

曹亮. 基于 Adam 算法 BP 神经网络的矿井环境瓦斯浓度预测模型[J]. 湖南科技大学学报(自然科学版), 2024, 39(1): 18-23. doi:10.13582/j.cnki.1672-9102.2024.01.003

CAO L. On Mine Environmental Gas Concentration Prediction Model Based on Adam Algorithm BP Neural Network [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2024, 39(1):18-23. doi:10.13582/j.cnki.1672-9102.2024.01.003

基于 Adam 算法 BP 神经网络的 矿井环境瓦斯浓度预测模型

曹亮

(中国煤炭科工集团 煤炭科学技术研究院有限公司, 北京 100013)

摘要:为了解决现有矿井环境瓦斯浓度预测方法无法处理大数据量、适应性差、误差较大、易陷入局部最优等问题,提出一种基于 Adam 算法的改进型 BP(Back Propagation)神经网络模型,模型适用于矿井多环境参数下,对某区域内环境瓦斯浓度进行预测.对监测监控系统采集到的真实数据进行归一化处理并形成数据集,通过将 Adam 算法与 BP 网络模型进行有效结合形成新的网络模型.运用训练集对模型进行训练及调优后,迭代次数在 1 200 次后损失率趋于平稳,验证集预测的结果整体平均误差率为 1.258%,结果表明:该优化模型提高了网络训练速度,且避免了传统 BP 模型容易陷入局部最小的缺点,同时降低了预测的相对误差.

关键词: 瓦斯浓度; BP 神经网络; 激活函数; Adam 算法; 梯度消失

中图分类号: TD824.2 **文献标志码:** A **文章编号:** 1672-9102(2024)01-0018-06

On Mine Environmental Gas Concentration Prediction Model Based on Adam Algorithm BP Neural Network

CAO Liang

(Coal Science and Technology Research Institute Co., Ltd., China Coal Science and Industry Group, Beijing 100013, China)

Abstract: In order to solve the problems of existing mine environment gas concentration prediction methods that cannot handle large data volume, poor adaptability, large errors, and easy to fall into local optimality, an improved BP (Back Propagation) neural network model based on Adam algorithm is proposed. The model is suitable for multiple environments in mines. Under the parameters, it predicts the environmental gas concentration in a certain area. The real data collected by the monitoring system is normalized and formed into a data set, and a new network model is formed by effectively combining the Adam algorithm with the BP network model. After training and tuning the model with the training set, the loss rate stabilized after 1200 iterations, and the overall average error rate of the verification set prediction is 1.258%. The results show that the optimized model has improved the network training speed and avoided the shortcoming that the traditional BP model is easy to fall into local minimum, and meantime it reduces the relative error of prediction.

Keywords: gas concentration; BP neural network; activation function; Adam algorithm; vanishing gradient

收稿日期: 2021-11-26

基金项目: 煤炭科学技术研究院科技发展基金资助项目(2020CX-II-27)

* 通信作者, E-mail: caoliang@ccrise.cn

煤炭行业无论是在优化能源结构政策实施之前,还是之后,都是我国经济与社会发展的支柱性产业,甚至未来较长一段时间内,其地位也并不会会有实质性改变^[1].地位虽未改变,但伴随着“碳达峰、碳中和”等措施的实施,在新的形势发展下,煤炭行业唯有提质增效实现优化升级,提升企业整体能力才能在新形势下实现稳定、健康、可持续发展。

智慧矿山概念的提出将矿山信息化提高到了“智慧”的高度^[2],但如何将“智慧”概念落地矿山信息化系统将是一个挑战.煤炭行业管理粗放,数字化意识相对薄弱,企业信息化系统众多,各系统产生的数据量庞大,数据类型纷繁复杂,传统的支持向量机(SVM)算法难以应对大规模的训练样本,同时 SVM 对于解决多分类问题也存在困难^[3].神经网络(Artificial Neural Network, ANN)起源于 20 世纪 40 年代,在过去的 10 年是最具影响力的技术之一,它是深度学习算法的基本组成部分,是 AI 技术的前沿,ANN 的应用使“智慧”这个概念落地成为可能^[4].

