为8, 两个全连接层的隐藏层结点数目分别设置为[128, 128]. 图3展示了STWD-SFNN-PR与神经网络推荐的推荐质量关于推荐列表长度的曲线图.

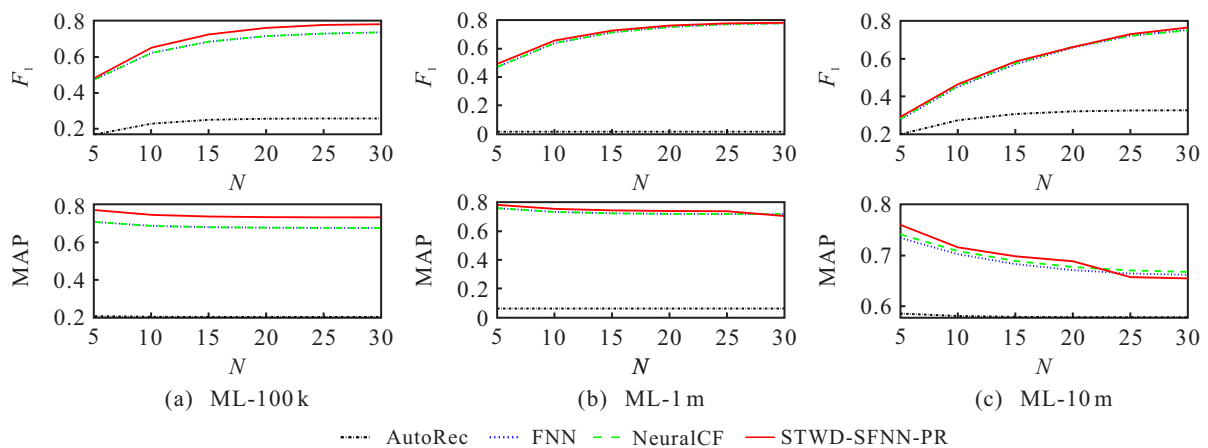


图3 与经典的深度学习推荐模型的对比

通过上述实验结果可以发现: 在不同的推荐列表长度上, STWD-SFNN-PR在绝大多数情形下具有较好的推荐质量. 图3显示STWD-SFNN-PR的 $F_1$ 曲线位于对比模型的上方, 且在ML-100k数据集上的差异更加显著. 例如, 在ML-1m数据集上, STWD-SFNN-PR的 $F_1@30$ (即 $0.785 \pm 0.001$ )比对比模型表现最优FNN的 $F_1@30$ (即 $0.781 \pm 0.002$ )提高了约0.51%. 类似地, 在ML-100k数据集上, STWD-SFNN-PR的 $F_1@30$ 为 $0.784 \pm 0.001$ , 对比模型表现较优的是FNN且 $F_1$ 为 $0.738 \pm 0.004$ , 此时STWD-SFNN-PR比对比模型提高了约6.23%. 此外, STWD-SFNN-PR在绝大多数情况下的MAP曲线位于对比模型的上方. 例如, 在ML-1m数据集上, STWD-SFNN-PR的MAP@10为 $0.761 \pm 0.003$ , 对比模型MAP表现最优的是NeuralCF且MAP@10为 $0.740 \pm 0.002$ , STWD-SFNN-PR比NeuralCF提高了约2.84%. 因此, STWD-

SFNN-PR在不同数据集的 $F_1$ 和MAP优于主流的神经网络推荐模型.

造成这一现象的原因分析如下: STWD-SFNN-PR在推荐过程中采用了序贯的三支策略, 可以更有针对性地学习难以推荐的样本, 而经典的深度学习推荐模型则没有. STWD-SFNN-PR在推荐过程中考虑延迟推荐, 通过设置延迟决策域的方式可以反复学习难以推荐样本的特征. 在当前网络拓扑结构的基础上, 在延迟决策域不为空的条件下增加隐藏层的结点数目, 从而补充更多有关难以推荐样本的信息. 此外, 结合在不同的粒度层设置的序贯阈值参数, STWD-SFNN-PR可以在有限轮次内实现难以推荐的样本划分. 然而, 经典的深度学习推荐模型不考虑推荐样本划分时的难易程度. 同时, AutoRec和NeuralCF仅使用用户ID和项目ID, 没有考虑更多有关用户和项目的其他特征; FNN只能学习高阶特征组合, 忽略了对

低阶特征建模. 因此, STWD-SFNN-PR的推荐质量在多数情况下优于对比模型.

