

基于深度森林的用户购买行为预测模型

葛绍林^{1,3} 叶 剑^{2,3} 何明祥¹

(山东科技大学计算机科学与工程学院 山东 青岛 266590)¹

(中国科学院计算技术研究所泛在计算系统研究中心 北京 100190)²

(移动计算与新型终端北京市重点实验室 北京 100190)³

摘 要 近年来,网络零售保持高速增长,网站中富含大量的用户行为数据。电商平台中的用户对商品的操作行为可以体现用户偏好,如何利用用户行为挖掘用户偏好已经成为学术界和工业界的关注焦点,并已经取得了众多研究成果。然而,目前用户操作行为预测方法研究通常只针对用户某一类操作行为进行分析,无法完备反映用户行为的整体特征。因此,提出一种基于深度森林的用户购买行为预测模型,通过构建用户行为特征工程建立整体用户行为特征模型;基于此,提出基于深度森林的用户购买行为预测方法,实现高效的行为预测训练效果。该方法的训练时间为 43s, F1 值为 9.73%, 相对其他模型取得了更好的效果。实验结果表明,该模型在降低时间开销的同时,提高了预测准确率。

关键词 用户行为特征,深度森林,特征工程,购买行为预测

中图分类号 TP181 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.09.027

Prediction Model of User Purchase Behavior Based on Deep Forest

GE Shao-lin^{1,3} YE Jian^{2,3} HE Ming-xiang¹

(College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong 266590, China)¹

(Research Center for Ubiquitous Computing Systems, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)²

(The Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100190, China)³

Abstract In recent years, online retail kept growing at a high speed. There exist redundant user behavior data in website. User's behavior can embody user's preference in the e-commerce platform. How to utilize user behavior to mine user preferences has become the focus of attention in academia and industry, and has formed a number of research results. The prediction methods of user behavior only aims at a certain type of user behavior, which is not able to reflect the overall characteristics of user behavior. Therefore, this paper proposed deep forest based prediction model of purchase behavior. By constructing feature engineering of user behavior, a whole user behavior feature model was built. In order to achieve efficient training, a deep forest based prediction method of purchase behavior was put forward to implement the behavior recognition training effect. The training time of this method is 43s, and the F1 value is 9.73%. Compared with other models, this method has achieved good results in both indexes. Finally, the experiments show that the model has an ability to reduce the time overhead and improve the prediction accuracy.

Keywords Characteristics of user behavior, Deep forest, Feature engineering, Prediction of purchase behavior

1 引言

近年来,随着互联网技术的飞速发展,电商消费者规模不断扩大。第 41 次中国互联网络发展状况统计报告的数据显示,截至 2017 年 12 月,我国网络购物用户规模达到 5.33 亿,较 2016 年增长 14.3%,占网民总体的 69.1%。电子商务在给人们衣食住行带来便捷的同时,也因其所涵盖商品的类别丰富且营销方式多样化,给用户挑选商品带来了诸多不便。因此,通过技术手段深入挖掘并析用户的购买行为,为用户

抉择和商家运营提供可靠的技术支撑,已成为工业界和学术界的重要研究课题,对于提高用户体验、实现商家与用户的互利共赢具有重要意义。

在此背景下,协同过滤、内容推荐等定向技术应运而生。协同过滤的大致思想就是基于评分相似的最近邻居的评分数据向用户推荐其可能感兴趣的项目^[1],然而某些场景无法获取到用户评分数据,因此无法完成推荐。基于内容推荐的大致思想是挖掘用户的特征来与商品的特征进行匹配,把两者相似度最高的项目推荐给用户。这些推荐技术主要依据用户

到稿日期:2018-07-27 返修日期:2018-10-24 本文受国家重点研发计划项目(2016YFB1001100),国家自然科学基金青年基金课题(61401040)资助。

葛绍林(1993-),男,硕士生,主要研究方向为数据挖掘;叶 剑(1974-),男,高级工程师,主要研究方向为普适计算、数据挖掘;何明祥(1969-),男,博士,副教授,硕士生导师,主要研究方向为信息处理、人工智能, E-mail: hmx0708@163.com(通信作者)。

