

doi:10.16104/j.issn.1673-1891.2021.03.016

# 粒子群优化深度交叉神经网络推荐算法设计研究

高心乐, 邱煜炎

(蚌埠医学院卫生管理学院, 安徽 蚌埠 233000)

**摘要:**深度交叉网络(DCN)因无法针对不同的数据集准确设置超参数,导致点击率预估模型无法达到最优解。基于此,提出粒子群优化深度交叉神经网络推荐算法(PSO-DCN),该算法通过使用粒子群优化算法的全局搜索特性,对深度交叉神经网络层数及每层神经元个数进行全局搜索,并将其最优值放入 DCN 模型中运行进行参数优化。在不同的数据集上,进行数次实验对比,实验结果表明 PSO-DCN 在模型训练速度以及测试集上的表现都较优于当前流行的基于深度学习框架的推荐算法。

**关键词:**深度学习;神经网络;粒子群优化;推荐系统;人工智能

**中图分类号:**TP183 **文献标志码:**A **文章编号:**1673-1891(2021)03-0075-08

## Study on Recommendation Algorithm Design of Deep Crossover Neural Network Based on Particle Swarm Optimization

GAO Xinle, QIU Yuyan

(School of Health Management, Bengbu Medical College, Bengbu, Anhui 233000, China)

**Abstract:** Deep crossover network (DCN) cannot accurately set super parameters for different data sets, which leads to the click rate prediction model being unable to achieve the optimal solution. Based on this, this paper proposes the deep crossover neural network recommendation algorithm based on particle swarm optimization (PSO-DCN), which uses its global search characteristics to globally search the number of layers and neurons in each layer of the deep crossover neural network, and puts its optimal value into the DCN model for parameter optimization. The experimental results show that the pso-dcn is better than the current popular recommendation algorithms based on deep learning framework in the model training speed and in the test set.

**Keywords:** deep learning; neural network; particle swarm optimization; recommendation system; artificial intelligence

## 0 引言

近年来,5G 技术、大数据、物联网和云计算的发展迅速,导致互联网网络用户数量大幅度增加,网络资源的数字化和信息容量也呈现出爆炸式的增长。推荐系统具备了解决信息的过载和挖掘客户数据的能力,并且已经得到了普遍的应用。推荐系统是通过用户对物品之间的关系进行分析和评价,根据用户的喜好和产品相关特征为其提供所需要的推荐信息,有效地解决了用户信息过载的问题<sup>[1]</sup>。深度学习算法,作为推荐系统的核心,其性能直接决定推荐的质量。

## 1 文献综述

推荐算法的研究起源于 20 世纪 90 年代,由美国明尼苏达大学 GroupLens 研究小组最先开始研究,他们想要制作一个名为 MovieLens 的电影推荐系统,从而实现对用户进行电影的个性化推荐。推荐系统中的排序算法从 20 世纪 90 年代开始,至今已经历了以下 3 个阶段:第 1 阶段是人工特征+线性模型(专家系统)阶段,具有模型简单、参数空间小、性能高、效果好的特点。第 2 阶段是自动特征交叉+线性模型(增强记忆)阶段,具有监督自动进行二阶、高阶特征交叉的特点,相对于第 1 阶段效果提升显著,而且只需对个别超参数进行人工调整,简单

收稿日期:2021-03-25

基金项目:安徽省教育厅高等学校自然科学研究重点项目(2019A0393);安徽省 2019 大学生创新训练项目(201910367100)。

作者简介:高心乐(2000—),男,安徽亳州人,本科生,研究方向:深度学习、人工智能。

高效。第 3 阶段是深度学习(深度发展)阶段,此阶段离散特征处理向 Embedding 空间发展,采用低维空间更加合理的表达特征;采用了 DNN 进行高阶特征挖掘,融合了低阶、二阶、高阶进行联合学习,低阶和二阶主要强化记忆能力,高阶主要强化泛化能力。第 3 阶段也是当前流行阶段,产生了许多当前流行的深度学习推荐算法。

DeepFM<sup>[2]</sup>模型是华为诺亚方舟实验室和哈工大在 2017 年合作提出的一个新型 CTR 预估模型,该模型有效结合了因子分解机与神经网络在特征学习中的优点,可以同时提取到低阶组合特征与高阶组合特征,而且除了原始特征之外无需其他特征工程。然而,其隐藏层层数和每层神经元个数超参数无法精确设置,导致了其预测准确度下降。

Feature Importance and Bilinear Feature Interaction Network (FiBiNET)<sup>[3]</sup>模型是新浪微博机器学习团队发表在 2019 推荐系统大会(RecSys19)上的一项工作,该工作中提出通过使用 Squeeze-Excitation network (SENET)<sup>[4]</sup>结构动态学习特征的重要性以及使用一个双线性函数来更好地建模交叉特征。然而,其隐藏层层数和每层神经元个数超参数无法精确设置,导致了 SENET 学习效率和 Bilinear Interaction 学习效率下降。

