



基于需求预测的主动服务推荐方法

刘志中¹, 张振兴¹, 海燕², 郭思慧¹, 刘永利¹

(1. 河南理工大学 计算机科学与技术学院, 河南 焦作 454002; 2. 华北水利水电大学 信息工程学院, 郑州 450045)

摘 要: 在智能计算领域, 网络中可用服务数量与类型的快速增长, 使用户更依赖于服务完成各种业务, 然而当前“请求-响应”被动式的服务模式严重影响了用户体验与资源利用率。为智能感知用户需求并主动为用户推荐合适的服务, 通过引入需求预测过程, 提出一种主动服务推荐方法。利用矩阵分解算法从大量历史服务使用数据中提取用户特征和服务特征, 据此训练深度学习模型并预测用户的服务需求, 进而为用户推荐其所需要的服务。基于真实数据的实验结果表明, 该方法较单一的矩阵分解模型和深度神经网络模型具有更高的服务推荐准确性和稳定性。

关键词: 需求预测; 主动服务; 服务推荐; 矩阵分解; 深度学习

开放科学(资源服务)标志码(OSID):



中文引用格式: 刘志中, 张振兴, 海燕, 等. 基于需求预测的主动服务推荐方法[J]. 计算机工程, 2020, 46(2): 96-102.

英文引用格式: LIU Zhizhong, ZHANG Zhenxing, HAI Yan, et al. Active service recommendation method based on requirement prediction[J]. Computer Engineering, 2020, 46(2): 96-102.

Active Service Recommendation Method Based on Requirement Prediction

LIU Zhizhong¹, ZHANG Zhenxing¹, HAI Yan², GUO Sihui¹, LIU Yongli¹

(1. College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan 454002, China;

2. College of Information Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450045, China)

[Abstract] In the intelligent computing field, the rapid growth of available Internet services makes users increasingly dependent on services to complete various businesses, but the passive “request-response” service model seriously decreases user experience and resource utilization. To intelligently perceive user requirements and proactively recommend appropriate services to users, this paper proposes a method of active service recommendation based on user requirement prediction. The method firstly extracts user features and service features from massive data of historical services by using matrix factorization. On this basis, the extracted data is used to train the deep learning model and predict service demands of users, so as to recommend appropriate services to users. Experimental results on real data show that the proposed method has higher accuracy and stability of service recommendation than simply a matrix factorization model or deep neural network model.

[Key words] requirement prediction; active service; service recommendation; Matrix Factorization (MF); deep learning
DOI: 10.19678/j.issn.1000-3428.0053379

0 概述

随着服务计算、云计算、边缘计算以及移动计算等新型计算模式的快速发展, 网络上出现了大量跨网跨域的可用服务, 同时无线网络和智能终端的快速发展与普及, 也使更多的用户可以随时随地调用各个领域的服务来完成工作业务与生活事务。

然而在海量服务为用户带来方便的同时, 也导致了服务发现与服务选择难题。此外, 当前服务系统通常采用“请求-响应”被动式的服务模式来处理用户的服务需求, 该模式缺乏智能性与主动性, 严重影响了服务资源的利用率与用户体验, 不能满足当前大数据与人工智能时代下用户对服务应用的体验要求。

基金项目: 国家自然科学基金(61872126, 61772159)。

作者简介: 刘志中(1981—), 男, 副教授、博士, 主研方向为服务计算、智能服务; 张振兴, 硕士研究生; 海燕, 教授、博士; 郭思慧, 硕士研究生; 刘永利, 副教授、博士。

收稿日期: 2018-12-11

修回日期: 2019-04-12

E-mail: lzzmff@126.com

针对当前被动服务模式存在的不足,研究者提出主动服务的思想^[1-2],以使大规模用户能够更快捷地获取到其迫切需要的服务。主动服务现已成为服务应用领域的研究热点^[3],其是一种运行于 Web 上的新型信息服务模式,可以解决服务信息过载以及用户需求多变而带来的服务可扩展性问题,在当前动态、互联、智能的信息环境下,主动捕获用户的个性化服务需求,并为用户提供满足其个性化需求的服务^[4-5]。以智慧养老服务为例,老人在发现服务、使用服务等方面比常人存在更多的困难,而主动服务模式则可以智能、实时地感知老人的需求,主动为老人推荐亟需使用并且最适合其自身情况的服务。因此,深入开展主动服务相关技术的研究,具有重要的理论意义与应用价值。