发生购买行为的结果完成推荐,而忽略了用户在购买过程中产生的操作行为。因此,定向推荐技术能够得出用户可能会购买哪一种类型的商品,却不能精确预测用户在将来某一天内最终购买该商品。近年来,海量的用户操作行为数据在各行各业的服务平台中产生,深入挖掘用户的操作行为数据,可以发现用户的购物习惯与偏好。利用用户购买过程中产生的操作行为数据进行推荐成为了一种可行的方法,并已经得到了一些学者的研究。

鉴于此,文章提出一种基于深度森林的用户购买行为预测模型。第2节介绍传统推荐算法存在的问题及其相关解决方案;第3节介绍了问题场景以及模型;第4节为实验部分,并对结果进行了分析和讨论;最后总结全文并展望未来。

2 相关工作

传统的推荐算法有:协同过滤推荐算法、基于内容的推荐算法以及混合推荐算法。协同过滤算法存在评分稀疏性、对新用户和新商品无法进行推荐,以及算法扩展性差等问题^[2-3],并且研究人员已经发现仅仅依靠用户对项目的评分来产生推荐结果并不能准确体现用户偏好^[4]。基于内容的推荐存在新用户冷启动、两个不同的商品特征词相同时无法区分、仅限于推荐与用户购买过的物品相似的商品,以及推荐多样性差等问题^[5]。混合推荐算法难以定义各个推荐算法的权重和推荐结果,且推荐框架复杂^[6-7]。

随着排序技术的发展,科研人员将排序学习技术应用到推荐算法中,认为商品的排序远比传统推荐算法依靠商品评分大小进行排序推荐更为重要。其主要思路是对用户的历史行为数据提取特征进行学习,得到商品的排序函数,最终对用户生成商品推荐列表^[8]。基于排序学习的推荐算法中,为了提高排序结果的准确性,需要集中优化排序函数和损失函数模型,然而这将导致模型复杂度呈指数级增长,同时加大了整个模型的时间开销与空间开销,进而导致推荐算法的时效性差。由于排序学习的推荐算法依赖于训练样本学习获得排序模型,因此无法用统一模型来表示不同类型的训练样本^[9]。

随着深度学习的发展,研究人员尝试用深度学习的方法来解决传统推荐算法的问题。其中,Strub等^[10]采用两个栈式降噪自编码器将商品评分和用户特征分别作为输入,用来学习用户和项目的隐表示,并使用隐表示对缺失评分进行预测。该方法有效解决了协同过滤的数据稀疏性问题,但是该方法在训练过程中由于将评分矩阵中的缺失值直接归零以减少网络的连接数量,导致了未评分信息被忽略。Xu等^[11]通过DSSM模型研究标签感知的推荐问题,利用用户和商品标签数据定义用户的输入特征,然后学习用户和商品的隐表示,通过计算用户和商品隐表示的相似度进行推荐,以解决传统基于内容的推荐系统中用户特征提取困难的问题。Wei等^[12]通过栈式降噪自编码器SDAE与组合协同过滤方法TimeSVD++,提出一种混合推荐算法,利用SDAE从商品辅助信息中学习商品隐表示,用TimeSVD++拟合用户和商品之间的评分矩阵,从而解决混合推荐中辅助数据表示困难的问题。以上3种深度学习方法很好地解决了传统推荐系统存在的问题,然而由于深度学习模型复杂,往往需要很长的训练时间,且深度学习模型往往类似于一个黑盒,容易带来推荐

系统可解释性不足的问题^[13]。

近年来,研究者考虑将用户在一些购物网站上的操作行为(点击、浏览、加入购物车、收藏、购买)用到推荐上,依据用户的历史操作行为来挖掘用户的购物习惯、行为偏好等。其中,曾宪宇等^[14]提出一种基于用户操作序列和选择模型的方法,用行为序列效用函数推断用户在购买周期中的最佳替代品,对购买商品和最佳替代品建立基于潜在因子的选择模型,从而获得用户购买偏好。丁哲等^[15]提出基于用户浏览行为的推荐模型,该模型可以对移动用户浏览互联网的行为进行预测,并在预测基础上对移动用户进行内容推荐。以上两种模型虽然很好地解决了数据稀疏性、项目冷启动等问题,但是考虑用户的行为数据都只是单一的点击(浏览)操作,在真实业务场景下,用户行为有点击、浏览、收藏、加入购物车、购买等操作,挖掘用户在购买过程中产生的整体行为,将会提高推荐的准确率。

