

BP 神经网络的优化研究与应用

潘文婵, 刘尚东

(南京邮电大学 计算机学院, 江苏 南京 210023)

摘要:针对 BP 神经网络自身的一些局限性, 诸如易陷于局部极小、网络收敛速度慢、训练时间长等, 提出一种改进 BP 神经网络的研究方案, 通过改变传统的固定学习率, 引入动态变化, 根据均方误差的变化而改变学习率。在误差曲面平坦区域增大学习率, 在误差变化剧烈的区域减小学习率, 从而加快算法的收敛速度, 避免陷入局部极小值。文中在传统 BP 神经网络中使用动态学习速率, 并融合参数可调激活函数来改进 BP 神经网络。采用公认完备、性能优异的 KDD Cup99 数据集, 分别对改进算法和传统 BP 算法进行了对比实验。实验结果表明, 与传统 BP 神经网络算法相比, 改进算法极大地提高了训练速度, 具有训练误差更小、预测精度更高的优点。

关键词:BP 神经网络; 学习率; 均方误差; 深度学习

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2019)05-0074-03

doi:10.3969/j.issn.1673-629X.2019.05.016

Optimization Research and Application of BP Neural Network

PAN Wen-chan, LIU Shang-dong

(School of Computer Science, Nanjing University of Posts and Telecommunications,
Nanjing 210023, China)

Abstract: Aiming at some limitations of BP neural network, such as easy to be trapped in local minimum, slow convergence speed and long training time, we propose a scheme of improving BP neural network. By changing the traditional fixed learning rate and introducing dynamic change, the learning rate can be changed according to the change of mean square error. The learning rate is increased in the flat area of the error surface, and decreased in the area where the error changes sharply, so as to accelerate the convergence speed of the algorithm and avoid falling into the local minimum. In the traditional BP neural network, we use the dynamic learning rate with the parameter adjustable activation function to improve the BP neural network. The improved algorithm and the traditional BP algorithm are compared by using the KDD Cup99 data set, which is recognized to be complete and excellent in performance. Experiment shows that compared with traditional algorithms based on BP neural network, the proposed method has greatly improved training speed and has the advantages of smaller training error and higher prediction accuracy.

Key words: BP neural network; learning rate; mean square error; deep learning

0 引言

网络中存在大量复杂的攻击行为, 这些攻击已经成为网络和信息安全的主要威胁。防火墙技术、VPN 技术等安全技术立足于防, 并不能完全阻止入侵的发生。入侵检测是防火墙之后的第二道“安全闸门”, 这种动态主动防御技术弥补了传统安全技术的不足。入侵检测系统 (intrusion detection system, IDS) 通过检查操作系统的审计数据或捕获网络数据包信息, 检测网络或系统中违反安全策略的行为, 防止系统数据的泄漏、篡改和破坏^[1]。IDS 能够对系统中未经授权的访

问或异常现象、活动与事件进行跟踪、审计、识别和检测。

随着人工智能技术的发展, 神经网络技术开始应用于入侵检测系统中。神经网络具有并行计算、自适应学习和调整, 可以处理不完整有失真的数据, 抗干扰能力强等特性, 在入侵检测领域应用广泛^[2]。1986 年, Rumelhart 和 McClelland 提出了 BP 神经网络的概念, 机器学习进入快速发展时期。在这一阶段, 不断涌现出更优秀的算法, 推动了人工智能在语音识别、图形图像处理以及自然语言处理方面的进展。

收稿日期: 2018-05-11

修回日期: 2018-09-13

网络出版时间: 2018-12-21

基金项目: 国家重点研发计划 (2017YFB1401301); 南邮实验室工作研究课题 (2018XSG06)

作者简介: 潘文婵 (1983-), 女, 硕士, 实验师, 研究方向为网络安全。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20181221.1445.004.html>

神经网络算法有很多种,其中比较常用的是反向传播(back propagation,BP)神经网络。但是随着应用范围的逐步扩大,BP 神经网络也暴露出一定的缺点和不足,比如处理复杂问题时易陷入局部最小、BP 算法学习收敛速度慢等^[3]。对此,文中提出一种 BP 神经网络的改进算法,以改变传统固定学习率,提高 BP 网络收敛速度和入侵检测准确度。