Operation-aware Neural Networks (ONN)是在 Operation-aware Neural Networks for User Response Prediction<sup>[5]</sup>中提出的一种新的用于点击率预测/广告推荐的深度学习网络模型,相比之前的同类模型,不同点在于丰富了 Embedding 表达的处理:对于每个不同的操作(复制或内积)所使用的 Embedding 方法不同,具体来说是每个特征都要在 Embedding 层训练足够多的系数,产生足够多的中间结果向量,用于后面一次性的内积或其他操作。

Deep & Cross Network (DCN)<sup>[6]</sup>模型是谷歌和斯坦福大学在 2017 年提出的用于 Ad Click Prediction 的模型。DCN (Deep Cross Network)在学习特定阶数组合特征的时候效率非常高,而且同样不需要特征工程,引入的额外的复杂度也是微不足道的。

尽管 DCN 模型在数据高维度、数据稀疏性的问题表现优异。但是深度交叉神经网络层数和每层神经元个数超参数无法精确设置,可能会出现过拟合或者欠拟合现象,这将导致模型无法达到全局最优解。

针对以上问题,本文提出一种粒子群优化深度交叉神经网络推荐算法(PSO-DCN)。首先在不同的数据集上运用粒子群优化(Particle Swarm Optimi-

zation, PSO)<sup>[7-8]</sup>算法快速获得全局最优深度交叉神经网络层数和每层神经元个数,然后将其放入 DCN 训练模型进行训练,最后在 DCN 测试模型上进行验证。实验结果表明,相对于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN),粒子群优化深度交叉神经网络推荐算法(PSO-DCN)在模型训练速度和测试集上的 ROC 曲线下的面积(Area Under Curve, AUC)、准确率(Accuracy, Acc)、精确率(Precision, Pre)、召回率(Recall)和 F-measure 值(F-score)上都有较好的表现。

## 2 模型算法介绍

### 2.1 深度神经网络

神经网络(Neural Networks, NN)是基于感知机的扩展,而深度神经网络(Deep Neural Networks, DNN)则可以简单地理解为一种具有许多隐含层的神经网络<sup>[9-10]</sup>。深度神经网络(DNN)别名又叫作多层感知机(Multi-Layer Perceptron, MLP)。

DNN 内部的神经网络层大致可以划分为 3 类:输入层,隐藏层和输出层,如图 1 所示。一般而言,第 1 层指的是输入层,最后 1 层指的是输出层,而中间的层数都指的是输入隐藏层。

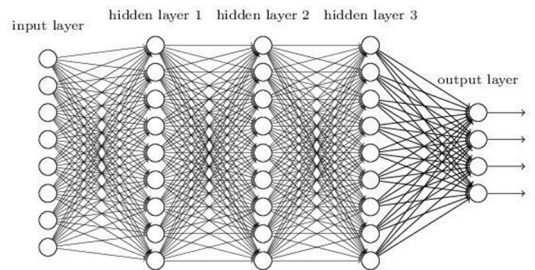


图 1 DNN 网络结构

层与层之间必须是全连接的,也就是说,第  $l$  层的任意一个神经元一定与第  $l + 1$  层的任意一个神经元相连。并且满足一个线性的关系

$$z = \sum w'_{jk} x_l + b_l + \sigma(z) \quad [11-17],$$

式中: $z$  表示线性关系; $w'_{jk}$  表示第  $(l-1)$  层的第  $k$  个神经元到第  $l$  层的第  $j$  个神经元;  $x_l$  表每层神经元对应的参数; $l$  表示隐藏层层数; $b_l$  表示每一层的偏倚; $\sigma(z)$  指代激活函数。

输入层和输出层是根据数据的特征和标签改变的,不受人工控制,而隐藏层却是受人工控制的。隐藏层的层数和每一层神经元的个数超参数无法精确设置,直接影响深度学习模型的预测准确率和精确率。

### 2.2 DCN 模型

Deep & Cross Network (DCN)模型是 Google 对

Wide & Deep 工作的一个后续研究, 将 Wide & Deep 中的 Wide 部分替换为由特殊网络结构实现的 Cross, 自动构造有限高阶的交叉特征, 并学习对应权重, 省去烦琐复杂的人工叉乘步骤, 如图 2 所示。

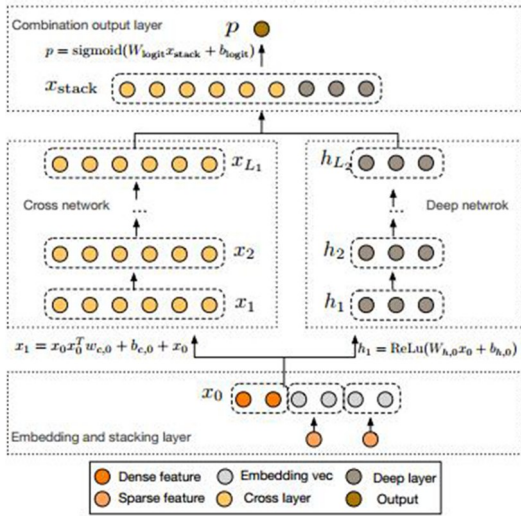


图 2 深度和交叉网络

DCN 模型底层主要分为 Embedding 和 Stacking 层, 利用 Embedding 将二值化特征转变为实值的稠密向量, Embedding 过程中用到的矩阵参数和网络中的其他参数一起训练, 然后将 Embedding 向量和经过归一化的稠密向量 Stack 到一起作为网络的输入。其数学表达式为:

$$x_0 = [x_{embed,1}^T, \dots, x_{embed,k}^T, x_{dense}^T] \quad (1)$$

式中:  $x_0$  表示网络的输入;  $x_{embed,k}^T$  表示 Embedding 向量;  $x_{dense}^T$  指代经过归一化的稠密向量。