综上所述,这些技术多基于用户与商品关系完成推荐,即便考虑了用户行为,也只是用户操作行为的一种,没有将用户的整体操作行为考虑在内。本文针对传统推荐的数据稀疏性问题,深度学习解决传统推荐问题时出现的训练时间长、模型复杂、需要大规模训练数据,以及不能全面考虑用户的整体操作行为完成推荐等问题,基于集成学习的方法,提出基于深度森林^[16]的用户购买行为预测模型,其能够很好地提高预测性能^[17]。深度森林是基于树的学习模型,相对于深度学习的推荐模型能更容易地进行理论分析,并且模型的参数相对于深度学习要少得多,这使得参数调节变得简便。该模型的参数设定具有很强的鲁棒性,多数情况下,针对不同领域数据,使用默认值也可以取得不错的效果,且模型训练速度快、准确率高。因此,本文利用深度森林方法,并结合用户操作行为数据建立模型,完成推荐。

3 推荐模型

3.1 问题场景

电商平台中,用户购物时会浏览不同种类的商品,当挑选和比较这些商品时,会产生一系列操作行为,如点击、浏览、收藏、加入购物车、购买等,这些用户操作行为数据保存在用户行为日志中。不同用户在购物时会有自己的操作习惯,例如:用户浏览次数达到一定数目后就会发生购买行为,用户喜欢从收藏的商品中购买商品,用户每次购买前都将商品加入购物车等。如表1所列,用户1一个月内只有浏览行为,无收藏与加购行为,说明该用户做事情可能很果断。用户3一个月内的浏览次数、收藏次数、加购次数、购买次数比较多,说明该用户可能是购物狂。图1是阿里平台的所有用户在2014年11月18日到12月18日期间每天购买、加购、收藏、浏览行为的数量变化图。从中看出,所有用户的日加购行为、收藏行为的数量比购买行为数量多,说明用户加入购物车或收藏不一定发生购买行为;所有用户每天加购行为比收藏行为多30%,说明用户在选择商品时,更喜欢加入购物车;浏览行为数量是其他行为的几十倍,说明用户在购买商品时存在大量的挑选和比较行为。由以上分析可以发现用户的操作行为间接体现了用户的偏好和习惯,因此可以从这些方面挖掘特征,据此对用户在未来一天是否发生购买行为进行预测。

表1 某电商平台一个月内的用户操作行为统计

Table 1 User behavior statistics of electronic business platform with in one month

用户	浏览次数	收藏次数	加购次数	购买次数
用户 1	207	0	0	4
用户 2	2035	129	185	14
用户 3	18023	438	244	24
用户 4	474	0	33	9

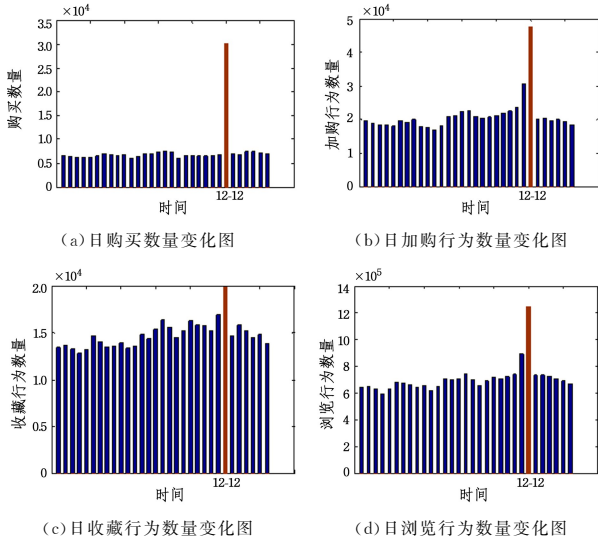


图1 用户日操作行为数量变化图

Fig. 1 Change of number of users' operation behavior per day

3.2 特征工程

对于机器学习模型,特征抽取的优劣是影响最终结果的关键。所谓特征提取,就是将原始特征转换为一组具有明显统计意义或物理意义的特征,以发现更有价值的潜在变量,帮助人们产生深刻理解。