BP 算法作为传统训练多层网络的典型算法,易出现局部最优。2006 年,加拿大多伦多大学教授机器学习领域的泰斗 Geoffrey Hinton 等提出了深度学习(deep learning)的概念,开启了深度学习在学术界和工业界的浪潮。深度学习本质上是深层人工神经网络,通过构建具有很多隐层的机器学习模型和海量的训练数据,以更少的参数与更深的结构来学习更有用的特征,从而提升分类或预测的准确性^[4-6]。

1 BP 神经网络

BP 神经网络是一种多层前馈网络,连接权值调整采用反向传播学习算法,BP 神经网络算法即误差逆传播算法。它的学习规则是使用梯度下降法,通过反向传播来不断调整网络的权值和阈值,使网络的误差平方和最小^[7-8]。BP 神经网络模型的拓扑结构包括输入层(input)、隐含层(hidden layer)和输出层(output layer)。

BP 神经网络模型一般由三层构成,隐含层可以是一层或多层。令 X 代表输入矢量, Y 代表输出矢量,定义各层间连接权矩阵, ω 、 W 分别代表相邻两层间各节点的连接权矩阵。 ω_{ij} 表示输入层神经元与隐含层神经元之间的连接权值, W_{jk} 表示隐含层神经元和输出层神经元之间的连接权值。网络的学习过程,通过对连接权值和阈值进行调整修正,减小输出误差,使网络实现给定的输入输出映射关系^[9-10]。

一般而言,只需包含一个足够多神经元的隐层,就能以任意精度逼近任意复杂度的连续函数。BP 神经网络从输入到输出,具有很强的非线性映射能力。理论研究证明,可以用三层神经网络以任意精度逼近任意非线性连续函数,这使得其特别适合于解决内部机制复杂的问题。

2 BP 神经网络算法

由于 BP 算法存在一定的局限性,降低了 BP 神经网络在实际应用中的效率,因此许多专家学者提出了多种改进的措施^[11-12],从学习率、误差函数、激励函数、网络结构以及优化算法等方面进行研究。

标准 BP 算法的学习率固定不变,通常将学习率定为常量,因此其收敛速度慢,易陷入局部最小值。选

择一个合理的学习率,需要根据学习情况来调整学习率。学习率控制着沿反梯度方向下降的步长,一般常把学习率设置为 0.1,有时更新权重时会将输出层与隐含层设置为不同的学习率。如果学习率过小,收敛速度会很慢,如果学习率过大,可能引起震荡而使网络不稳定。

在均方误差变化较平坦的区域,为了加快收敛速度,则增大学习率。而在误差变化剧烈的区域,训练出现振荡,则减小学习率,有助于使网络从误差曲面的局部极小值中跳出。

3 改进的 BP 神经网络优化算法

对 BP 神经网络的有效改进是一个值得研究的问题。为了提高 BP 神经网络的收敛速度和预测精度,选择合适的学习率很重要^[13-14]。文中从以下几个方面来探究改进 BP 神经网络的方法。

3.1 学习率的改进

在训练开始时,随机设置初始学习率,使用误差反向传播算法来训练权值参数,寻找误差函数的最小值。以隐含层节点与输出层节点的连接权值调整为例,用动态学习速率法来缩短训练时间。具体步骤如下:在更新权值之前,设置初始学习率 η 。如果均方误差减小,则学习率的取值偏小,学习率将乘以一个大于 1 的因子 α 。如果均方误差在权值更新后增加,且该值超过了预设的某个百分数,如:1%~5%,则需要减小学习率,学习率乘以一个大于零且小于 1 的因子 β 。如果均方误差增加,但是增加值小于预设的数值区域 ξ ,则学习率保持不变。

将对动态学习速率 η 的调整公式定义如下:

$$\eta(n+1)=\begin{cases}1.05\eta(n), & E(\eta+1)<E(\eta) \\ 0.75\eta(n), & E(\eta+1)>E(\eta)\times\text{er} \\ \eta(n), & E(\eta)\leq E(\eta+1)\leq E(\eta)\times\text{er}\end{cases}$$