我国煤矿事故中环境因素造成的事故占比很高,因此监控系统是煤矿安全生产必不可少的组成部分,系统根据预先设定好的规则对单一环境指标进行超限报警^[5].但矿井内环境复杂,瓦斯浓度受多种环境因素影响^[6],因此井下各类环境指标与瓦斯浓度之间形成了一个典型的非线性关系。

ANN 尤其是反向传播(Back Propagation, BP)类型的神经网络对于处理复杂的非线性数据有很好的拟合能力,本文旨在建立一个 BP 类型的神经网络模型^[7],并对环境数据进行标注,使用标注后的数据对神经网络模型进行训练,通过神经网络模型建立起环境数据与瓦斯浓度之间的非线性关系,再以实际环境数据为基础,对未来一段时间内瓦斯浓度进行预测预警,为井下生产提供安全保障。

1 模型结构

对人类生物神经系统的观察启发了 ANN 的概念原型,ANN 的最小单位为神经元,每个神经元都由输入向量、权值、偏置、激活函数和输出组成,整个网络则由输入层、隐藏层和输出层组成,各层中的神经元与邻近层的神经元完全连接,这样就整体构成一个完整的 ANN 模型^[8].将数据集分为自变量和因变量,ANN 模型的输入层神经元个数与自变量维度相同,输出层神经元个数与因变量维度相同.自变量数据集送入模型对模型进行训练,训练结果会与输入的因变量做比对,比对结果反向反馈给模型,模型根据反馈结果来调整各神经元的权值,经过自变量内数据集对模型的反复训练,使结果趋向于因变量^[9].如图 1 所示。

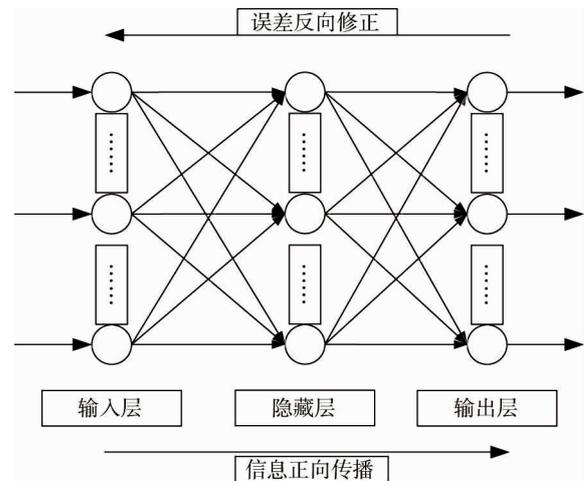


图 1 BP 神经网络模型

2 模型构建与模拟实验

2.1 数据归一化

本文研究内容中使用的数据集,以监测监控系统收集的环境数据为基础.监测监控系统在井下重点区域部署瓦斯浓度、一氧化碳、二氧化碳、氧含量、粉尘浓度、风速及风压传感器,传感器数据以秒为间隔单位上传至监测监控系统,本研究选用鄂尔多斯某矿井 2019 年 1 月—6 月数据作为研究对象,首先对数据进行清洗工作,将脏数据、异常数据及标校数据去除,然后选取前 4 个月数据作为训练集,后 2 个月数据作为验证集。

原始数据的标准化程度对于模型的学习效果有非常大的影响,监控系统采集的环境指标众多,各种指标的量纲不同范围也不同,这就造成数据不标准、不统一的情况出现,进而会造成模型训练时梯度下降慢、收敛慢、最优解精度低等问题,归一化处理可以有效地避免这一问题.常见的归一化方法有 min-max 方法、log 函数、atan 函数及 z-score 标准化等方法,各方法优缺点不同,适用的场景及数据的特点也不同^[10].这里

选择 z-score 归一化方法,此方法基于原始数据的均值和标准差进行数据的标准化,可将各种不同量纲与单位的数据统一转化为一个量纲,很好地解决了监控系统数据种类繁多的问题,提高了数据可比性,削弱了数据解释性^[11].z-score 方法的计算公式为

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

式中: x 为需要被归一化的原始数据集; μ 为原始数据集的平均值; σ 为原始数据集的标准差; z 为标准化后的数据集.