3.1节的用户行为分析可为用户在电商网站上的操作提供一些特征构建思路。本文主要从5方面考虑特征构建,把这些特征组合在一起形成训练集、验证集及测试集,如图2所示。

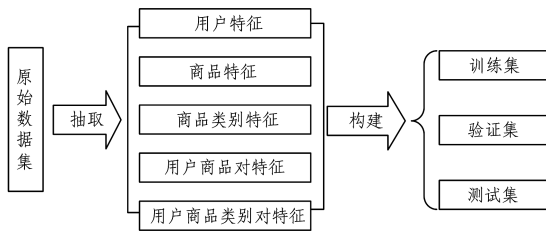


图2 构建特征的模型图

Fig. 2 Model diagrams for building features

假设用户集合为 U ,商品集合为 I ,用户对商品的交互关系 $R \subseteq U \times I \times O \times T$,其中 $O = \{1, 2, 3, 4\}$,1,2,3,4分别代表浏览、收藏、加购、购买操作, T 为操作行为发生的时间,若 $r \in R, r = \{u, i, o, t\}$ 表示用户 u 在 t 时刻对商品 i 操作 o 。 T_d 为预测日当天操作行为的时间集合, T_{d-j} 是预测日前 j 天的当天操作行为的时间集合。

(1) 用户商品对特征

关注用户对商品最近的行为,其能代表用户打算购买或者不购买商品。抽取特征如下:用户在距离预测日前 n 天的各种行为计数,记为 $|M_n^o|, M_n^o = \{y | y \in \bigcup_{j=1}^n T_{d-j}, (u, i, o, y) \in R\}$,其中 $o \in O, u \in U, i \in I$;用户在距离预测日前 n 天的所有行为总数,记为 $|M_n|, M_n = \bigcup_{o=1}^4 M_n^o$;由于用户的购物行为具有周期性,因此取一周为一个考察时间段, n 可以取 $\{1, 4, 7\}$;用户上次对商品发生的所有行为的时间距预测日的时间差,记为 $T_d - T_o$,其中 T_o 为用户上次对商品发生的所有行为的时间。

(2) 用户商品类别对特征

用户将更频繁与某一商品类别交互(商品类别指商品属于的类别,如冰箱属于家电类),因此可以挖掘用户对某一商品类别的偏好,抽取的特征与用户商品对抽取特征一致,即用户与商品类别构成一组特征,也可用与用户商品对相同的公式来表示用户商品种类对特征。

(3) 用户特征

用户特征是描述用户的属性。抽取特征如下:用户 u 距预测日前 n 天的所有行为的计数,记为 $|C_n^o|, C_n^o = \{(x, y) | y \in \bigcup_{j=1}^n T_{d-j}, (u, i, x, o, y) \in R\}$,其中 $o \in O, u \in U$;距预测日前 n 天用户的所有行为总数,记为 $|C_n|, C_n = \bigcup_{o=1}^4 C_n^o$;用户购买行为的平均时间间隔;用户点击购买转化率 $|R_1|/|R|$ (用户购买次数占点击商品次数,其中点击为用户与商品交互行为的统称)。

用户 u 购买行为的总数为 $|R_1|, R_1 \subseteq \{(u, i, 4, t) | (u, i, 4, t) \in R\}$ 。

(4) 商品特征

为了描述用户对商品的关注度,抽取了商品在距预测日前 n 天被访问的用户总数、距预测日前 n 天商品被访问的所有行为总数、距预测日前 n 天商品被访问的所有行为计数、商品被购买的平均时间间隔、商品点击转化率,也可用与用户特征相同的公式表示商品特征。

(5) 商品类别特征

商品类别特征与商品特征的抽取方法一致,也是从行为次数、时间粒度及转化率方面考虑特征的构建,也可用与用户特征相同的公式表示商品的类别特征。

3.3 深度森林算法

图3所示是深度森林的整体结构图。其输入是3.2节特征构成的原始样本集(其中一行样本数据由用户id、商品id、5类特征构成的数据、标签构成),模型每层由2个完全随机森林(椭圆框)和2个随机森林(矩形框)组成^[18],其中每个完全随机森林包含 n 棵完全随机树^[19],每个随机森林包含 n 棵决策树。每棵完全随机树通过随机选择一个特征在树结点进行分割生长,直至每个叶结点只包含同类实例或不大于10个实例。不同于完全随机树,决策树随机选取 \sqrt{d} 个特征(其中 d 是输入特征个数),然后选择其中Gini值^[20]最佳的特征进行分割生长,直至每个叶结点只包含相同的实例。由于是对用户将来某一天是否发生购买行为进行预测,而用户可能出现购买和未购买(购买标记为1,未购买标记为-1)两种情况,因此可以将此问题归结为一个二分类问题,每个森林输出一个二维类向量。为了减少过拟合风险,每个森林输出的类向量经 K 折交叉验证产生,即每个实例会被当作训练数据 $K-1$ 次,生成 $K-1$ 个类向量,然后对其求均值,将其与原始样本集中的特征进行拼接后重新作为下一层的原始输入。