为了高效地应用显式的交叉特征, DCN 模型引入了交叉网络 (Cross Network)。交叉网络由交叉层组成, 每个层具有以下公式:

$$x_{l+1} = x_0 x_l^T w + b_l + x_l = f(x_l, w_l, b_l) + x_l \quad (2)$$

式中:  $x_0$  是公式 (1) 的输出, 同时是公式 (2) 的输入;  $x_l, x_{l+1}$  是列向量 (column vectors), 分别表示来自第  $l$  层和第  $(l+1)$  层 cross layers 的输出;  $w_l, b_l \in R^d$  是第  $l$  层 layer 的 weight 和 bias 参数。在完成一个特征交叉  $f$  后, 每个 cross layer 会将它的输入加回去, 对应的 mapping function  $f: R^d \rightarrow R^d$ , 刚好等于残差  $x_{l+1} - x_l$ , 这里借鉴了残差网络的思想。

交叉网络 (Cross Network) 的独特结构使得交叉特征的阶 (the degree of cross features) 随着 layer 的深度而增长。对于第  $l$  层 layer, 它的最高多项式阶 (在输入  $x_0$  上) 是  $l+1$ 。实际上, 交叉网络 (Cross Network) 由这些交叉项  $x_1^{\alpha_1} x_2^{\alpha_2} \dots x_d^{\alpha_d}$  组成, 对应的阶从 1 到  $l+1$ 。

交叉网络的参数数目少, 从而限制了模型的能力 (capacity)。为了捕获高阶非线性交叉, 本文引入了一个深度网络 (Deep Network)。每个深度层具有如下公式:

$$h_{l+1} = f(w_l h_l + b_l) \quad (3)$$

式中:  $h_l \in R^{n_l}, h_{l+1} \in R^{n_{l+1}}$  分别第  $l$  层和第  $(l+1)$  层 hidden layer;  $w_l \in R^{n_{l+1} \times n_l}, b_l \in R^{n_{l+1}}$  第  $l$  个 deep layer 的参数;  $f(\cdot)$  是 ReLU function。

最后, 将交叉网络 (Cross Network) 的输出和深度网络 (Deep Network) 的输出进行拼接 (concatenate), 然后将该拼接向量 (concatenated vector) 喂给一个标准的逻辑回归模型, 并使用梯度下降算法进行优化, 最后经 sigmoid 函数输出。

$$p = \sigma([x_{L1}^T, h_{L2}^T] w_{logits}) \quad (4)$$

式中:  $p$  是最终的预测概率;  $x_{L1}^T$  表示 Cross Network 的最终输出;  $h_{L2}^T$  表示 Deep Network 的最终输出;  $w_{logits}$  是 Combination Layer 的权重。

DCN 模型使用交叉网络在每一层都运用了 Feature Crossing, 高效学习高阶特征。网络结构简单且高效。相比 DNN, DCN 的 Logloss 值更低, 而且参数的数量少了一个数量级。然而, 交叉网络和深度网络的深度交叉神经网络层数和每层神经元个数超参数无法精确设置, 在数据高维度、数据稀疏性的问题上, 会导致了 DCN 模型预测效率下降。

## 2.3 粒子群优化算法

### 2.3.1 算法描述

粒子群算法是通过模拟鸟群捕食行为设计的一种群智能算法。鸟群通过种群交流以及自身经验改变位置和飞行速度以及调整个体搜寻路径, 每只鸟所能找到的最好位置, 称之为个体最优 (pbest)。群鸟中所有个体所能找到的最好位置, 称为全局最优 (gbest), 整个鸟群的觅食中心都趋向全局最优移动<sup>[18]</sup>。相比较于其他算法, 粒子群算法的优势在于简单, 容易实现, 无须梯度信息, 参数少。

### 2.3.2 算法流程

粒子群算法主要适用于连续函数的极值问题求解, 具体流程如图 1 所示。

根据图 3 的流程, 粒子群算法中所涉及的参数有: (1) 种群数量: 初始种群取值 50~200; (2) 惯性权重: 取值范围 0.4~1, 该参数反映了个体历史成绩对现在的影响; (3) 学习因子: 取值范围 0~5, 分为个体和群体 2 种学习因子; (4) 迭代次数: 取值 50~100, 要在时间耗费和搜索广度进行权衡; (5) 空间维数: 取值 2, 是指粒子在网络中搜索的空间维数, 即代表一个自变量的个数; (6) 位置限制: 是指限制