### 3.3.3 与最新三支协同过滤推荐模型的对比(RQ<sub>3</sub>)

为了验证 STWD-SFNN-PR 在个性化推荐时考虑样本间非线性关系的有效性, 选择 DTWGR-SVD<sup>[20]</sup>、DTWGR-SVD++<sup>[20]</sup> 和 DTWGR-NFM<sup>[20]</sup> 作为对比模型. 对比模型的主要超参数设置如下所示: 粒度层的上界为 10, 优化器为 Adam, 优化步长为 1024, 优化学习率为  $10^{-3}$ ,  $L_2$  正则化系数为  $10^{-4}$ . 图4展示了 STWD-SFNN-PR 与三支决策推荐算法的推荐质量关于推荐列表长度的曲线图.

通过上述实验结果可以发现: 在不同的推荐列表长度上, STWD-SFNN-PR 在绝大多数情形下具有较好的推荐质量. 图4显示 STWD-SFNN-PR 的  $F_1$  曲线位于对比模型的上方. 例如, 在 ML-100k 数据集上, STWD-SFNN-PR 的  $F_1@20$  为  $0.763 \pm 0.002$ , 对比模型表现较优的是 DTWGR-NFM 且  $F_1@20$  为  $0.717 \pm 0.003$ , STWD-SFNN-PR 对比模型提高了约 6.42%. 此外, STWD-SFNN-PR 在绝大多数情况下的 MAP 曲线位于对比模型的上方, 且在 ML-100k 数据集上的差异更加显著. 例如, 在 ML-1m 数据集上, STWD-SFNN-PR 的 MAP@20 (即  $0.746 \pm 0.003$ ) 对比模型表现最优 DTWGR-NFM 的 MAP@20 (即  $0.727 \pm$

$0.003$ ) 提高了约 2.61%. 同时, 在 ML-100k 上, STWD-SFNN-PR 的 MAP@10 (即  $0.745 \pm 0.002$ ) 对比模型表现最优 DTWGR-NFM 的 MAP@10 (即  $0.690 \pm 0.009$ ) 提高了约 7.97%. 因此, STWD-SFNN-PR 在不同数据集上的  $F_1$  和 MAP 优于三支决策推荐模型.

造成这一现象的原因分析如下: 虽然 STWD-SFNN-PR 与三支决策推荐模型均是利用序贯三支决策指导推荐过程, 但它们仍存在不同之处. STWD-SFNN-PR 不仅使用用户对项目的评分数据, 还使用用户和项目的特征, 使得模型能挖掘更多隐藏的用户兴趣模式. 同时, STWD-SFNN-PR 是针对所有推荐样本的学习, 即所有目标用户的延迟决策域为空集作为模型停止学习的条件之一. 相应地, 在召回阶段, STWD-SFNN-PR 的召回对象是由推荐概率大于序贯阈值的项目构成的, 并按照推荐概率的大小选取前  $N$  个项目作为最终的推荐项目. 然而, DTWGR-SVD 等对比模型仅使用用户对项目的评分数据, 且是针对推荐列表长度的学习过程, 即向目标用户推荐的项目数目达到  $N$  时终止对比模型的学习. 相较而言, 对比模型难以发现丰富的推荐信息, 且召回对象仅是模型训练好的  $N$  个推荐项目, 具有一定的局限性. 因此, STWD-SFNN-PR 的推荐质量优于最新的三支推荐模型.