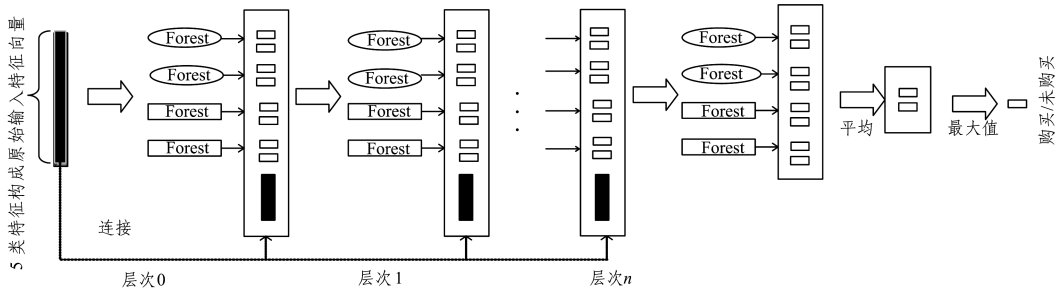


图 3 深度森林的整体结构

Fig. 3 Overall structure of deep forest

深度森林算法具体如算法 1 所示。

算法 1 深度森林算法

输入:数据集 D

输出:深度森林模型

1. 初始化 $D_0 = D$;
2. 用 $D_i (i=0, 1, 2, 3, \dots)$ 训练得到 2 个随机森林与 2 个完全随机森林, 将它们组合在一起形成深度森林的第 i 层;
3. 对第 i 层的每个森林输出一个二维类向量 X , 将其叠加求均值, 记为 \bar{X} , 令第 i 层的输出类别为 $\arg\max \bar{X}$;
4. 在当前模型上用验证集进行预测, 得到准确率 θ_i , 当 $\theta_i - \theta_{i-1} < 0$, ($i > 1$) 时, 输出深度森林模型并终止;
5. 对第 i 层的每个森林输出二维类向量 X , 并与 D_0 中的特征进行拼接, 得到新的 D_{i+1} ; 跳转至步骤 2。

由深度森林结构图和算法可知, 深度森林每层使用了不同种类的森林, 提高了模型的容错率。上节提取的特征可能并不充分, 而深度森林下一层输入是上一层特征的强化, 并且每层都与原始特征进行拼接, 因此避免了信息丢失。

4 实验

本文采用 Anaconda4.4 + PyCharm2017 作为实验平台, Anaconda 是一个用于科学计算的 Python 发行版。实验使用了 scikit-learn 包中的机器学习工具函数, 降低了实验实现的难度。实验环境为 Ubuntu 系统, Corei7 处理器 3.4 GB 主频, 16 GB 内存。

4.1 实验数据与数据分析

本文实验基于阿里平台提供的移动推荐大赛数据集¹⁾。该数据集由两部分组成, 第一部分数据集包含约 2 万用户 2014 年 11 月 18 日到 2014 年 12 月 18 日期间在商品全集上的所有移动行为数据 D , 共 2000 万条行为记录, 每个行为记录由用户 id、商品 id、行为类型、用户位置信息、商品类别和时间组成; 第二部分商品子集 P 包含约 60 多万条商品信息, 每条记录由商品 id、商品类别及位置信息组成。

由于原始数据中约 31% 的地理位置信息是缺失的, 因此抽取特征时, 将不考虑用户地理位置信息特征, 需要将这些用户地理信息从原始数据集中抽离出去。由于“双 12”当天的数据无法代表用户的真实购物习惯, 又由于用户购物具有周期性, 因此选取一周为一个时间段。通过选取 11-18 至 11-25 与预测日 11-26 构成训练集, 11-27 至 12-4 与预测日 12-5 构成验证集, 去除“双 12”前 7 天的用户行为数据, 取 12-13, 12-