被检测到搜索空间的范围,即代表自变量的取值范围;(7)速度约束:通过设置合理的粒子速度约束可以有效地防止由于速度过快而导致的粒子跳过最优求解位置以及由于速度过慢而导致的最优收敛速度缓慢的问题<sup>[19]</sup>。

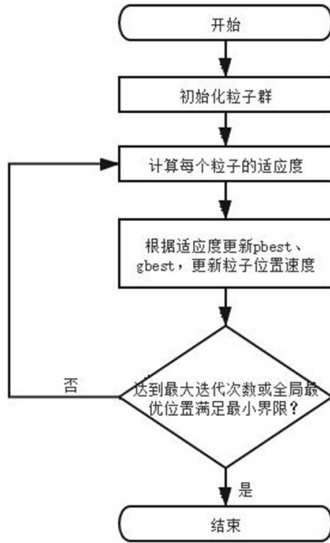


图 3 粒子群优化算法流程

### 2.3.3 算法规则与公式

粒子群算法首先初始化为一群随机粒子(随机解)。然后通过迭代找到最优解。每次迭代后,粒子通过对比 2 个极值(pbest 和 gbest)来更新自己的参数<sup>[20]</sup>。当粒子确定最优位置后,更新粒子的速度和位置,公式如下:

$$v_i = w \times v_i + c_1 \times rand() \times (pbest_i - x_i) + c_2 \times rand() \times (gbest_i - x_i) \quad (5)$$

$$x_i = x_i + v_i \quad (6)$$

式中: $i = 1, 2, \dots, N$ ,  $N$  是此群中的粒子总数; $v_i$  表示粒子的速度; $w$  表示惯性权重; $rand()$  表示介于(0, 1)之间的随机数; $x_i$  表示粒子所在当前位置; $pbest_i$  表示个体最优参数; $gbest_i$  表示全局最优参数; $c_1$  和  $c_2$  表示学习因子,通常  $c_1 = c_2 = 2$ <sup>[21]</sup>。

## 3 基于粒子群优化深度交叉神经网络算法

### 3.1 算法思路

为了解决上述深度交叉神经网络层数和每层神经元个数超参数无法精确设置,导致在数据高维度、数据稀疏性的问题上深度学习模型预测结果达不到最优解的问题,本文提出粒子群优化深度交叉神经网络推荐算法(PSO-DCN)。利用粒子群算法全局搜索能力寻找最优深度交叉神经网络层数和每层所对应的神经元个数,以达到模型最优解。

首先,通过粒子群算法初始化粒子群,并将粒

子取整作为参数放入到深度交叉神经网络中。然后,将深度交叉神经网络嵌入 DCN 训练模型中进行训练,将训练输出的验证集的对数损失函数值放入到个体最优适应度集合中,并选取其中最小的值最为全局最佳适应度。如果粒子没有达到最大迭代次数,更新粒子的位置和速度,并将新的粒子取整作为参数放入到深度交叉神经网络,并将其嵌入到 DCN 训练模型中进行训练,并将其结果放入个体新一代最优适应度集合中,选取其中最小的值与上一代全局最佳适应度进行比较,若新一代最优适应度集合中的最小值小于上一代全局最佳适应度,则更新全局最佳适应度为新一代最优适应度集合中的最小值。继续循环直达到最大迭代次数<sup>[22]</sup>。如果达到最大迭代次数,将输出全局最佳适应度对应的粒子,并将其作为参数放入 DCN 测试模型中进行测试,最终得到 DCN 最优解。

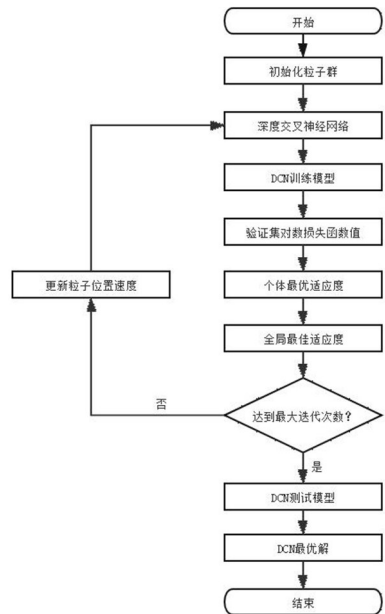


图 4 粒子群优化深度交叉神经网络推荐算法流程

### 3.2 算法模型

对于含有  $M$  条样本数据  $(X^{(i)}, y^{(i)})$  的数据集,其中  $X^{(i)} = [x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}] \in R^n$  表示特征向量,其中每个  $X^{(i)}$  含有  $n$  个特征, $y^{(i)}$  表示  $n$  个特征所对于的标签。对于二分类问题  $y^{(i)} \in \{0, 1\}$ ,对于回归问题  $y^{(i)} \in R$ 。对于样本中的  $n$  个特征,首先构建粒子群算法模型:

$$y(\hat{x}) = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n + \langle V_1, V_2 \rangle x_1 x_2 + \dots + \langle V_{n-1}, V_n \rangle x_{n-1} x_n \quad (7)$$