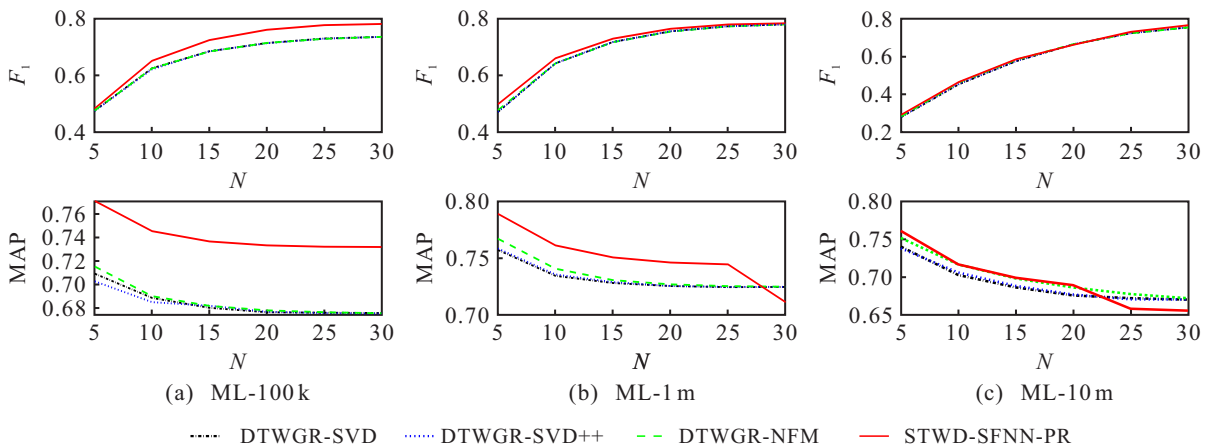


图4 与最新的三支决策推荐模型的对比

## 4 结论

为了综合考虑有关用户信息和项目信息等特征, 并避免二支决策的缺陷, 本文提出了一种新的推荐模型 STWD-SFNN-PR. 该模型基于网络拓扑结构比较简单的单隐藏层神经网络, 从而有效地降低了推荐模型的复杂度. 此外, 在目标用户的动态推荐过程中, STWD-SFNN-PR 采用序贯三支的策略, 使用嵌入处理高维稀疏的特征向量, 并将 SFNN 难以决策的推

荐样本划分至延迟决策域, 通过反复学习这类样本的特征, 可以有效地提高模型的泛化能力. 更重要的是, 在多种电影数据集上验证, 与二支决策的 SFNN、深度学习推荐和最新序贯三支指导下的协同过滤推荐的对比模型相比, 本文方法具有更高的推荐质量, 表明了 STWD-SFNN-PR 的可行性和有效性.

虽然 STWD-SFNN-PR 提高了在个性化推荐的质量, 但仍存在以下几点值得进一步研究:

1) STWD-SFNN-PR 融合了序贯三支决策思想和神经网络模型实现推荐, 这样的推荐模型略显复杂, 如何进一步简化推荐模型以提高其可操作性, 值得深入研究;

2) 在 3 个不同大小的电影推荐数据集上验证了 STWD-SFNN-PR 的推荐性能, 如何进一步探索 STWD-SFNN-PR 在其他更多数据集上的推荐性能, 值得深入探索;