2.2 模型构建

根据人工神经网络的定义及本模型的应用场景,将模型设计为输入层、隐藏层和输出层共 3 层的结构^[12].输入层神经元个数与输入数据集维度相同,输出层神经元个数与预测维度相同.依据神经网络领域研究人员及学者,经过大量实验总结出的经验,此模型隐藏层神经元个数设置为 8 个,因此整个模型如图 2 所示.

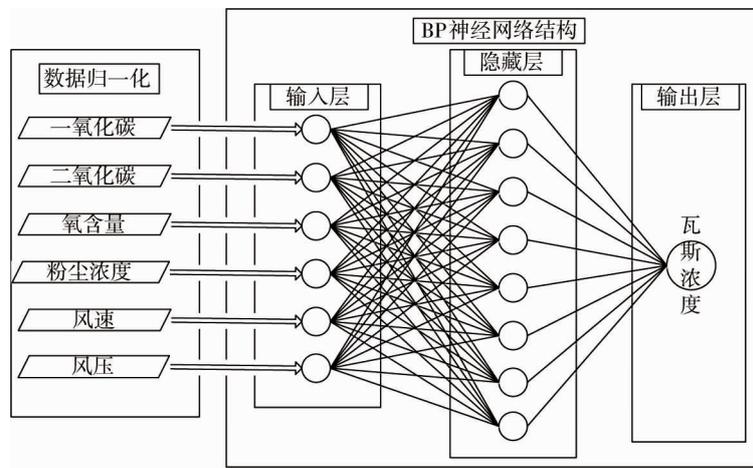


图 2 模型结构

激活函数同样也是影响模型学习效率及预测效果的关键因素,因此综合考虑激活函数特性、数据集特性、网络结构等因素,最终确定合适的激活函数可显著提高网络学习效率,加快网络收敛速度,同时还可提高网络泛化能力^[13].Sigmoid, Tanh, ReLU 等都是比较常用的激活函数,但函数特性各有不同,从函数曲线可以看到 Sigmoid 函数的输出均为正值,随着输入越大或越小,在该处的导数越接近于 0,这样直接带来的问题便是在反向传播过程中,可能部分权重更新的非常小,甚至某些情况下导致部分权重基本不更新,也就出现了梯度饱和或梯度消失的问题.Tanh 函数相对 Sigmoid 函数来说,避免了输出均为正值的情况,但对于梯度饱和与梯度消失问题并未改善.ReLU 函数相对其他函数来说有如下优缺点:

优点:

- 1) 计算速度快,求导简单.
- 2) 不再梯度弥散,ReLU 函数与 Sigmoid 和 Tanh 函数不同,输入大于 0 的情况下,不存在梯度饱和区,几乎不会造成梯度弥散^[14].

3) ReLU 函数,在负轴上输出为 0,虽然在负轴方向上也出现了梯度消失问题,但这也带来了一些好处:即这样会使得被激活的神经元非常稀疏,这和人们熟知的认知神经科学中关于人脑学习时生物神经元激活的稀疏性具有相似性,此特点可以加大网络稀疏性,减少模型的过拟合.

缺点:

- 1) ReLU 的输出不是 0 均值.
- 2) 神经元坏死现象,ReLU 在训练时很“脆弱”,当 $x < 0$ 时,梯度为 0,这个神经元及之后的神经元梯度永远为 0,不再对任何数据有所响应,相应参数永远不会被更新.导致此现象的原因是设置了太大的学习率,因此可以通过设置合理的、较小的学习率进行规避.