(1)

其中,er 为最大误差率。

用同样的算法来调整输入层与隐含层各节点间的连接权,找到最优学习率,使得网络输出更符合实际。

3.2 采用带有可调因子的激活函数

BP 神经网络的重点是网络误差与权值调整。网络误差 $E(n)$ 是各层权值 ω_{ij} 、 W_{jk} 的函数,因此调节权值可改变误差。

隐含层至输出层权值调整量为:

$$\Delta\omega_{ij}=-\eta\frac{\partial E(n)}{\partial\omega_{ij}}=-\eta\cdot\frac{1}{\partial\omega_{ij}}\cdot\frac{1}{2}\sum_{j=0}^{n-1}(d_j-y_j)^2=-\eta(d_j-y_j)\cdot\frac{\partial d_j}{\partial\omega_{ij}}=\eta(d_j-y_j)$$

$$y_j) \cdot f_2' \cdot v_i = \eta \delta_{ij} \cdot v_i \quad (2)$$

其中, $\delta_{ij} = (d_j - y_j) \cdot f_2' = (d_j - y_j)y_j(1 - y_j)$ 表示隐含层节点的局部梯度; f_2' 表示输出层转移函数对输入变量的导数; δ 为学习规则, 通过改变神经元之间的连接权值来减少系统实际输出和期望输出的误差。在网络学习的过程中, 各层神经元连接权值的学习沿输出层的输出误差 E 的负梯度方向变化。

通过式 2 可得, $\partial E(n)/\partial \omega_{ij} = (d_j - y_j)f_2' \cdot v_i$, 当网络权值进入误差曲面底部的平坦区时, $\Delta \omega_{ij}(k+1) = \Delta \omega_{ij}(k)$, k 为训练次数, $f_2' \approx 0$, 则 $\partial E(n)/\partial \omega_{ij} \approx 0$ 。可能使网络权值的调节过程几乎停顿, 延长了网络的收敛时间。

计算输出层的实际输出时, 如果 $|E(\eta+1) - E(\eta)| < \varepsilon$, 则导入可调函数 $f_2 = 1/(1 + \theta \cdot \exp(-\text{net}))$, 否则 $f_2 = 1/(1 + \exp(-\text{net}))$ 。 θ 是可调因子, $0 < \theta < 1$, 根据具体实验而定, ε 一般取值较小, 0.000 1 即可。计算误差 ΔE , 判断是否位于局部最小点。当误差变化很小时, 使用带可调因子的激活函数, 避免误差函数长时间处于平坦区, 从而防止 $\Delta \omega_{ij} = 0$ 的出现, 有助于使网络从误差曲面局部极小值中跳出。

4 仿真实验和结果分析

运用 MATLAB 仿真来验证改进之后的 BP 神经网络的性能。文中采用美国麻省理工学院林肯实验室提供的 KDD Cup99 数据集, 以其中的 corrected. gz 数据集为研究对象。数据集中每个网络连接记录有 41 个属性值, 利用信息熵理论, 最终选取所含信息量较大的前 14 维特征用于实验。

数据预处理分为两步: 数值标准化和数值归一化。在实验中, 最大迭代次数 3 000, 初始学习率 0.01, 精度值 0.001, 输入节点数为 10, 输出节点数为 1, 并设置不同的隐含层节点、参数进行对比实验。

从 2 000 组样本数据中随机提取 1 900 组作为训练数据, 另外 100 组作为测试数据, 用于测试网络的性能。分别对传统 BP 神经网络以及改进算法进行比较, BP 网络学习率的优化模型采用动态学习率与参数可调激活函数相结合的方法对网络进行训练。

从迭代次数与误差精度两个方面来分析传统 BP 神经网络与改进后的 BP 神经网络的拟合性能。

标准 BP 算法的误差曲线如图 1 所示, 可以看出逼近效果不理想, 训练 3 000 次还未达到所要求的误差精度, $MSE=0.026$ 。采用不同的参数对改进后的算法进行了多次对比实验, 结果发现当学习率增长比为 1.05, 学习率下降比为 0.75, 误差函数增加为迭代前的 1.02, 隐含层节点数为 11 时, 训练效果最佳。