主动服务推荐是实现主动服务的主要技术之一^[6-7]。目前,国内外学者针对服务推荐问题展开了大量的研究工作^[8-10],其中部分基于协同过滤^[7,11]、内容^[8,12]或关联规则^[13-14]为用户推荐服务,部分则基于服务质量(Quality of Service, QoS)预测为用户推荐服务^[10,15]。然而,当前的研究工作多数仅关注对服务模式的研究或基于 QoS 预测的服务推荐,忽略了服务推荐过程中用户需求对服务推荐效果的影响,导致服务推荐的准确性较差。

针对上述问题,本文提出一种基于用户需求预测的主动服务推荐方法。利用矩阵分解(Matrix Factorization, MF)算法从大量历史服务使用数据中快速地提取用户特征和服务特征,基于此训练深度学习算法实现对用户服务需求的精确预测,并根据所预测的服务需求选择最合适的服务主动推荐给用户。

1 相关工作

近年来,研究者针对主动服务开展了一系列的研究。文献[16]提出一种在移动互联网下基于情景感知的透明主动服务架构,利用传感器设备对用户所处情景信息进行收集并发送给云服务器端,经过服务器端的服务选择算法处理数据,最终返回给用户一个合适的服务推荐列表。该文以智能手机代替传感器设备,在云服务器端进行了模拟实验,验证了所设计架构的有效性。文献[17]提出一种基于熟人模型与最小二乘支持向量机的主动服务选择策略,利用最小二乘支持向量机预测随机服务列表,并基于熟人模型进行服务分配。实验结果表明,该服务选择策略不仅缩短了通信时间,而且还减少了服务组合的执行时间,显著提高了服务选择的成功率。

文献[4]在物联网环境下对服务模式进行研究,通过应用模糊逻辑对服务资源进行等级划分和评估,提出一种三级结构的主动服务资源选择策略,并在模拟物联网的分布式环境下,验证其性能优于随机策

略。文献[5]针对当前车联网环境下被动服务模式带来的误报、延报等问题,设计一种多层次感知的主动行车服务模型,其中每一层次都作为一个智能体存在。实验仿真验证了该模型的主动性与高效性。

此外,文献[18]设计了一种基于规则归纳推理的推荐机制来预测用户需求,利用智能手机收集用户的情景信息以及该情景下用户的活动信息,并基于收集到的信息建立<时间,条件,行为>规则模型。近年来研究者尝试引入新的方法和技术对主动服务进行创新和改进。文献[19]提出一种将边缘化去噪自动编码器(mDA)与隐含因子模型相结合的混合推荐算法。通过利用 mDA 算法引入辅助信息来改善传统 CF 模型的数据稀疏性,并在4个不同数据集上验证了其性能的提升。文献[20]将概率图模型与深度信念网络结合,提出一种在线音乐推荐方法。文献[21]使用循环神经网络提取文本数据的隐含特征以提升协同过滤算法的性能,并通过在2个不同稀疏程度的数据集上验证结果,证明了其显著的性能优势。文献[22]将层叠降噪自动编码器与隐含因子模型相结合,综合考虑用户评分信息和评论文本信息。在亚马逊公开数据集上验证结果表明,该混合推荐算法对评分预测的准确性高于传统推荐模型。

上述工作分别从服务模式、服务架构等方面对服务推荐进行研究,但未从用户需求预测的角度出发,缺乏对用户需求的主动感知,导致服务推荐的主动性与准确性较低。因此,本文提出一种基于用户需求预测的主动服务推荐方法,以弥补当前研究工作存在的不足。

2 基于矩阵分解与深度学习的用户需求预测

用户需求预测是实现主动服务推荐的关键。在网络服务平台上,存在着大量关于用户使用服务的数据,这些数据中隐藏着用户和服务的重要特征,可以通过数据挖掘与机器学习方法来获这些规则特征,从而预测用户未来对不同类型服务的需求概率。用户需求预测为实现主动的服务推荐提供了重要的信息支持。

基于用户使用服务的历史数据可以生成<用户-服务>矩阵。然而,<用户-服务>矩阵具有很强的数据稀疏性,严重影响了用户服务需求预测的准确性。针对这一问题,本文采用隐语义模型的矩阵分解方法,从大量用户使用服务的历史数据中快速提取用户特征和服务特征,形成用户特征矩阵和服务特征矩阵,从而有效解决数据稀疏性问题。在此基础上,利用用户特征矩阵与服务特征矩阵训练深度学习模型并完成对用户需求的预测。本文方法流程如图1所示。图中实线部分是对用户需求的正向预测,虚线部分则是利用预测值与真实值的误差反方向训练用户需求预测模型。