综合以上优缺点,考虑到本次研究中训练集数据多为正小数这一特性,在隐藏层选用 ReLU 函数作为激活函数,如图 3 所示

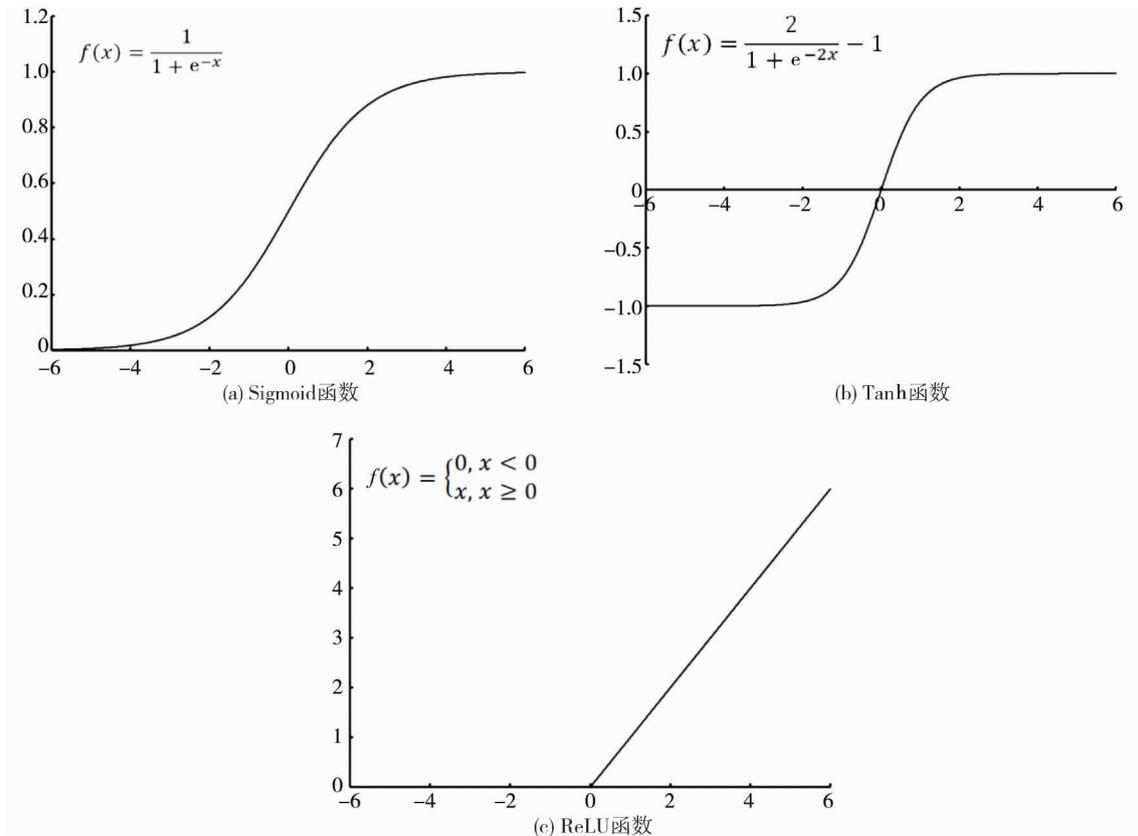


图3 激活函数曲线

模型参数包含迭代次数、初始权值、学习率、误差。初始迭代数为 3 000 次,权值设置为(0,0.01)的随机数,初始学习率为 0.05,以均方根误差作为评判标准。

2.3 模型优化

为了避免算法中出现局部极小问题,同时为了提高模型的预测速度及准确度,对模型的算法进行优化处理。Adam 算法是一种对随机梯度下降法的扩展,基于低阶矩阵的自适应估计,算法计算模型的一阶矩估计和二阶矩估计,根据结果为各个模型设定不同的动态学习率,通过每一次迭代来持续更新模型参数。Adam 不仅保存了历史梯度平方指数衰减平均值 v_t , 还保存了过去梯度指数衰减平均值 m_t , 这一点与物体运动时产生的动量相似,是一个矢量,因此更新时在一定程度上保留了之前的方向^[15]。

在实际应用中,大多数情况下 Adam 方法效果良好,与其他自适应学习率算法相比,其收敛速度更快,学习效果更为有效,而且可以纠正其他优化技术中存在的问题,如学习率消失、收敛过慢或是高方差的参数更新导致损失函数波动较大等问题。虽然有些研究表示,某些极端情况下 Adam 算法会出现不收敛或错过最优解的情况^[16],但分析其研究内容后不难发现,多数都构造了一些较为极端的例子来证实 Adam 失效的可能性。如构造了一个数据集,其拥有少量频次非常低、梯度却非常大的数据点^[17],在实际应用中这种情况发生的可能性非常低,但这也提醒了我们,理解数据对于设计算法的必要性。纵观优化算法的演变历史,均为基于对数据的某种假设而进行的优化,因此算法是否有效,依赖于数据特点是否符合算法特性。