后的 BP 神经网络迭代 118 次就达到了均方误差精度的要求, $MSE=0.000 6$ 。改进后的 BP 神经网络与传统神经网络相比, 可以提高收敛速度, 避免陷入局部极小, 而且较为明显地改善了学习精度。

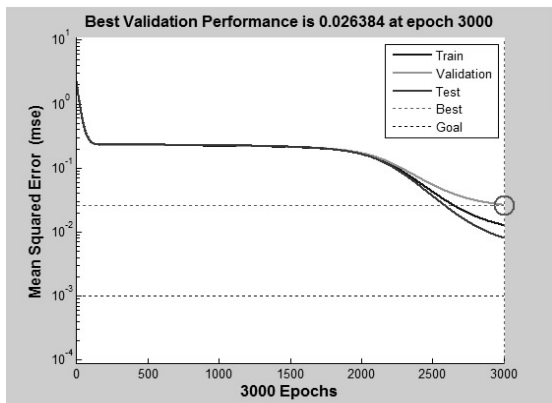


图 1 标准 BP 算法的误差性能曲线

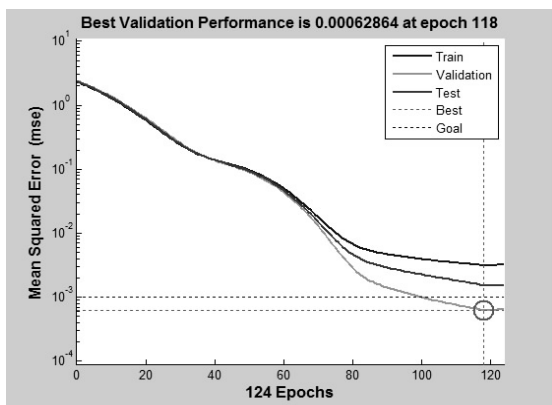


图 2 改进学习率算法的误差性能曲线

5 结束语

学习率是影响神经网络训练速度和性能的重要因素。针对传统 BP 神经网络收敛速度较慢, 预测精度不高, 极易陷入局部最小点等缺陷, 提出了一种动态学习率与参数可调激活函数相结合的改进方法。实验结果表明, 该算法提高了 BP 网络的收敛速度和误差精度, 避免陷入局部极小值。下一步的研究将考虑深度学习, 在应用方面, 深度学习在图像识别、语音识别等领域展现出巨大的优势, 以进一步推动人工智能技术的发展。

参考文献:

- [1] ZHANG Li, WANG Fulin, SUN Ting, et al. A constrained optimization method based on BP neural network[J]. Neural Computing and Applications, 2018, 29(2): 413-421.
- [2] HOU Yue, ZHAO Long, LU Huaiwei. Fuzzy neural network optimization and network traffic forecasting based on improved differential evolution[J]. Future Generation Computer Systems, 2018, 81: 425-432.

公开,这时需要提供对应的加密密钥才能看到私密信息具体内容,保证了信息数据的安全性,保护了资源提供者的隐私。在去中心化的特征下,保证了制造资源信息共享的同时,也保证了在某一节点可能存在数据丢失的情况下,对全网影响几乎微乎其微,使得云制造服务平台中的数据安全得到了保证。

3 结束语

提出一种基于区块链的云制造服务数据记录与存储方法。首先,将区块链技术与云平台相结合,提出交易区块链体系结构,对交易区块链中的数据层、网络层、共识层、激励层、合约层及应用层进行了分析。然后,对应用层中数据记录及存储进行研究,提出了云制造平台中数据记录的方法,并将该方法应用于设备资源信息记录中。结果表明,该方法明显增强了数据记录的安全性,提高了云制造平台中的资源共享效率。

参考文献:

[1] 曹佳硕. 基于 RDF 的云制造资源数据存储及检索方法的研究与实现[D]. 北京:北京交通大学,2013.