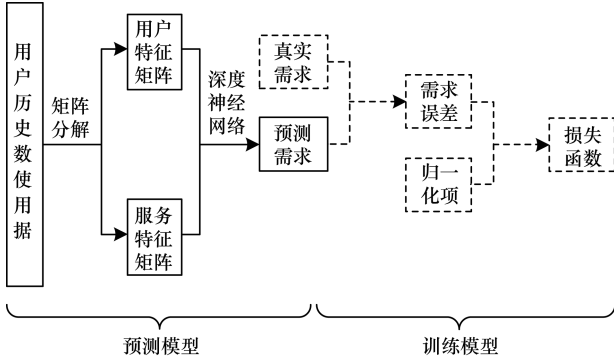


图 1 本文方法流程

Fig. 1 Procedure of the proposed method

本文方法能够弥补矩阵分解方法在特征提取方面的不足,其具有以下特点:

- 1) 从用户与服务的交互矩阵中提取每个用户与服务的向量,用户向量能够表征用户与所有服务的不同映射关系。
- 2) 服务向量能够表征服务与所有用户的不同关系映射。
- 3) 利用深度神经网络高效的特征提取能力,通过 3 个隐层不断向上的特征抽取,获得具有高表征能力的用户向量和服务向量表示。

2.1 基于矩阵分解的特征提取

矩阵分解 (MF) 是一种高效的特征提取方法^[19,23]。基于 MF 可以把用户与服务的交互矩阵分解为用户和服务特征的共同潜在空间,进而分别发现用户对服务需求的潜在因素。

给定用户与服务的交互矩阵 $R = (r_{u,i})_{U \times I}$, 矩阵的行表示一位用户对所有类型服务的历史使用情况,列则表示关于某一种服务的所有用户的历史使用情况。经过矩阵分解分别得到用户和服务的低秩矩阵 $M_{U \times D}$ 和 $N_{I \times D}$, 其中,矩阵 $M_{U \times D}$ 是 U 个用户的 D 维潜在空间表示,而 $N_{I \times D}$ 是 I 类服务的 D 维潜在空间表示,且矩阵 $M_{U \times D}$ 和 $N_{I \times D}$ 的乘积能还原成原始矩阵 $R = (r_{u,i})_{U \times I}$,即可以通过矩阵还原预测用户 u 对服务 i 的需求 $r'_{u,i}$:

$$r'_{u,i} = M_u N_i^T \quad (1)$$

从用户和服务的整体范围考虑,每个用户有自己的服务使用习惯,每个服务具有自身的服务特征,因此,也可以使用带有平均偏差的预测模型进行矩阵还原操作。对于每个用户服务,分别引入一个偏移量 b_u 和 b_i ,基础部分是所有需求数值的平均值 a ,考虑平均偏差的矩阵分解预测公式如下:

$$r'_{u,i} = a + b_u + b_i + M_u N_i^T \quad (2)$$

为尽可能准确地表达用户和服务特征,本文定义一个损失函数 f_{loss} 来衡量预测值与真实值之间的差距,并通过最小化损失函数来预测和修改参数。

损失函数的定义有很多种,本文通过计算预测值与真实值的均方误差作为损失函数,设用户对服务的所有交互集合为 $H = \{(u, i, r)\}$, 则最小化损失

函数如式(3)所示。

$$f_{\text{loss}} = \sum_{(u,i) \in H} e_{u,i}^2 \quad (3)$$

其中:

$$e_{u,i}^2 = (r_{u,i} - r'_{u,i})^2 = (r_{u,i} - \sum_{k=1}^D m_{u,k} n_{k,i})^2 \quad (4)$$

基于矩阵分解的用户特征与服务特征提取算法描述如下:

算法 1 基于矩阵分解的特征提取算法

输入 用户列表 U , 服务类型列表 I , 用户对不同服务类型的历史需求记录列表 R

输出 用户特征矩阵 M , 服务特征矩阵 N

1. Begin

2. 构造用户和服务初始特征

3. For $u = 1 \rightarrow U$ Do//遍历用户列表 U 中的每位用户 u

4. $r(u) = \text{embed}(1, 8)$

//通过嵌入矩阵方法初始化用户特征,维度为 1×8

5. $M = r(U)$ //矩阵 M 是由所有用户特征组成的

6. End For

7. For $i = 1 \rightarrow I$ Do//遍历每个服务类型 i

8. $r(i) = \text{embed}(1, 8)$ //初始化服务特征,维度为 1×8

9. $N = r(I)$ //矩阵 N 是由所有服务特征组成的

10. End For

11. 预测用户需求

12. 通过式(1)预测用户 u 对服务 i 的需求 $r'_{u,i}$

13. 矩阵分解

14. 通过式(3)、式(4)计算损失函数

15. $M, N = \text{gd}(f_{\text{loss}})$ //采用梯度下降法优化损失函数,得到用户特征矩阵 M 和服务特征矩阵 N

16. 迭代训练

17. For $j = 1 \rightarrow 1500$ Do//迭代次数

18. 重复步骤 11

19. 重复步骤 13

20. End For

21. End

2.2 基于深度神经网络的用户需求预测

深度神经网络 (Deep Neural Network, DNN) 是深度学习的通用型神经网络架构,其目的是通过确定性学习将原始数据表达为隐藏性表示。DNN 通过对原始数据进行重构输入,从数据中学习到高表征的输入和输出的关系映射。近年来, DNN 已在多个应用领域取得优异成绩^[24-25]。本文利用 DNN 3 个隐层的数学运算学习特征表达输出,将前一层网络的输出作为下一层的输入,其表达形式如式(5)所示。

$$a^l = \sigma(w^l a^{l-1} + b^l) \quad (5)$$

其中: a 是 DNN 的输出,表示预测的用户需求概率值; l 表示网络的层数; w 和 b 分别表示对应的权重矩阵和偏倚向量; σ 是激活函数。本文采用 ReLU 函数作为激活函数,该函数的单侧抑制性使神经网络中的神经元具有稀疏激活性,在线性函数中有更强的表达能力,而对于非线性函数, ReLU 函数非负区间为正值的特性,也解决了神经网络反向传播求解参数时经常会出现的梯度消失的问题。ReLU 激活函数如式(6)所示。

$$\text{ReLU}(a) = \begin{cases} a, & a > 0 \\ 0, & a \leq 0 \end{cases} \quad (6)$$

在 DNN 的输出层, 本文采用 Softmax 函数对输出结果进行分类, 进而得到用户对每种服务类型的需求概率。Softmax 函数模型可以看作是 Logistic 函数模型的推广, 其本质是把一个 n 维的向量映射成另一个具有相同维度表示的向量, 主要用于多分类问题, 对于不同类别的输入, 分别用一个概率值表示输出, 且所有类别的概率和为 1。Softmax 函数表达式如下:

$$\sigma_j(a) = \frac{a^{z_j}}{\sum_{k=1}^K a^{z_k}} \quad (7)$$

其中, $j=1, 2, \dots, K$ 。DNN 通过反向传播求解最适合的神经网络参数, 也需要选择一个损失函数来度量神经网络训练的输出和真实值的误差。在神经网络中, 损失函数的选取直接影响网络的训练输出结果, 本文采用均方差函数和交叉熵函数之和作为损失函数。设定真实值为 y , 预测值为 a , 下标 i 表示第 i 个样本, w_i 是权重, 则神经网络每次迭代训练的均方差可以表示为:

$$M_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^m w_i (y_i - a_i)^2 \quad (8)$$

交叉熵是用来度量神经网络预测用户需求实际输出和期望值接近程度的一种方法, 交叉熵的值越小, 表示实际输出和期望值越接近。交叉熵函数的公式表示为:

$$C = -\frac{1}{n} \sum_i [y \ln a + (1-y) \ln(1-a)] \quad (9)$$

利用 DNN 进行预测时, 为避免过拟合, 在损失函数中引入正则化惩罚项, 通过控制学习参数, 防止网络学习到特别复杂的模型, 减小过拟合发生的概率, 引入了正则化的损失函数如下:

$$f_{\text{loss2}} = M_{\text{MSE}} + C + \frac{\lambda}{2} \sum_w W^2 \quad (10)$$

其中, W 表示所有学习参数, λ 为一个可调节的参数, 用来控制正则化对损失函数的影响。基于 DNN 的用户需求预测算法描述如下:

算法 2 基于 DNN 的用户需求预测算法

输入 用户特征矩阵 M , 服务特征矩阵 N

输出 用户 u 对不同类型服务 i 的需求概率 $r'_{u,i}$

1. Begin

2. 初始化参数

3. For layer = 2 → 4 Do//DNN 的 3 个隐层

4. parameter[w, b] = init_param (layer) // 随机生成 // 初始权重值和偏置值, 3 个隐藏层中的节点个数分别为 // 64, 32 和 16

5. End For

6. DNN 前向传播

7. 加载矩阵 M, N 作为输入层的输入数据

8. 通过式 (5) ~ 式 (7) 预测用户 u 对每种类型服务 i 的需求概率 $r'_{u,i}$

9. DNN 反向传播

10. 通过式 (10) 计算损失函数

11. parameter[w, b] = gd(f_{loss}) // 采用梯度下降法优化 // 参数 w, b

12. 迭代训练

13. For $j = 1 \rightarrow 1\ 500$ Do//迭代次数

14. 重复步骤 6

15. 重复步骤 9

16. End For

17. End

3 基于需求预测的主动服务推荐

经过对用户服务需求的预测, 可以得到用户未来使用某一类服务的概率, 以此选取概率值较大的服务类作为该用户的服务需求, 并依据预测的用户服务需求, 确定满足该需求的候选服务集。在候选服务集内部, 按用户历史使用服务的频率并结合服务的评价数据, 选择综合评价最优的服务并主动地推荐给用户。

设满足用户 u 服务需求的候选服务集合为 $S = \{S_i, i \in 1, 2, \dots, n\}$, 其中, n 表示候选服务的总数。对于每个候选服务 S_i , 计算其针对用户 u 的使用频率 $f(S_i)$, 并计算每个候选服务的平均评价 $g(S_i)$, 在此基础上, 计算每个候选服务的综合推荐指数 $e(S_i)$, 将综合推荐指数最大的服务 S_i 推荐给用户, 如式 (11) 所示。

$$e(S_i) = \lambda f(S_i) + (1-\lambda) g(S_i) \quad (11)$$

其中, λ 是一个可变参数, 用于调整候选服务的评价值和使用频率在综合推荐指数中所占的比重。在推荐过程中, 对于每个用户都需要进行上述操作, 且每个候选服务的实际评价数值较大, 在进行综合推荐指数计算时要对其进行归一化处理。基于需求预测的服务推荐算法描述如下:

算法 3 基于需求预测的主动服务推荐算法

输入 用户 u 对每种类型服务 i 的需求概率 $r(u, i)$, 每个服务的平均评价 g

输出 推荐服务 S_i

1. Begin

2. 确定候选服务集合 s

3. $S \leftarrow \max(r(u, i))$

4. 计算服务使用频度

5. For $j = 1 \rightarrow \text{len}(S)$ Do

6. $f(S(j)) = \text{count}(S(j)) / \text{len}(S)$

7. EndFor

8. 推荐服务

9. 采用式 (11) 计算各服务的综合推荐指数

10. 将综合推荐指数最大的服务 S_i 推荐给用户 u

11. End

4 实验与结果分析

4.1 实验环境和数据集

为验证本文方法的有效性, 笔者进行了模拟实验。从公开的数据网站 MovieLens 获取到用户对电影的评分数据, 该实验把“电影”看作“服务”, 把“评

分”看作“需求”的概率值。选取 MovieLens 1M 数据集模拟用户对服务的需求,该数据包括约 6 000 名用户对 4 000 部电影产生的 100 万条评分数据,评分共分为 5 个等级,用 1~5 表示。为便于实验,对数据做如下处理:将用户对电影的前三个评分等级视为用户未使用过此服务,标记为 0;将后两个评分等级视为 1,表示用户使用过本服务。2 种交互类型数据之比约为 1:1.35,每位用户交互过的服务记录至少 20 条,以保证真实数据的独立性和稀疏性。

实验环境为个人电脑,操作系统:Windows 10 家庭版 64 位,CPU:Intel 酷睿 i5 4210M;RM:4 GB。实验应用 Python3 版本编程语言,基于谷歌开源深度学习框架 TensorFlow 1.6.0,在 Pycharm 社区版编程软件上完成。

4.2 实验评价指标

该实验将数据集随机划分成两部分,分别用于训练和测试。其中,数据集的 70% 作为训练数据,30% 用于对模型进行评估。在介绍评估方法前,首先引入混淆矩阵,如图 2 所示,其中, True、False 分别表示分类是否正确, Positive、Negative 分别表示正负样本的分类情况。