计算公式为

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t; \quad (2)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2. \quad (3)$$

式中: m_t 为当前梯度的一阶动量; v_t 为当前梯度的二阶动量; g_t 为当前的梯度值; β_1 为一阶矩阵指数衰减率,默认值为 0.9; β_2 为二阶矩阵指数衰减率,默认值为 0.999。

由于 m_t 和 v_t 的初始均为零向量,因此衰减率在初始时间可能会偏向零向量,故需要进行偏差校正,公式如下:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1'}; \quad (4)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2'}. \quad (5)$$

式中: \hat{m}_t 为校正后梯度带权平均; \hat{v}_t 为校正后梯度带权有偏方差.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\varepsilon + \sqrt{\hat{v}_t}} \hat{m}_t. \quad (6)$$

式中: ε 为默认值为 10^{-8} ; η 为学习率,初始值为 0.05.

3 网络训练及预测

迭代次数会对模型的学习效果产生一定影响,关于迭代次数、学习率的确定,需要结合结果的准确率,通过不断试验,在试验过程中不断调整迭代次数及学习率,观察此参数下的预测准确率,进而再对迭代次数和学习率进行调整,通过不断的试验在迭代次数、学习率及准确率之间寻找一个平衡,获取最优迭代次数.根据上文中网络参数的设置,对网络进行训练实验,其迭代次数与损失率的关系如图 4 所示.

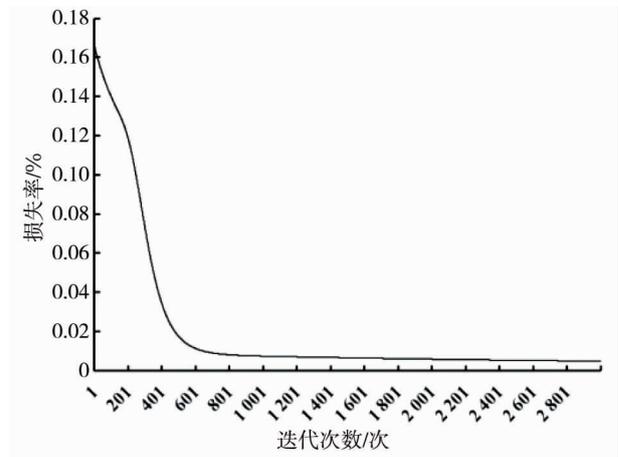


图 4 损失率曲线

通过图 4 可知:迭代次数大于 800 后,损失率看似趋于平稳,但当将曲线局部放大后会观察到,在迭代 1 200 次后,损失率曲线的变化趋于平稳,因此将模型训练迭代数最终确定为 1 200 次.利用前文所述数据集中的训练集对模型进行训练,利用验证集对训练完毕的模型进行预测准确度的比对验证.本文模型基于 Python 语言开发,Python 语言有非常强大的数据及数学运算的能力,能够以最简单的方式实现模型内的各种运算.使用 VS code 作为 IDE 工具,由于数据量较大,故只选取验证集中部分数据进行比对,具体见表 1,瓦斯浓度实测值与预测值的折线对比见图 5,验证集的整体平均误差率为 1.258 %.

表 1 瓦斯浓度实测与预测对比 %

编号	瓦斯浓度		相对误差
	实测	预测	
1	0.237	0.232	2.108
2	0.232	0.227	2.156
3	0.223	0.219	1.790
4	0.221	0.219	0.907
5	0.212	0.215	1.414
6	0.207	0.210	1.449
7	0.213	0.216	1.410
8	0.222	0.217	2.253
9	0.222	0.219	1.353
10	0.211	0.212	0.473
11	0.205	0.210	2.441