18 与预测日 12-19 为测试集进行实验验证, 最后的评分数据是这些用户在预测日当天对商品子集 (P) 的购买数据。

4.2 评价指标

采用经典的精确度 (Precision)、召回率 (Recall) 和 F1 值作为评估指标。精确度是指预测正确的购买数据集占有所有预测的购买数据集的比例; 召回率是指预测正确的购买数据集占有所有实际购买数据集的比例, 召回率越高, 说明对购买数据集的预测越准确。但精确度和召回率是相互作用的一对指标, 一个指标的增加将导致另一个指标的减小, 为此选用 F1 值作为最终的唯一评测标准来评价模型的整体性能。具体评判公式如下:

$$Precision = \frac{|\cap (PredictionSet, ReferenceSet)|}{|PredictionSet|}$$

$$Recall = \frac{|\cap (PredictionSet, ReferenceSet)|}{|ReferenceSet|}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

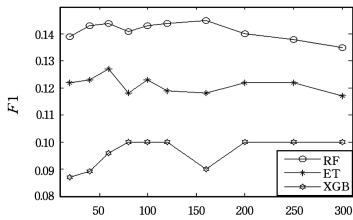
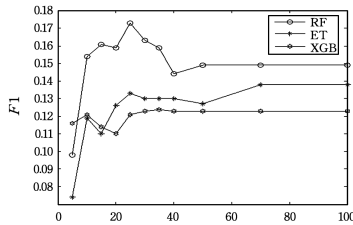
其中, $PredictionSet$ 为算法预测的购买数据集, $ReferenceSet$ 为真实的购买数据集。

4.3 实验结果

深度森林模型中森林的构建是模型建立的核心, 而决策树的构建是森林的核心, 因此决策树的数量及深度对深度森林算法的运算效率和分类效果至关重要。由于可以在深度森林中级联不同的模型, 因此尝试级联极端随机树 (ET)、随机森林 (RF)、完全提升树 (XGB)。如图 4(a) 所示, 可以看出当级联层是随机森林时, 模型的 F1 值比其他级联模型要高, 且随着 $n_estimate$ 参数的逐渐增加, 各级联模型在验证集上的预测结果并未持续改善, 反而呈下降趋势; 又由于 $n_estimate$ 参数的增大会带来时间上的开销, 通过综合分析: 对于级联 RF 模型, 选取 $n_estimate = 160$; 对于级联 ET, 选取 $n_estimate = 60$; 对于级联 XGB 模型, 选取 $n_estimate = 60$ 。

如图 4(b) 所示, 级联各模型随参数 $maxdepth$ 的增加呈现出先增加后保持稳定状态的趋势。由于参数 $maxdepth$ 的增加也会导致模型时间开销的增加, 因此对于级联 RF 模型, 选取 $maxdepth = 25$; 对于级联 ET, 选取 $n_estimate = 70$; 对于级联 XGB 模型, 选取 $n_estimate = 30$ 。实验验证, 当级联 RF 模型时, 选取 $n_estimate = 160$, $maxdepth = 25$ 能获取最高的 F1 值, 为 9.73%。

¹⁾ <https://tianchi.aliyun.com/competition/information.html>.

(a)各模型中 $n_estimate$ 参数对 F1 值的影响(b)各模型中 $maxdepth$ 参数对 F1 值的影响图4 各模型中 $n_estimate$ 和 $maxdepth$ 参数对 F1 值的影响Fig. 4 Influence of $n_estimate$ and $maxdepth$ on F1 in each model

为了突出深度森林模型的优点,对于上述数据集,用传统的机器学习方法及深度神经网络进行预测。虽然传统的机器学习方法的训练时间很短,但 F1 的效果不理想;深度神经网络的训练时间是传统机器学习方法的近 100 倍。可能由于数据集或者特征构建的原因,深度神经网络在 F1 值上并没有取得较好的结果。深度森林在训练时间上比传统机器学习模型要长,但是在 F1 值上相比其他模型有提升。表 2 中的实验数据表明,使用深度森林预测模型取得了更好的效果。

表2 深度森林与其他模型比较

Table 2 Comparison of deep forest and other models

模型	训练时间/s	F1 值/%
LR	7	5.6
SVM	35	7.2
Xgboost	28	6.8
随机森林	17.45	9.08
深度神经网络	1392	7.86
深度森林	43	9.73