	Positive	Negative
True	True Positive(T_P)	True Negative(T_N)
False	False Positive(F_P)	False Negative(F_N)

图 2 混淆矩阵结构

Fig. 2 Confusion matrix structure

本文实验设置预测结果 1 为正样本,0 为负样本。评估方法如下:

1) 准确率,即总样本中分类正确的样本个数所占比重,表达式为:

$$A = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (12)$$

2) 精确率,即正样本预测结果中正确预测为正样本个数所占比重,表达式为:

$$P = \frac{T_P}{T_P + F_P} \quad (13)$$

3) 召回率,即正样本总数中被正确预测为正样本个数所占比重,表达式为:

$$R = \frac{T_P}{T_P + F_N} \quad (14)$$

由于分类问题是回归模型输出结果的离散化表现,因此本文还引入平均绝对误差和均方根误差 2 种回归模型的评估方法,用于衡量预测值与真实值的误差。

4) 平均绝对误差表达式为:

$$M_{MAE} = \frac{1}{N} \sum_{i=1}^N |(y_i - y'_i)| \quad (15)$$

5) 均方根误差表达式为:

$$R_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2} \quad (16)$$

4.3 实验结果与分析

根据实验评价指标,依次对比矩阵分解(MF)模型、DNN 模型和基于矩阵分解的 DNN 混合模型的性能,以 1 500 次迭代为机器学习终止条件,预测准确率对比如图 3 所示。可以看出,与 MF 模型相比,DNN 模型在预测准确率上效果并不理想,最终结果略低于 MF 模型。在 1 500 次的迭代训练中,MF 模型的预测准确率从 42.5% 左右开始逐步增加,最终稳定在 73.5%,DNN 模型的预测精确度则从 56.3% 左右增加至 73%。混合模型 MF + DNN 的表现要优于 2 种单一的模型,不仅比 MF 拥有较高的起点,并且最终准确率达到 78.8%,相比于 MF 和 DNN 分别提高了近 5% 和 6%,表明基于矩阵分解的 DNN 方法具有更高的用户需求预测准确性。

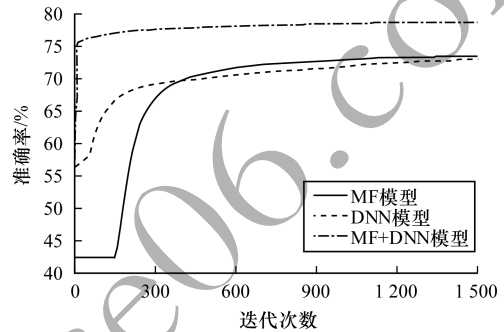


图 3 3 种模型的准确率对比

Fig. 3 Comparison of accuracy of three models

3 种模型的精确率和召回率对比如图 4 所示,可以看出,相比于 MF 方法,DNN 模型取得了较好的结果。在精确率的对比上,MF 为 75%,而 DNN 为 72%,两者相差仅为 3%,并无明显的差异。在召回率的对比上,DNN 为 86.6%,高出 MF 的 80.2% 约 6.5%。混合模型 MF + DNN 整体表现要优于 2 种单一的模型,其精确率为 79.85%,高于 MF 和 DNN,而召回率为 84.4%,仅低于 DNN 约 2.2%。上述实验结果表明,基于矩阵分解的 DNN 预测模型具有较强的预测能力,可弥补矩阵分解算法预测的不足。

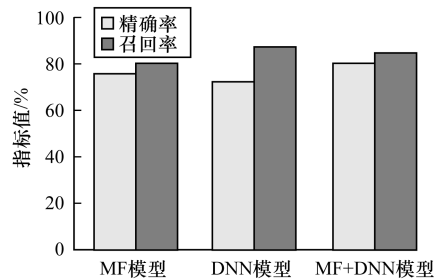


图 4 3 种模型的精确率和召回率对比

Fig. 4 Comparison of precision rate and recall rate of three models

3种模型的平均绝对误差和均方根误差对比如图5所示,可以看出,MF模型和DNN模型几乎相等,而混合模型MF+DNN效果要优于这2种模型,在平均绝对误差和均方根误差上均低于MF和DNN模型约0.05,可见由数据稀疏性导致的预测误差对MF+DNN模型影响较小,因此,该模型具有更好的稳定性。