式中: $y(\hat{x})$  表示标签结果; $w_n$  表示权重; $x_n$  表示样本数据特征; $V_n$  表示粒子速度。

在利用粒子群算法求解时, 首先需要构建粒子群  $P = \{p_1, p_2, \dots, p_m\}$ , 即包含  $m$  个粒子群, 每个粒子  $p_i$ , 按照公式 (7) 可以表示为:

$$p_i = \{\omega_0, \omega_1, \omega_2, \dots, \omega_n, v_{1,1}, v_{1,2}, \dots, v_{1,k}, \dots, v_{n-1}, v_{n-2}, \dots, v_{n-k}\} \quad (8)$$

根据公式 (5) 利用粒子群优化算法进行求解最优深度交叉神经网络层数和每层神经元个数的过程中, 将 DCN 训练模型中输出的值作为个体的最优适应度。个体的最优适应度的最优值作为全局最佳适应度。

### 3.3 算法步骤

根据以上描述, 粒子群优化深度交叉神经网络推荐算法 (PSO-DCN) 的目的就是通过公式 (7) 快速获得全局最优区域, 并获得 DCN 训练模型的全局最优解, 将 DCN 训练模型的全局最优解所对应的深度交叉神经网络层数和每层所对应的神经元个数放入 DCN 测试模型中进行测试, 从而获得 DCN 最优解。粒子群优化深度交叉神经网络推荐算法具体步骤如下:

**Step 1:** 设置超参数, 包括粒子规模  $m$ , 惯性权重  $\omega$ , 学习因子  $c_1, c_2$ , 迭代次数  $i$ , 正则化参数  $\lambda$ , 随机种子  $n$ , 激活函数。

**Step 2:** 初始化粒子群, 即粒子的速度和位置。

**Step 3:** 将随机生成的粒子取整作为参数放入深度交叉神经网络, 并嵌入到 DCN 训练模型中进行训练。

**Step 4:** 计算个体最优适应度和全局最佳适应度。

**Step 5:** 根据公式 (5) 和公式 (6) 更新每个粒子的位置和速度。

**Step 6:** 当满足迭代次数  $i$ , 返回最优参数。

**Step 7:** 将最优参数放入到 DCN 测试模型进行测试, 得到 DCN 最优解。

## 4 实验结果与分析

### 4.1 实验准备

本实验针对二分类问题对粒子群优化深度交叉神经网络推荐算法 (PSO-DCN) 进行验证, 选取了 3 个常用的数据集, 分别是 Criteo 数据集、乳腺癌数据集和信用卡数据集, 并与当前流行的深度学习推荐算法 (DCN、FM、FiBiNET、ONN) 进行对比。其中, Criteo 数据集包含 1 000 条数据, 特征个数 39, 分类结果为是否已单击此广告 ( $\{1, 0\}$ )。乳腺癌数据集包含 569 条数据, 特征个数 30, 分类结果为良性恶性 ( $\{1, 0\}$ )。信用卡数据集包含 800 条数据, 特征

个数 20, 分类结果为信用好坏 ( $\{1, 0\}$ )。按照实验要求, 将数据集按 16:4:5 的比例分为训练集、验证集和测试集。

实验算法采用 Python3.6 编写的代码, 硬件环境采用英特尔 第二代酷睿 i7 处理器, 8GB 内存。

粒子群优化深度交叉神经网络推荐算法运行超参数设置如表 1 所示。

表 1 粒子群优化深度交叉神经网络推荐算法运行超参数

$m$	$\omega$	$c_1$	$c_2$	$i$	$\lambda$	$n$
100	0.6	2	2	100	0.000 01	1 024

当前流行的深度学习推荐算法 (DCN、FM、FiBiNET、ONN) 运行参数如表 2 所示。

表 2 当前流行的深度学习推荐算法运行超参数

$\lambda$	$n$	$k$	激活函数
0.000 01	1 024	(128, 128)	<i>relu</i>

表 1 中的超参数已在本文 3.3 中说明, 不再重复。表 2 中的参数  $\lambda$  表示正则化参数,  $n$  表示随机种子,  $k$  表示深度交叉神经网络层数和每层神经元个数。

### 4.2 结果及分析

为验证粒子群优化深度交叉神经网络推荐算法 (PSO-DCN) 比当前流行的深度学习推荐算法 (DCN、FM、FiBiNET、ONN) 训练时间更短, 以及在测试集上的对数损失函数值 (Logloss)、ROC 曲线下的面积 (Area Under Curve, AUC)、准确率 (Accuracy, Acc)、精确率 (Precision, Pre)、召回率 (Recall) 和  $F$ -measure 值 ( $F$ -score) 指标都比较优秀。因此, 在 3 个不同的数据集上进行实验, 并得到如下结果。

#### 4.2.1 Criteo 数据集

图 5 对比了在 Criteo 数据集上粒子群优化深度交叉神经网络推荐算法 (PSO-DCN) 与当前流行的深度学习推荐算法 (DCN、FM、FiBiNET、ONN) 在测试集所得损失函数值的变化情况。

