

基于模拟退火算法优化的BP神经网络预测模型

蒋美云

(南京工业职业技术学院计算机与软件学院, 江苏 南京 210023)

摘 要:神经网络算法是深度学习研究的重点,遗传算法是一种自适应优化搜索算法,模拟退火算法是寻找最优解的算法,本文主要分析了神经网络,遗传算法和模拟退火算法的特点和缺陷,研究BP神经网络和遗传模拟退火算法相结合的技术,从发挥算法的优点基础上,提出了一个基于模拟退火遗传算法的BP神经网络模型,并应用于某观影俱乐部,作为新电影上映预测和用户推荐,实验结果表明:该算法在收敛性和准确率上都有较好的效果。

关键词:BP神经网络;模拟退火;遗传算法;收敛

中图分类号:TP391 **文献标识码:**A

A Predication Model of BP Neural Network Based on Genetic Simulated Annealing Algorithms

JIANG Meiyun

(Department of Computer and Software, Nanjing Institute of Industry Technology, Nanjing 210023, China)

Abstract:The Neural Network Algorithm is a focus of deep learning, Genetic Algorithm is to search adaptive optimization and Genetic Simulated Annealing Algorithm is to find the optimal solution. Analyzing the characteristics and defects of these three algorithms as well as studying the technology of combining BP neural network with genetic simulated annealing algorithm, this paper proposes a new BP neural network model based on genetic simulated annealing algorithm, from getting the advantages of the algorithm, and the new algorithm has been used in a movie club, where the experimental results show that this algorithm has better effects on the convergence and accuracy.

Keywords:BP neural network; simulated annealing; Genetic Algorithm; convergence

1 引言(Introduction)

神经网络作为人工智能的重要组成部分,一直是神经科学、认知科学、行为科学、计算机科学等关注的焦点。自AlphaGo击败职业围棋选手,第一次战胜人类围棋冠军以来,人类再次掀起了研究机器学习,深度学习的热潮。神经网络作为深度学习的基本模型,吸引了大批研究人员的注意,在神经网络的理论模型,开发工具,学习算法等方面经过多年研究,在语音识别,图像识别,非线性预测模型,模式识别等方面取得令人瞩目的成就。我们国家决定从今年开始大力发展基于大数据的人工智能,即机器学习和深度学习。

人工神经网络^[1]是能模拟和反映人脑部分特征的一种仿真信息处理系统,该系统基于模仿人脑系统的结构和功能而设计。从神经网络建模或神经网络学习预测的角度分析,神经网络都必须包含神经网络拓扑结构^[2],即学习模型,神经元特征函数和目标函数,机器学习算法,存在对训练样本要求高,数据量要求大的困难,而对于小数据训练样本则容易导

致误差大,预测精度低等局限性。

遗传算法^[3,4]GA是一种仿生算法,该算法模拟生物在自然环境中的遗传和进化过程而形成的一种自适应全局优化概率搜索算法。遗传算法以求得最适应环境个体为目的,从任意一个初始群体出发,经过选择、交叉、变异操作来产生一个更适应环境的新的个体,经过多次迭代,使得新群体进化到搜索空间中越来越好的区域,经过一代代的进化,最后收敛到一群最适应环境的个体,求得最优解。

模拟退火算法^[4]SA从设定一个较高温度开始,伴随温度参数的不断下降,结合概率突跳特性在解空间中随机寻找目标函数的全局最优解的过程。

由于BP神经网络易于陷入局部极小值,收敛速度慢等问题^[2,5],采用梯度下降寻优迭代次数多,效率低下,而GA和SA都是基于概率分布机制的优化算法,两者结合,使用SA可以避免GA早熟问题,从而增强全局和局部搜索能力,本文在此基础上提出模拟退火遗传算法优化的BP神经网络算法,并

建立模型用于预测。

2 基于遗传模拟退火的BP神经网络(BP neural network based on genetic simulated annealing algorithm)

2.1 遗传模拟退火算法

遗传模拟退火算法是将遗传算法和模拟退火算法相结合, 模拟退火算法具有较强的全局搜索能力寻得最优解的效率高, 并能使搜索过程避免陷入局部最优解, 但模拟退火算法对全局解空间了解不多。而遗传算法的局部搜索能力较差, 易于陷入局部最优解, 但把握搜索过程的总体能力较强, 遗传模拟退火算法以提高算法效率为目的, 结合了遗传算法和模拟退火算法的优点。

遗传模拟算法^[3]从一组随机产生的初始解开始进行全局最优解的搜索过程, 该过程正好符合反向神经网络解空间最优值的搜寻, 通过选择、交叉、变异等遗传操作产生一组新的个体, 然后对所产生的各个个体进行模拟退火过程, 类似随机梯度下降法, 以其结果作为下一代群体中的个体。