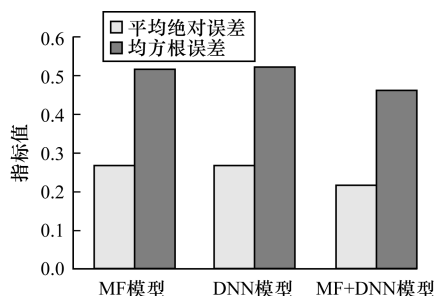


图5 3种模型的平均绝对误差与均方根误差对比

Fig.5 Comparison of MAE and RMSE of three models

3种模型的实验数据对比如表1所示。综合5项评价指标对比可以得出结论:基于矩阵分解的DNN模型性能优于单一使用的矩阵分解模型。

表1 3种模型的实验数据对比

Table 1 Comparison of experimental data of three models

模型	准确率 /%	精确率 /%	召回率 /%	平均绝对 误差	均方根 误差
MF模型	73.46	75.26	80.23	0.2653	0.5152
DNN模型	73.01	72.11	86.64	0.2696	0.5195
MF+DNN模型	78.77	79.85	84.40	0.2123	0.4608

5 结束语

为实现服务的智能化主动推荐,提高用户满意度与资源利用率,本文提出一种基于用户需求预测的主动服务推荐方法。该方法基于隐语义模型的矩阵分解算法,从服务应用历史数据中快速提取用户特征与服务特征,为精确预测用户的服务需求提供了更有效的数据,同时解决了数据稀疏性问题。此外,其基于深度学习算法建立用户与服务之间的非线性关系映射,有助于挖掘数据中隐含的特征,能够提高用户服务需求预测的准确性以及服务需求预测模型的可扩展性。由于本文模型仅基于用户的历史服务使用数据进行预测,并未考虑用户使用时的场景信息,因此下一步将研究情景感知的主动服务推荐方法。

参考文献

[1] JIN Xin, ZHONG Shiqiang, LI Youjiang, et al. A domain knowledge based method on active and focused information service for decision support within big data

environment [J]. Procedia Computer Science, 2015, 60(1):93-102.

[2] XU Xiangzheng, WANG Shiqi, WU Baihong. Research and implementation of active service for distribution network based on large data analysis [J]. Bulletin of Science and Technology, 2017, 33(6):105-108. (in Chinese)

徐祥征,王师奇,吴百洪.基于大数据分析的配电网主动检修业务应用研究与实现[J].科技通报,2017,33(6):105-108.

[3] SOO S, CHANG C, SRIRAMA S N. Proactive service discovery in fog computing using mobile ad hoc social network in proximity [C]//Proceedings of IEEE International Conference on Internet of Things. Washington D. C., USA: IEEE Press, 2017:561-566.

[4] JIA Bing, YANG Yongjian, SUN Yongxiong. Resource selection method for active service based on Internet of things [J]. Journal of Huazhong University of Science and Technology (Nature Science Edition), 2012, 40(S):47-50. (in Chinese)

贾冰,杨永健,孙永雄.基于物联网的主动服务资源选择方法[J].华中科技大学学报(自然科学版),2012,40(增刊):47-50.

[5] LIANG Jun, ZHAO Zhenchao, CHEN Long, et al. The driving active service system model based on MAS for Internet of vehicle environment [J]. Scientia Sinica (Technologica), 2016, 46(12):1303-1314. (in Chinese)

梁军,赵振超,陈龙,等.车联网环境下基于MAS的行车主动服务模型[J].中国科学(技术科学),2016,46(12):1303-1314.

[6] TAI Chunliang, XIE Yi, SUN Zhixin. A recommendation algorithm of location service based on user interest [J]. Computer Technology and Development, 2017, 27(9):64-69. (in Chinese)

邵淳亮,谢怡,孙知信.一种基于用户兴趣的位置服务推荐算法[J].计算机技术与发展,2017,27(9):64-69.

[7] WU Xiaokun, CHENG Bo, CHEN Junliang. Collaborative filtering service recommendation based on a novel similarity computation method [J]. IEEE Transactions on Services Computing, 2017, 10(3):352-365.

[8] SUBERCAZE J, GRAVIER C, LAFOREST F. Real-time, scalable, content-based Twitter users recommendation [C]//Companion of the Web Conference 2018. Lyon, France: [s.n.], 2018:1367-1367.