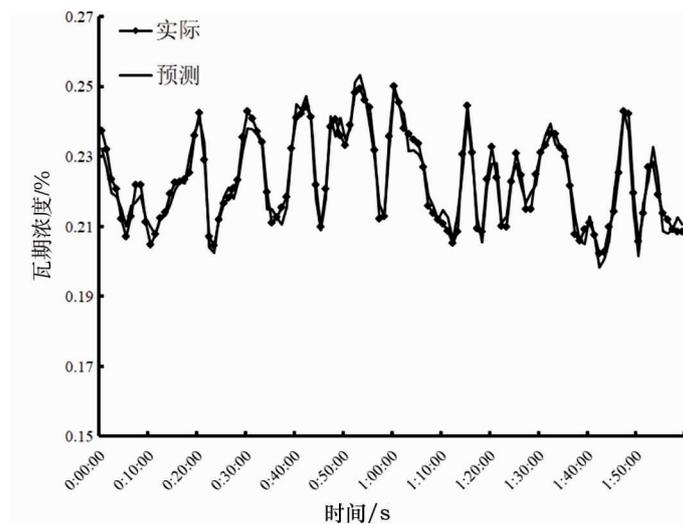


图 5 瓦斯浓度实测值与预测值

4 结论

1) BP 神经网络通过构建多层神经元的连接与反向传播算法,用来学习输入数据与瓦斯浓度之间的复杂关系,而 Adam 算法则被用来优化神经元权重。

2) 此类模型若使用不同场景的训练数据进行训练,则可应用于煤矿各种环境监测预警场景的建模。

参考文献:

- [1] 刘新杰, 宋高峰, 蒋斌斌. 煤炭行业发展历程及展望[J]. 矿业安全与环保, 2019, 46(3): 100-103.
- [2] 阙建立. 智能矿山平台建设与实践[J]. 工矿自动化, 2018, 44(4): 90-94.
- [3] 高高. 支持向量机增量学习算法研究[D]. 鞍山: 辽宁科技大学, 2018.
- [4] 张雨. 基于深度学习的井下瓦斯浓度预测系统设计与实现[D]. 徐州: 中国矿业大学, 2019.
- [5] 董丁稳. 基于安全监控系统实测数据的瓦斯浓度预测预警研究[D]. 西安: 西安科技大学, 2012.
- [6] 潘少波. 基于 RNN 的瓦斯浓度多参数融合预测研究[D]. 西安: 西安科技大学, 2020.
- [7] 杨磊. 基于 BP 神经网络的重力异常分离[J]. 工程地球物理学报, 2021, 18(1): 90-97.
- [8] 阮晓钢. 神经计算科学: 在细胞的水平上模拟脑功能[M]. 北京: 国防工业出版社, 2006.
- [9] 姚青华, 邱本花. 基于改进 BP 神经网络的矿井瓦斯浓度预测算法[J]. 煤炭技术, 2017, 36(5): 182-184.
- [10] 雷思友, 刘洁, 方信. 煤与瓦斯突出影响因素综合评价研究: 基于改进熵权法+灰色系统关联法的研究[J]. 煤, 2021, 30(7): 21-24.
- [11] 梁路, 黎剑, 霍颖翔, 等. 一种非均匀分布数据的非线性标准化方法[J]. 计算机科学, 2016, 43(4): 264-269.
- [12] 段梦楠, 刘泽功. 基于神经网络的煤与瓦斯突出规模预测研究[J]. 淮阴工学院学报, 2021, 30(1): 33-39.
- [13] 牟晋娟. 神经网络中激活函数的研究[J]. 电脑编程技巧与维护, 2019(12): 59-61.
- [14] 刘宇晴. 一种新型自适应神经网络激活函数用于深度学习研究[D]. 长春: 吉林大学, 2020.
- [15] 刘航舵. 基于循环神经网络的瓦斯灾害预测预警研究[D]. 西安: 西安科技大学, 2020.
- [16] WILSON A C, ROELOFS R, STERN M, et al. The marginal value of adaptive gradient methods in machine learning[J]. Neural Information Processing Systems, 2017: 3273477.
- [17] REDDI S J, KALE S, KUMAR S. On the convergence of Adam and beyond[J]. ArXiv e-Prints, 2019; arXiv: 1904.09237.