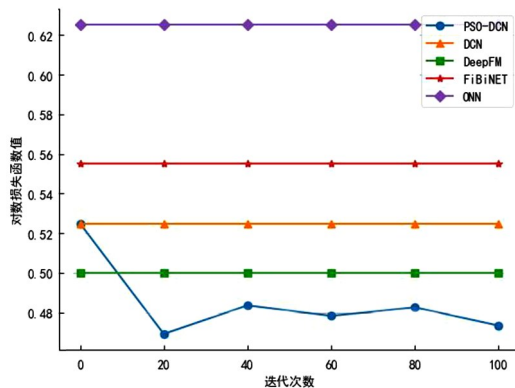


图 5 Criteo 数据集算法对数损失函数迭代对比

如图 5 所示,粒子群优化深度交叉神经网络推荐算法(PSO-DCN)的对数损失函数值随着粒子迭代次数的增加,呈现减小后增加最后稳定的趋势,并且小于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)。表 3 列出了 5 种算法训练时间(TraingTime, TT)、在测试集上的 ROC 曲线下的面积(Area Under Curve, AUC)、准确率(Accuracy, Acc)、精确率(Precision, Pre)、召回率(Recall)和  $F$ -measure 值( $F$ -score)指标。

表 3 Criteo 数据集算法运行指标

	DCN	DeepFM	FiBiNET	ONN	PSO-DCN
TT/s	3.809 6	3.578 8	13.564 6	31.757 4	3.544 8
AUC	0.699 8	0.697 9	0.675 1	0.643 4	0.704 1
Acc	0.780 0	0.805 0	0.805 0	0.800 0	0.800 0
Pre	0.272 7	0.500 0	0.500 0	0.333 3	0.363 6
Rec	0.076 9	0.076 9	0.051 3	0.025 6	0.102 6
$F$ -score	0.120 0	0.133 3	0.093 0	0.047 6	0.160 0

从表 3 可以看出,粒子群优化深度交叉神经网络推荐算法(PSO-DCN)的训练时间明显低于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN),而且在测试集上的 ROC 曲线下的面积、召回率和  $F$ -measure 值全部优于后者,但在测试集上的准确率和精确率只优于了 DCN 和 ONN 模型。可能是 DeepFM 算法和 FiBiNET 算法更容易处理稀疏稠密混合型数据。

#### 4.2.2 乳腺癌数据集

图 6 对比了在乳腺癌数据集上粒子群优化深度交叉神经网络推荐算法(PSO-DCN)与当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)在测试集所得损失函数值的变化情况。

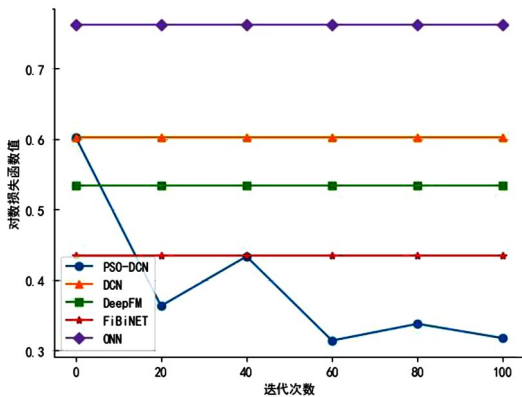


图 6 乳腺癌数据集算法对数损失函数迭代对比

如图 6 所示,粒子群优化深度交叉神经网络推荐算法(PSO-DCN)的对数损失函数值随着粒子迭

代次数的增加,整体上呈现减小的趋势,并且小于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)。表 4 列出了 5 种算法训练时间(TraingTime, TT)、在测试集上的 ROC 曲线下的面积(Area Under Curve, AUC)、准确率(Accuracy, Acc)、精确率(Precision, Pre)、召回率(Recall)和  $F$ -measure 值( $F$ -score)指标。

从表 4 可以看出,粒子群优化深度交叉神经网络推荐算法(PSO-DCN)的训练时间明显低于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN),而且在测试集上准确率、精确率和  $F$ -measure 值全部优于后者,但在测试集上的 AUC 值却只大于了 DCN 算法、DeepFM 算法和 ONN 算法,而召回率是最小的。可能是当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)更容易处理内容比较简单的数据集。

表 4 乳腺癌数据集算法运行指标

	DCN	DeepFM	FiBiNET	ONN	PSO-DCN
TT/s	3.301 0	1.981 8	3.710 7	7.562 3	1.929 8
AUC	0.954 9	0.961 2	0.973 2	0.568 5	0.968 4
Acc	0.649 1	0.657 9	0.842 1	0.578 9	0.850 9
Pre	0.622 6	0.628 6	0.785 7	0.578 9	0.802 5
Rec	1.000 0	1.000 0	1.000 0	1.000 0	0.984 8
$F$ -score	0.767 4	0.771 9	0.880 0	0.733 3	0.884 4