[2] 赵 淳,张 霖,任 磊,等. 面向云制造交易过程的仿真平台[J]. 计算机集成制造系统,2016,22(1):25-32.

[3] 林小驰,胡叶倩雯. 关于区块链技术的研究综述[J]. 金融市场研究,2016,45:97-109.

[4] 沈 鑫,裴庆祺,刘雪峰. 区块链技术综述[J]. 网络与信息

(上接第 76 页)

[3] LIU Tianhua, YIN Shoulin. An improved particle swarm optimization algorithm used for BP neural network and multimedia course-ware evaluation [J]. Multimedia Tools and Applications,2017,76(9):11961-11974.

[4] SUN Aixi, JIN Xue, CHANG Yubo. Research on the process optimization model of micro-clearance electrolysis-assisted laser machining based on BP neural network and ant colony [J]. International Journal of Advanced Manufacturing Technology,2017,88(9-12):3485-3498.

[5] PENG J S. Multi-objective optimization of vibration characteristics of steering systems based on GA-BP neural networks[J]. Journal of Vibroengineering,2017,19(5):3216-3229.

[6] YANG Xinshe, KARAMANOGLU M, HE Xingshi. Flower pollination algorithm:a novel approach for multiobjective optimization[J]. Engineering Optimization,2013,46(9):1222-1237.

[7] 邓文杰. 基于聚粒子群算法的神经网络权值优化方法[J].

安全学报,2016,2(11):11-20.

[5] CHRISTIDIS K,DEVETSIKIOTIS M. Blockchains and smart con-tracts for the Internet of Things[J]. IEEE Access,2016,4:2292-2303.

[6] SWAN M. Block chain thinking:the brain as a decentralized autonomous corporation [J]. IEEE Technology and Society Magazine,2015,34(4):41-52.

[7] 李伯虎. 云制造[M]. 北京:清华大学出版社,2015.

[8] 陈友玲,刘传彪,阳玮琦,等. 云制造环境下能力资源需求的评价与选择[J]. 计算机集成制造系统,2017,23(10):2304-2312.

[9] 黄 辉,王 哲,纪玉娇,等. 分布式网络化云制造模式分析[J]. 现代制造工程,2017(11):36-43.

[10] 朱李楠,王万良,沈国江. 基于改进差分进化算法的云制造资源优化组合方法[J]. 计算机集成制造系统,2017,23(1):203-214.

[11] 张 霖,罗永亮,范文慧,等. 云制造及相关先进制造模式分析[J]. 计算机集成制造系统,2011,17(3):458-468.

[12] XU Xun. From cloud computing to cloud manufacturing[J]. Robotics and Computer-Integrated Manufacturing,2012,28(1):75-86.

[13] 常瑞云,周井泉,许 斌,等. 基于离散人工群算法的云制造服务组合[J]. 计算机技术与发展,2016,26(7):177-182.

[14] 齐 轩,刘茜萍. 基于用户偏好的可信 QoS 服务选择方法[J]. 计算机技术与发展,2016,26(8):43-47.

计算机技术与发展,2017,27(10):16-18.

[8] 张 志,杨清海. 基于 BP 神经网络和改进 D-S 证据理论的目标识别方法[J]. 计算机应用与软件,2018,35(3):151-156.

[9] 付晓明,王福林,尚家杰. 基于多子代遗传算法优化 BP 神经网络[J]. 计算机仿真,2016,33(3):258-263.

[10] 沈夏炯,王 龙,韩道军. 人工蜂群优化的 BP 神经网络在入侵检测中的应用[J]. 计算机工程,2016,42(2):190-194.

[11] 吴文铁,宋日聪,李 敏. 蚁群优化神经网络的网络流量混沌预测[J]. 计算机工程与应用,2012,48(34):97-101.

[12] 严 旭,李思源,张 征. 基于遗传算法的 BP 神经网络在城市用水量预测中的应用[J]. 计算机科学,2016,43:547-550.

[13] 高 隽. 人工神经网络原理及仿真实例[M]. 北京:机械工业出版社,2007.

[14] 周志华,曹存根. 神经网络及其应用[M]. 北京:清华大学出版社,2004.