结束语 本文提出用深度森林模型对用户行为进行预测,从数据处理、特征工程、模型建立、模型调优、结果预测等常用机器学习流程角度进行建模。每一个流程对最终结果的预测都是至关重要的,如设计有效的特征,对缺失值或异常值进行处理等等。由于深度森林可以级联不同的模型,如把级联的森林模型换成线性回归模型,因此本文模型在分类预测效果上还有进一步提升的可能。

参考文献

[1] LI G, CHEN Q, LI L. Collaborative Filtering Recommendation Algorithm Based on Rating Prediction and Ranking Prediction [J]. Chinese Journal of Electronics, 2017, 45(12): 3070-3075.

[2] LENG Y J, LU Q, LIANG C Y. Survey of Recommendation Based on Collaborative Filtering [J]. Pattern Recognition and Artificial Intelligence, 2014, 27(8): 720-734.

[3] WANG H, WANG N, YEUNG D Y. Collaborative deep learning for recommender systems [C] // Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015: 1235-1244.

[4] KARATZOGLOU A, BALTRUNAS L, SHI Y. Learning to

rank for recommender systems [C] // Proceedings of the 7th ACM Conference on Recommender Systems. ACM, 2013: 493-494.

[5] VESIN B, BUDIMAC Z. E-Learning personalization based on hybrid recommendation strategy and learning style identification [J]. Computers & Education, 2011, 56(3): 885-899.

[6] WEN H, FANG L, GUAN L. A hybrid approach for personalized recommendation of news on the Web [J]. Expert Systems with Applications, 2012, 39(5): 5806-5814.

[7] HUANG Z H, ZHANG J W, ZHANG B, et al. Survey of Semantics-Based Recommendation Algorithms [J]. Chinese Journal of Electronics, 2016, 44(9): 2262-2275.

[8] LI H. Learning to rank for information retrieval and natural language processing [J]. Synthesis Lectures on Human Language Technologies, 2014, 7(3): 1-121.

[9] HUANG Z H, ZHANG J W, TIAN C Q, et al. Survey on Learning-to-Rank Based Recommendation Algorithms [J]. Journal of Software, 2016, 27(3): 691-713.

[10] STRUB F, MARY J. Collaborative filtering with stacked denoising autoencoders and sparse inputs [C] // NIPS workshop on machine learning for ecommerce. 2015.

[11] XU Z, CHEN C, LUKASIEWICZ T, et al. Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling [C] // Proceedings of the 25th ACM International Conference on Information and Knowledge Management. ACM, 2016: 1921-1924.

[12] WEI J, HE J, CHEN K, et al. Collaborative filtering and deep learning based recommendation system for cold start items [J]. Expert Systems with Applications, 2017, 69: 29-39.

[13] HUANG L W, JIANG B T, LV S Y, et al. Survey on Deep Learning Based Recommender Systems [J]. Chinese Journal of Computers, 2018, 42(7): 191-219.

[14] ZENG X Y, LIU Q, ZHAO H K, et al. Online Consumption Prediction via Modeling User Behaviors and Choices [J]. Journal of Computer Research and Development, 2016, 53(8): 1673-1683.

[15] DING Z, QIN Z, ZHENG W T, et al. A Recommendation Model Based on Browsing Behaviors of Mobile Users [J]. Journal of University of Electronic Science and Technology of China, 2017, 46(6): 907-912.

[16] ZHOU Z H, FENG J. Deep Forest: Towards An Alternative to Deep Neural Networks [C] // Twenty-Sixth International Joint Conference on Artificial Intelligence. 2017: 3553-3559.

[17] YU Y, ZHOU Z H. Research on Complete Random Learning Scheme in Ensemble Learning [J]. Computer Engineering, 2006, 32(17): 100-102.

[18] GRAY K R, ALJABAR P, HECKEMANN R A, et al. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease [J]. NeuroImage, 2013, 65: 167-175.

[19] LIU F T, KAI M T, YU Y, et al. Spectrum of variable-random trees [J]. Journal of Artificial Intelligence Research, 2008, 32(1): 355-384.

[20] DING J M, LIU G Q, LI H. The Application of Improved Random Forest in the Telecom Customer Churn Prediction [J]. Pattern Recognition and Artificial Intelligence, 2015, 28(11): 1041-1049.