#### 4.2.3 信用卡数据集

图 7 对比了在信用卡数据集上粒子群优化深度交叉神经网络推荐算法(PSO-DCN)与当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)在测试集所得损失函数值的变化情况。

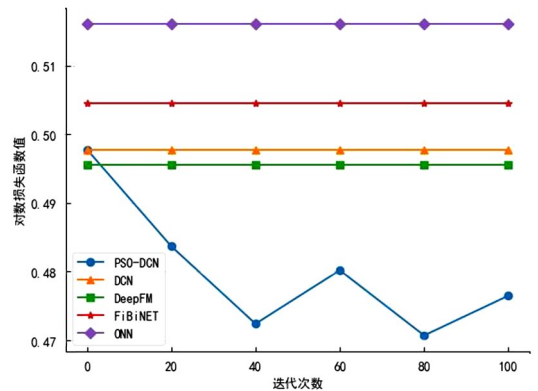


图 7 信用卡数据集算法对数损失函数迭代对比

如图 7 所示,粒子群优化深度交叉神经网络推荐算法(PSO-DCN)的对数损失函数值随着粒子迭代次数的增加,呈现先减小,后波动的趋势,并且小

于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)。表5列出了5种算法训练时间(TrainingTime, TT)、在测试集上的ROC曲线下的面积(Area Under Curve, AUC)、准确率(Accuracy, Acc)、精确率(Precision, Pre)、召回率(Recall)和F-measure值(F-score)指标。

从表5可以看出,粒子群优化深度交叉神经网络推荐算法(PSO-DCN)的训练时间略高于DCN算法和DeepFM算法,可能深度交叉神经网络层数增多导致了时间略高一点。在测试集上的AUC值、准确率和F-measure值全部优于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN),但在测试集上的召回率却低于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)。可能是当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)更容易处理内容比较简单的数据集。

表5 信用卡数据集算法运行指标

	DCN	DeepFM	FiBiNET	ONN	PSO-DCN
TT/s	1.401 1	1.365 2	2.279 6	4.702 1	1.441 1
AUC	0.808 9	0.818 3	0.819 1	0.815 9	0.820 2
Acc	0.737 5	0.743 8	0.731 2	0.750 0	0.762 5
Pre	0.744 8	0.743 2	0.724 4	0.745 0	0.769 8
Rec	0.955 8	0.973 5	1.000 0	0.982 3	0.946 9
F-score	0.837 2	0.842 9	0.840 1	0.847 3	0.849 2

#### 参考文献:

- [1] BOBADILLA J, ORTEGA F, Hernando A, et al. Recommender systems survey [J]. Knowledge - Based Systems, 2013, 46: 109-132.
- [2] GUO H, TANG R, YE Y, et al. DeepFM: a factorization-machine based neural network for CTR prediction [C]. Melbourne: Twenty-Sixth International Joint Conference on Artificial Intelligence, 2017: 1725-1731.
- [3] HUANG Tongwen, ZHANG Zhiqi, ZHANG Junlin. FiBiNET: combining feature importance and bilinear feature interaction for click-through rate prediction [C]. New York: Proceedings of the 13th ACM Conference on Recommender Systems, 2019: 169-177.
- [4] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. New Orleans: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [5] YANG Y, XU B, SHEN S, et al. Operation-aware Neural Networks for user response prediction [J]. Neural Networks, 2020, 121(1): 161-168.
- [6] WANG R, FU B, FU G, et al. Deep & cross network for ad click predictions [C]. Halifax: Proceedings of the ADKDD'17, 2017: 1-7.
- [7] KENNEDY J, EBERHART R. Particle swarm optimization [C]. Washington: Proceedings of ICNN'95-international conference on neural networks, 1995: 1942-1948.
- [8] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory [C]. Nagoya: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995: 39-43.
- [9] 管超. 基于深度学习的物流道路场景分割系统的设计与实现 [D]. 成都: 电子科技大学, 2019.
- [10] 王振业. 基于深度学习的公共数据预测研究 [D]. 西宁: 青海师范大学, 2020.
- [11] 赵振宇. 阜新地区最低气温预报订正方法对比分析 [D]. 沈阳: 沈阳农业大学, 2019.

#### 4.3 实验总结

综上实验,针对不同的数据集粒子群优化深度交叉神经网络推荐算法(PSO-DCN)比当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)表现出更优秀的运行指标。本实验数据集具有训练样本量少,特征属性多的特点,对于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN)很容易产生过拟合或者欠拟合。凭借 PSO 算法的全局搜索能力,找到最优深度交叉神经网络层数和每层神经元个数,使深度学习算法表现出更强的推荐能力。

#### 5 结语

本文引入粒子群优化算法嵌入到深度学习模型中,在全局搜索下寻找最优深度交叉神经网络层数和每层神经元个数。通过在 Criteo 数据集、乳腺癌数据集和信用卡数据集上进行实验,结果表明,相对于当前流行的深度学习推荐算法(DCN、FM、FiBiNET、ONN),粒子群优化深度交叉神经网络推荐算法(PSO-DCN)表现出更快的模型训练速度,以及更高的模型预测 AUC 值、准确率、精确率、召回率和 F-measure 值。本文实验数据比较小,初始化粒子群和粒子迭代次数也比较小,后续的研究工作中,将针对初始化粒子量或者粒子迭代次数比较大的情况,解决粒子群优化深度交叉神经网络推荐算法(PSO-DCN)迭代时间比较长的问题。