2.2 BP神经网络

BP神经网络^[6,7], 也叫反向传播神经网络, 按反向误差传播算法训练权重和阈值的多层神经网络, 采用梯度下降算法作为优化函数, 对权重和阈值的误差求导, 获得新的权重和阈值来更新整个神经网络的权重和阈值, 以求得全局最小值。误差计算通常用LSM^[8]公式, 计算如下:

$$E(w) = \frac{1}{2} \sum_{d \in D} \sum_{k \in outputs} (t_{kd} - o_{kd})^2$$

其中, $outputs$ 是神经网络输出集合, D 是所有训练样本的集合, t_{kd} 是和训练样本 d 对应的第 k 个真实值, o_{kd} 是与训练样例 d 对应的第 k 个输出值。这个函数的解空间有对应无数函数, 构成了一个搜索误差曲面^[8], 需要在这个解空间上找到这个曲面上的最小值点即最优解。问题转化为在曲面空间求极值, 为了保证能找到最小值, 确保算法快速收敛, 优化效率提升, 合适的学习步长是一个重要参数, 不宜太大也不宜太小。

2.3 基于遗传模拟退火的BP神经网络

利用模拟退火算法^[3]优化遗传算法并利用遗传算法的全局寻优搜索功能, 获得每一次的最优种群(权重和阈值), 接着利用BP的局部寻优特征反向寻得最优值^[9]。

3 基于遗传模拟退火的BP神经网络算法及模型(The BP neural network algorithm and model based on genetic simulated annealing algorithms)

3.1 数据模型

假设输入节点为三个 i_1 、 i_2 、 i_3 , 隐藏节点为两个 h_1 、 h_2 , 输出节点为一个 o 的三次神经网络, 随机生成初始BP神经网络权值和阈值。

输入层到隐藏层权值 $W = [[ih_{11}, ih_{12}], [ih_{21}, ih_{22}], [ih_{31}, ih_{32}]]$

隐藏层的阈值 $t = [t_1, t_2]$

隐藏层到输出层权值 $OW = [[ho_{12}], [ho_{21}]]$

输出层的阈值 $OT = [ot_3]$

为了使用遗传模拟退火对BP神经网络的权值和阈值进行训练优化, 根据任务, 可以建立以下四元组:

$$\Sigma = (I, Q, T, X)$$

其中, $I = \{i_1, i_2, \dots, i_m\}$ 是初始输入集合, 即神经元节点的值;

Q 是一个权值 W 的矩阵, 其原始值随机生成;

T 是一个阈值的矩阵, 原始值随机生成;

X 是一个 $w \times t$ 的矩阵, 对应输入层到第一个隐藏层的阈值矩阵, 获知第一个隐藏层到第二个隐藏层的阈值矩阵……

设计输出误差作为目标函数: $\Delta di = \sum_{m=i}^j (y_{xm} - y_m)^2$, 算法求解的过程就是利用模拟退火算法寻求最优解的过程。

3.2 算法

Begin

While 不满足性能要求 do

Temp = 10;

‘产生初始任务集;

$\Sigma = (I, Q, T, X)$;

$TA = [I, R[1..n]]$;

当 temp == 1 循环

$i' = \text{crossover}(i)$; ‘对样本进行交叉运算

$i'' = \text{mutation}(i')$; ‘样本进行变异运算

‘根据目标函数对每个个体进行模拟退火即选择;

$\Delta di = \sum_{m=i}^j (y_{xm} - y_m)^2$; ‘目标函数

如果 ($\Delta di < 0$)

则采用新方案

否则 { 计算 $\text{prob} = \exp\left\{\frac{\Delta di}{k * \text{temp}}\right\}$;

随机产生一个 $\text{rand}(0 < \text{rand} < 1)$;

如果 ($\text{prob} < \text{rand}$) 则采用新方案;

否则放弃新方案;

}

$\text{temp} = \text{temp} * \alpha$;

BP神经网络算法训练

end;

3.3 算法分析

该算法的优化过程是基于对目标函数的最小化过程, Δdi 的最小化过程模拟自然界的退火过程, 由一个逐步冷却的温度 temp 来控制, 一个冷却机制 $\text{Temp}_{i+1} = \alpha * \text{temp}_i$, α 称为冷却率, $0 < \alpha < 1$, 此处 α 用作神经网络的步长, 步长的大小影响最小化的过程, 模拟退火算法易于避免陷入局部最优, 结合遗传算法全局搜索提高算法效率。

K 是玻尔兹曼常数, temp 也可以看成是一个常数。

3.4 数据模型

首先需要建立一个强大的数据银行来建立模型, 这些数据从 MovieLens 提供的数据库获得^[10], 这些数据包括六个 csv 文件, 一个影评文件, 包括 userId 、 movieId 、 rating 、 timestamp ; 一个标签文件, 包括 userId 、 movieId 、 tag 、 timestamp ; 一个电影文件, 包括 movieId 、 title 、 genres ; 一个链接文件(电影来源), 包括 movieId 、 imdbId 、 tmdbId ; 两个标记相关性文件, 包括 movieId 、 tagId 、 relevance 、 tagId 、 tag 。这些数据显示每个观众观影信息及影评, 本文的目标是利用遗传模拟退火的BP神经网络模型, 对这些数据建立模型, 通过对模型的训练进行预测。