[9] YU Chunhua, LIU Xuejun, LI Bin. Mobile service recommendation based on context similarity and social network [J]. Acta Electronica Sinica, 2017, 45(6):1530-1536. (in Chinese)

俞春花,刘学军,李斌,等.基于上下文相似度和社交网络的移动服务推荐方法[J].电子学报,2017,45(6):1530-1536.

[10] LU Feng, LI Hairong, HAN Yan. QoS collaborative filtering recommendation of Web service based on spatial and temporal similarity sensing [J]. Computer Engineering, 2017, 43(4):28-33. (in Chinese)

- 卢凤,李海荣,韩艳. 基于时空相似度感知的 Web 服务 QoS 协同过滤推荐[J]. 计算机工程,2017,43(4):28-33.
- [11] VIJAYAKUMAR V, CHALLA A, CHANDRASHEKHAR M, et al. A personalized movie recommendation system based on collaborative filtering [J]. International Journal of High Performance Computing and Networking, 2017, 10(1/2):54.
- [12] SHU Jiangbo, SHEN Xiaoxuan, LIU Hai, et al. A content-based recommendation algorithm for learning resources [J]. Multimedia Systems, 2018, 24(2):163-173.
- [13] ZHANG Yongjie, YANG Pengfei, DUAN Qun. An algorithm of intelligent recommendation based on association rules[J]. Modern Computer, 2016(10):25-27. (in Chinese)
张勇杰,杨鹏飞,段群. 基于关联规则的商品智能推荐算法[J]. 现代计算机, 2016(10):25-27.
- [14] SHEN Jiajie, CHENG Xianyi. Research on intelligent recommendation algorithm of e-commerce based on association rules [J]. IOP Conference Series: Materials Science and Engineering, 2017, 242(1):1-5.
- [15] REN Di, WAN Jian, YIN Yuyu. Web services QoS prediction method based on Bayes classification [J]. Journal of Zhejiang University (Engineering Science), 2017, 51(6):1242-1251. (in Chinese)
任迪,万健,殷昱煜,等. 基于贝叶斯分类的 Web 服务质量预测方法研究[J]. 浙江大学学报(工学版), 2017, 51(6):1242-1251.
- [16] GUO Kehua, HUANG Yujian, KUANG Li, et al. CASP: a context-aware transparent active service provision architecture in a mobile internet environment [J]. Computing in Science and Engineering, 2016, 19(1):38-45.
- [17] HU Jingjing, CHEN Xiaolei, ZHANG Changyou. Proactive service selection based on acquaintance model and LS-SVM [J]. Neurocomputing, 2016, 211:60-65.
- [18] LEE W P, LEE K H. Making smartphone service recommendations by predicting users' intentions: a context-aware approach [J]. Information Sciences, 2014, 277:21-35.
- [19] LI S, KAWALE J, FU Y. Deep collaborative filtering via marginalized denoising auto-encoder [C]//Proceedings of ACM International Conference on Information and Knowledge Management. New York, USA: ACM Press, 2015:811-820.
- [20] WANG Xinxi, WANG Ye. Improving content-based and hybrid music recommendation using deep learning [C]//Proceedings of ACM International Conference on Multimedia. New York, USA: ACM Press, 2014:627-636.
- [21] BANSAL T, BELANGER D, MCCALLUM A. Ask the GRU: multi-task learning for deep text recommendations [C]//Proceedings of ACM Conference on Recommender Systems. New York, USA: ACM Press, 2016:107-114.
- [22] ZHANG Min, DING Biyuan, MA Weizhi, et al. Hybrid recommendation approach enhanced by deep learning [J]. Journal of Tsinghua University (Science and Technology), 2017, 57(10):1014-1021. (in Chinese)
张敏,丁弼原,马为之,等. 基于深度学习加强的混合推荐方法[J]. 清华大学学报(自然科学版), 2017, 57(10):1014-1021.
- [23] WANG Juan, XIONG Wei. Collaborative filtering recommender system based on matrix factorization and its application [J]. Statistics and Decision, 2019, 35(6):17-20. (in Chinese)
王娟,熊巍. 基于矩阵分解的最近邻推荐系统及其应用[J]. 统计与决策, 2019, 35(6):17-20.
- [24] WANG Y, KOSINSKI M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images [J]. Journal of Personality and Social Psychology, 2018, 114(2):246-257.
- [25] WU Kedi, WEI Guowei. Quantitative toxicity prediction using topology based multi-task deep neural networks [J]. Journal of Chemical Information and Modeling, 2018, 58(2):520-561.

编辑 金胡考