- [12] 葛胜. 腱驱动灵巧手自适应控制算法的研究[D]. 南京: 南京邮电大学, 2019.
- [13] 刘宇宁. 基于深度学习的 Web 社区推荐算法研究与应用[D]. 成都: 西南交通大学, 2018.
- [14] 卢春城. 基于深度学习的糖尿病决策算法研究[D]. 杭州: 浙江理工大学, 2019.
- [15] 李春. 基于 BP 和 LSTM 神经网络的基因表达预测研究[D]. 太原: 太原理工大学, 2019.
- [16] 靳晓岩. C-RAN 中基于深度学习的无线资源分配方法研究[D]. 南京: 南京邮电大学, 2020.
- [17] 杨越翔. 基于日志数据的 U2R 和 R2L 入侵检测研究[D]. 兰州: 西北师范大学, 2020.
- [18] 马烈. 混合优化算法及其在图像处理中的应用研究[D]. 武汉: 湖北工业大学, 2017.
- [19] 邱煜炎, 吴福生. 基于粒子群优化的因子分解机算法[J]. 山西师范大学学报(自然科学版), 2020, 34(1): 5-11.
- [20] 高海兵, 周驰, 高亮. 广义粒子群优化模型[J]. 计算机学报, 2005(12): 1980-1987.
- [21] 刁爱军. 基于粒群优化算法的云存储数据检索方法研究[J]. 激光杂志, 2016, 37(11): 98-102.
- [22] 朱瞳彤, 顾洁, 金之俭, 等. 规划与运行融合的配电网无功补偿智能协调配置[J]. 电力自动化设备, 2019, 39(2): 36-43.

(上接第 69 页)

较高。展示了本文方法在提高光栅传感网络数据分块无损压缩性能方面的优越性。

#### 参考文献:

- [1] 汪泽帆, 殷爱菡, 郭鑫, 等. 蜂窝形光纤光栅传感网络的可靠性研究[J]. 光通信技术, 2018, 42(6): 18-20.
- [2] 赵建君, 闫石, 陈红叶. 利用光栅探测激光信息的迭代算法研究[J]. 激光与红外, 2015, 45(7): 770-774.
- [3] 张明, 任建文, 陈文, 等. 光折变长周期波导光栅耦合器的设计和分析[J]. 光学学报, 2015, 35(3): 153-161.
- [4] 丁旭, 黄成, 吴晓蓓, 等. 基于压缩感知的传感器网络中概率负载均衡的数据路由协议[J]. 控制与决策, 2018, 33(6): 76-82.
- [5] 彭代锋, 张永军, 熊小东. 结合 LiDAR 点云和航空影像的建筑物三维变化检测[J]. 武汉大学学报(信息科学版), 2015, 40(4): 462-468.
- [6] 李国瑞, 孟婕, 彭三城, 等. 基于 Jacobi ADMM 的传感网分布式压缩感知数据重构算法[J]. 计算机研究与发展, 2020, 57(6): 164-171.
- [7] 周桥, 伊鹏, 门浩崧. 基于资源效用最大化的虚拟网络功能备份方法[J]. 计算机应用, 2017, 37(4): 948-953.
- [8] 森波, 付浩, 李小龙, 等. 基于局域表面等离子体共振效应的聚合物波导传感器特性研究[J]. 发光学报, 2016, 37(1): 112-116.
- [9] 杨军, 李震宇, 孙光才, 等. 一种新的大斜视 TOPSSAR 全孔径成像方法[J]. 西安电子科技大学学报, 2015, 42(1): 47-51.
- [10] 李永刚, 张治中, 李龙江. 一种 WLAN 与 eHRPD 系统间负载均衡算法[J]. 微电子学与计算机, 2017, 34(1): 44-47.
- [11] 熊风光, 李希, 韩燮. 基于整体最小二乘的椭圆拟合方法[J]. 微电子学与计算机, 2017, 34(1): 102-105.
- [12] REN N, YU Y, JIANGX, et al. Improved multi-grating filtering demodulation method based on cascading neural networks for fiber bragg grating sensor[J]. Journal of Lightwave Technology, 2019, 37(9): 2147-2154.
- [13] GERZ D, SCHWEINBERGER W, BUTLER TP, et al. Mid-infrared long-pass filter for high-power applications based on grating diffraction[J]. Optics Letters, 2019, 44(12): 3014-3017.
- [14] 孙其法, 闫秋艳, 闫欣鸣. 基于多样化 top-k shapelets 转换的时间序列分类方法[J]. 计算机应用, 2017, 37(2): 335-340.
- [15] 胡俊, 胡贤德, 程家兴. 基于 Spark 的大数据混合计算模型[J]. 计算机系统应用, 2015, 24(4): 214-218.