预测某观影俱乐部中的n位会员对于某部电影的评分范围,首先对电影tag和分类数据进行汇总处理,得到电影的特征(电影类型和n个最高的tag),计算电影A与某个用户曾经看过所有电影的相似度以相似度最高的影片的评分,作为评分。实验数据描述见表1。

表1 实验数据描述

Tab.1 Description of data

数据	条件属性	学习样本	测试样本	误差函数
5000	8	4500	50	MSE

3.5 实验结果与分析

通过特征提取降维操作,得到核心属性为五个(A1,A2,A3,A4,A5),见表2。

表2 实验数据核心属性

Tab.2 Key attributes of data

userId	movieId	tag	timestamp	relevance
18	4141	Mark Waters	1240597190	3302500000000000
65	208	Dark hero	1368150078	05699999999999995
65	521	Noir thriller	1368149983	67625

采用十折交叉选取数据进行实验^[11],该实验预测某一影片的评分,loss是损失函数的大小,val_loss是validation_loss,指验证集的loss,均方误差。并通过Keras搭建神经网络仿真平台获得以下实验结果^[12]。

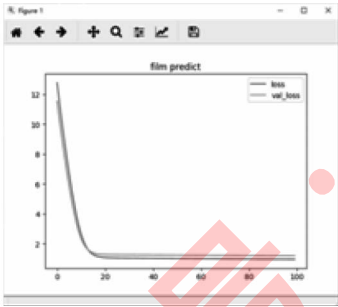


图1 迭代次数和误差函数

Fig.1 Iteration number and error function

从图1可以看出,基于模拟退火的BP神经网络有较好的收敛性,在迭代20次左右算法趋于收敛。收敛速度快,无锯齿现象。

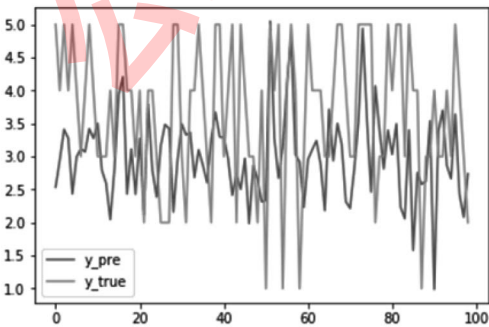


图2 真实评分和预测评分对比

Fig.2 Comparison of real score and predicted score

从图2可以看出,系统的预测值和真实值之间的比较,通

过调参后预测值可以更接近真实值。

4 结论(Conclusion)

传统的神经网络存在收敛效果差,易于陷入局部最小值等问题^[13],在本文中引入遗传算法和模拟退火算法,模拟退火符合随机梯度下降优化算法需求,能提高收敛速度和减少迭代次数,提高算法效率,遗传算法是一种全局搜索优化算法,能改进传统神经网络陷入局部最优解的问题,从而提高算法的效率。本算法结合遗传算法和模拟退火算法,使得传统BP神经网络算法在经过遗传模拟退火算法优化后不仅提高了收敛速度^[14],而且提高了预测精度。

参考文献(References)

[1] Lin Guangfeng,Caixia Fan,Hong Zhu,et al.Visual feature coding based on heterogeneous structure fusion for image classification[J].information Fusion,2017(36):275-283.

[2] 王崇骏,于汶潞,陈兆乾,等.一种基于遗传算法的BP神经网络算法及其应用[J].南京大学学报自然科学版,2003,5:1-4.

[3] 蒋美云.遗传模拟退火算法在云调度中的应用研究[J].计算机与现代化,2013(13):3-4.

[4] 何炎祥,陈萃萌.Agent和多Agent系统的设计与应用[M].武汉大学学术丛书:64-70.

[5] F Guzmán,S Joty,L Márquez,et al.Machine translation evaluation with neural networks[J].Computer Speech & Language[J].2017(45):180-200.

[6] 刘浩然,赵翠香,李轩,等.一种基于改进遗传算法的神经网络优化算法研究[J].仪器仪表学报,2016,37(7):1573-1580.

[7] 高玉明,张仁津.基于遗传算法和BP神经网络的房价预测分析[J].计算机工程,2014,40(4):187-191.

[8] 曾小华,李广含,宋大风,等.基于遗传算法优化的BP神经网络侧翻预警算法[J].华南理工大学学报自然科学版,2017,45(2):30-38.

[9] 何涛,张洪伟,邹书蓉.特征提取与多目标机器学习研究及应用[J].四川理工学院学报(自然科学版)2013(1):30-35.

[10] MR Mahdiani.A Modified Neural Network Model for Predicting the Crude Oil Price[J].Intellectual Economics,2017:22-30.

[11] <https://www.superdatascience.com>.

[12] AJD Silva,WRD Oliveira Comments on "Quantum artificial neural networks with applications"[J].Information Sciences,2016,370-371:120-122.

[13] JürgenSchmidhuber.Deep Learning in neural networks:An overview[J].Neural Networks,2015(61):85-117.

[14] 王倩,王莹,邱继勤.基于灰色GM(1,1)和BP神经网络的房价预测[J].时代金融,2012(36):332-333.

作者简介:

蒋美云(1977-),女,硕士,讲师.研究领域:人工智能,机器学习,深度学习.