3) 推荐系统存在冷启动和数据稀疏性等问题, 如何采用 STWD-SFNN-PR 的原理, 有效地解决推荐系统中存在的这些问题, 值得深入思考。

#### 参考文献(References)

- [1] 胡春华, 赵慧, 童小芹, 等. 推荐系统对消费者网购支出的影响研究[J]. 中国管理科学, 2020, 28(6): 158-170.  
(Hu C H, Zhao H, Tong X Q, et al. Research on the impact of intelligent recommendation system on consumer online shopping[J]. Chinese Journal of Management Science, 2020, 28(6): 158-170.)
- [2] Duan R, Jiang C Q, Jain H K. Combining review-based collaborative filtering and matrix factorization: A solution to rating's sparsity problem[J]. Decision Support Systems, 2022, 156: 113748.
- [3] He X R, Pan J F, Jin O, et al. Practical lessons from predicting clicks on ads at facebook[C]. Proceedings of the 8th International Workshop on Data Mining for Online Advertising. New York, 2014: 1-9.
- [4] Pang G Y, Wang X M, Hao F, et al. ACNN-FM: A novel recommender with attention-based convolutional neural network and factorization machines[J]. Knowledge-Based Systems, 2019, 181: 104786.
- [5] Sedhain S, Menon A K, Sanner S, et al. AutoRec: Autoencoders meet collaborative filtering[C]. Proceedings of the 24th International Conference on World Wide Web. Florence, 2015: 111-112.
- [6] Zhang W N, Du T M, Wang J. Deep learning over multi-field categorical data[C]. European Conference on Information Retrieval. Cham: Springer, 2016: 45-57.
- [7] He X N, Liao L Z, Zhang H W, et al. Neural collaborative filtering[C]. Proceedings of the 26th International Conference on World Wide Web. New York: ACM, 2017: 173-182.
- [8] 张文, 崔杨波, 李健, 等. 基于聚类矩阵近似的协同过滤推荐研究[J]. 运筹与管理, 2020, 29(4): 171-178.  
(Zhang W, Cui Y B, Li J, et al. F-cluMA: A collaborative filtering model based on clustered matrix approximation[J]. Operations Research and Management Science, 2020, 29(4): 171-178.)
- [9] Chen L, Cao J, Wang Y Q, et al. Multi-view graph attention network for travel recommendation[J]. Expert Systems With Applications, 2022, 191: 116234.
- [10] Cheng S H, Wu Y X, Li Y, et al. TWD-SFNN: Three-way decisions with a single hidden layer feedforward neural network[J]. Information Sciences, 2021, 579: 15-32.
- [11] Fang Y, Gao C, Yao Y Y. Granularity-driven sequential three-way decisions: A cost-sensitive approach to classification[J]. Information Sciences, 2020, 507: 644-664.
- [12] 张萌, 孙秉珍, 王婷, 等. 融合粗糙集与 GRA 的异构信息多准则三支推荐及其在医疗推荐中的应用[J]. 控制与决策, 2022, 37(7): 1883-1893.  
(Zhang M, Sun B Z, Wang T, et al. Multi-criteria three-way recommendation of heterogeneous information based on rough set and GRA and its application in medical recommendation[J]. Control and Decision, 2022, 37(7): 1883-1893.)
- [13] Belciug S. Logistic regression paradigm for training a single-hidden layer feedforward neural network. Application to gene expression datasets for cancer research[J]. Journal of Biomedical Informatics, 2020, 102: 103373.
- [14] Zhang X Y, Gou H Y, Lv Z Y, et al. Double-quantitative distance measurement and classification learning based on the tri-level granular structure of neighborhood system[J]. Knowledge-Based Systems, 2021, 217: 106799.
- [15] 向露, 朱军楠, 周玉, 等. 基于多粒度对抗训练的鲁棒跨语言对话系统[J]. 自动化学报, 2021, 47(8): 1855-1866.  
(Xiang L, Zhu J N, Zhou Y, et al. Robust cross-lingual dialogue system based on multi-granularity adversarial training[J]. Acta Automatica Sinica, 2021, 47(8): 1855-1866.)
- [16] Zhan J M, Ye J, Ding W P, et al. A novel three-way decision model based on utility theory in incomplete fuzzy decision systems[J]. IEEE Transactions on Fuzzy Systems, 2022, 30(7): 2210-2226.
- [17] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [18] Tong Q Q, Liang G N, Bi J B. Calibrating the adaptive learning rate to improve convergence of ADAM[J]. Neurocomputing, 2022, 481: 333-356.
- [19] Öztürk M M, Cavusoglu U, Zengin A. A novel defect prediction method for web pages using k-means++[J]. Expert Systems With Applications, 2015, 42(19): 6496-6506.
- [20] Liu D, Ye X Q. A matrix factorization based dynamic granularity recommendation with three-way decisions[J]. Knowledge-Based Systems, 2020, 191: 105243.

#### 作者简介

成淑慧(1991—), 女, 博士生, 从事机器学习的研究, E-mail: shuhui\_cheng@163.com;

武优西(1974—), 男, 教授, 博士生导师, 从事数据挖掘、机器学习等研究, E-mail: wuc567@163.com.